Online Supplement to:
"Immigration, Wages, and Education: A Labor Market Equilibrium Structural Model"

Joan Llull*
Universitat Autònoma de Barcelona, MOVE, and Barcelona GSE

This appendix includes supplementary materials for Llull (2017). More specifically, it includes a more detailed exploration of the patters of immigration in the United States since 1960s; a detailed version of the solution and estimation algorithm; a detailed description of the data construction, variable definitions, and data sources; a robustness exercise that compares the baseline model fit, with that of a version in which the autoregressive parameters from the expectation rule of skill prices are set to zero; and the equilibrium parameters of the expectation rules in each of the counterfactual simulations.

S1. Exploring the U.S. mass immigration: further details

Table 1 of Section I in Llull (2017) shows that the share of immigrants among the less educated increased faster than among any other group during the last four decades. The increase in the share of immigrants among high school dropouts was twice as large as the average increase. In absolute terms, however, this does not mean that immigrants are less educated than four decades ago; instead, it is the result of a slower increase in their education compared to natives. Table S1 shows that the share of immigrants with less than a high school diploma decreased from 49.8 to 27.4 percent, whereas it decreased from 41 to 10.7 percent for natives. An interesting insight from Table S1 is that most of this slower increase in education is driven by the substitution of Western immigrants by Latin Americans and, to a lesser extent, Asians and Africans (see the trends in Figure **??** below). Indeed, if we constructed the counterfactual evolution of the distribution of education for immigrants aggregating the distributions of education by region of origin in each period from Table S1 keeping the distribution of immigrants by region of origin constant to the one in 1970, we would have obtained a distribution of education that would have evolved very similarly to the one for natives.

Another important conclusion from Table 1 of Section I in Llull (2017) is that immigrants are (increasingly) more clustered in blue collar jobs, even conditional on educational levels. This is also true for more disaggregated occupational levels. Table S2 shows that in all categories included in the blue collar aggregate, the

* Departament d'Economia i Història Econòmica. Facultat d'Economia i Empresa. Universitat Autònoma de Barcelona. Campus de Bellaterra – Edifici B, 08193, Bellaterra, Cerdanyola del Vallès, Barcelona (Spain). URL: http://pareto.uab.cat/jllull. E-mail: joan.llull [at] movebarcelona [dot] eu.

Table S1—Education of Natives and Immigrants (%)

	1970	1980	1990	2000	2008
A. Natives					
High school dropouts	41.0	28.2	16.7	12.8	10.7
High school graduates	35.5	38.7	34.8	32.4	37.5
Some college	13.5	18.2	29.0	31.7	26.2
College graduates	10.1	14.8	19.4	23.0	25.6
B. Immigrants					
High school dropouts	49.8	39.0	31.8	30.6	27.4
High school graduates	26.5	27.3	26.2	25.9	28.9
Some college	12.1	16.9	21.8	20.5	17.4
College graduates	11.6	16.8	20.1	23.0	26.3
a. Western Countries					
High school dropouts	49.1	32.2	18.7	11.6	7.7
High school graduates	28.8	33.7	31.2	27.6	29.8
Some college	11.9	17.9	27.1	28.1	24.1
College graduates	10.2	16.3	23.1	32.7	38.4
b. Latin America					
High school dropouts	61.4	56.4	49.4	47.6	42.7
High school graduates	21.8	22.4	25.8	28.1	32.2
Some college	10.0	13.1	16.7	15.7	14.2
College graduates	6.9	8.1	8.2	8.6	10.9
c. Asia and Africa					
High school dropouts	31.5	22.6	16.4	13.2	10.9
High school graduates	22.4	22.8	22.3	21.2	22.6
Some college	16.9	21.5	25.0	23.9	19.6
College graduates	29.2	33.1	36.3	41.7	46.9

Note: Figures indicate the percentage of the population working-age from each region of origin who has the corresponding educational level (columns for each region of origin add to 100%). Immigrants from Western countries include individuals from Canada, Europe and Oceania. *Sources*: Census data (1970-2000) and ACS (2008).

Table S2—Share of Immigrants in each Occupation (%)

	1970	1980	1990	2000	2008
A. Blue-collar	**6.03**	**7.83**	**11.21**	**17.53**	**24.08**
Farm laborers	8.32	14.06	26.08	40.08	51.11
Laborers	5.47	7.40	11.87	21.48	31.27
Service workers	7.58	9.62	13.65	19.58	25.59
Operatives	5.84	8.38	11.74	18.55	23.98
Craftsmen	5.38	6.06	8.16	12.69	18.24
B. White-collar	**4.96**	**5.76**	**7.70**	**10.78**	**13.34**
Professionals	6.29	6.90	8.64	11.95	14.50
Managers	5.02	5.93	7.76	10.75	13.37
Clerical and kindred	4.27	5.17	7.14	9.97	12.47
Sales workers	4.78	5.03	6.78	9.29	11.52
Farm managers	1.52	1.56	2.87	4.87	6.38

Note: Figures indicate the percentage of the workers employed in each occupation who are immigrants. *Sources*: Census data for 1970-2000 and ACS for 2008.

share of immigrants increased faster than the overall share, whereas the opposite is true for all white collar categories. The case of farming-related occupations is very illustrative: farm laborer (blue collar) is the occupation with the largest share of immigrants, whereas farm manager (white collar) is the occupation with fewer immigrants. This finding is in line with the argument of occupation/task specialization of Peri and Sparber (2009). The most important conclusion from Table S2 is that, although sometimes the blue/white collar classification is seen as too broad and heterogeneous (especially for a long period of time), in this case it seems enough to describe the differential supply shock across occupations.

S2. Detailed explanation of the solution and estimation algorithm

The equilibrium model presented in Llull (2017) does not have a closed form solution and needs to be solved numerically. To explain the solution and estimation algorithm, it is convenient to differentiate two types of parameters: *expectation parameters*, Θ_2, which are given by the forecasting rules described in Equation (13), and the process for the aggregate shock (7), and *fundamental parameters* of the model, Θ_1, which are the remaining parameters of the model. Forecasting rules are part of the solution of the model, in the sense that their parameters η^js are implicit functions of the fundamental parameters. Parameters from the aggregate shock process are fundamental by nature, but since the aggregate shock is estimated as a residual (i.e. an implicit function of the data and fundamental parameters), and it is used to forecast future skill prices in the same way forecasting rules given by Equation (13) are used, I treat (and estimate) them as expectation parameters. We can express Θ_2 as $\Theta_2(\Theta_1)$.

Parameters in Θ_1 are estimated by Simulated Minimum Distance. The Simulated Minimum Distance estimator minimizes the distance between a large number of statistics from the data (listed in the paper) and their simulated counterparts. $\Theta_2(\Theta_1)$ is obtained as the fixed point of an algorithm that simulate the behavior of individuals using a guess of Θ_2, and then estimates Equations (7) and (13) from the simulated data to update the guess. Thus, the estimator requires a nested algorithm with a procedure that estimates Θ_1, and another solving Θ_2 given Θ_1.

Lee and Wolpin (2006, 2010) describe a natural nested algorithm in which an inner procedure finds the fixed point in Θ_2 for every guess of Θ_1, and an outer loop solves the Θ_1 estimation problem with a polytope algorithm. The main drawback of this procedure is that it requires solving the fixed point problem in every evaluation of Θ_1, and this increases the computational burden significantly.[1]

[1] This problem is relatively exacerbated if one uses the parallel version of the Simplex Method developed by Lee and Wiswall (2007) in the minimization problem. The basic idea in Lee and Wiswall (2007) is to move the p worst parameters in each Simplex iteration. The problem is

I propose an alternative algorithm that avoids having to solve the fixed point in every iteration of Θ_1. In particular, I propose a swapping of the two procedures which is in the same spirit of the swapping of conditional choice probabilities and parameter estimation proposed by Aguirregabiria and Mira (2002). Θ_1 is estimated for every guess of Θ_2, which is updated at a lower frequency, i.e., I estimate $\Theta_1(\Theta_2)$ for every guess of Θ_2. The algorithm consists of the following steps:

1) Choose a set of parameters $[\Theta_1]^0$ and $[\Theta_2]^0$.[2]

2) Solve the optimization problem for each cohort that exists from $t = 1$ to $t = T$.[3] This dynamic programming problem (given by Equation (1)) is solved recursively by backwards induction from age 65 to age 16. This solution is not analytic. Moreover, the size of the state space is infinite, and even discretizing the continuous variables with a relatively small number of grid points, it still remains impossible to handle. As introduced by Keane and Wolpin (1994, 1997), in each period I solve the problem for a subset of the state space and then I estimate an interpolation rule as a function of the state variables.[4] Unlike in these papers and in Lee and Wolpin (2006, 2010), I use a Gaussian quadrature instead of Monte Carlo integration to numerically compute the multiple-dimensional integrals from the expectation of the value function in $t + 1$.[5]

3) Find the skill rental prices that clear the market and the aggregate shock that closes the production function simulating the economy from $t = 1$ to $t = T$. More specifically:

    a) Guess skill rental prices of period $t = 1$.

that if one of the processors takes more iterations to find the fixed point in $\Theta_2(\Theta_1)$ than all others, the latter will remain idle while the former performs further iterations.

[2] A very natural initial guess for Θ_2 is the one given by the solution of the fixed point algorithm described in Step 5 given the initial guess $[\Theta_1]^0$.

[3] I assume that the economy begins in 1860. This very early initial date is so to overcome the arbitrary initial conditions that I assign to all cohorts existing in $t = 1$. In 1967, the first estimation year, slightly more than two entire generations have gone by. Hence, the oldest individuals (the ones turning 65 that year) have never coexisted in the model with any of the initial cohorts. T is the last estimation year, which is 2007; the youngest individuals that are in the model that year will die in 2057.

[4] The model is solved at 1,280,000 different points of the state space. For each of this points, the expected value function at the selected alternative is obtained —the so-called Emax function. Then, the interpolation rule is estimated as a set of regressions of the log Emax on education, a quadratic in blue collar and white collar experience, the interaction of education, blue collar experience, and white collar experience, predicted skill prices using Equation (13), interactions of these predicted skill prices with education, blue collar experience, and white collar experience, a time trend, number of children, and foreign potential experience. These eighteen regression coefficients are estimated for every age, individual type, and for each of the four alternatives chosen potentially chosen in the preceding period. Hence, the interpolation rule that delivers the Emax at every potential point of the state space consists of 28,224 regression coefficients.

[5] Although being more time consuming, gaussian quadrature is known to be widely more accurate than Monte Carlo integration.

b) Find the supply of skills at this price using the solution of the individual optimization problem obtained in Step 2.

c) Plug the supply of skills into the production function and, together with data on capital and output, recover the aggregate shock as a residual.

d) Update skill rental prices with the demand Equations (9) and (10), using the supply of labor obtained in Step 3b and the aggregate shock from Step 3c.

e) Repeat Steps 3b to 3d using the prices obtained in 3d as the updated guess. Keep iterating to find a fixed point in skill prices. These are the skill prices that clear the market, since they equalize supply and demand.

f) Repeat Steps 3b to 3e for $t = 2, ..., T$.

4) Compare the statistics computed with simulated data and their observed counterparts. Update Θ_1 with a simplex iteration and repeat Steps 2 and 3 with $[\Theta_1]^1$. Keep updating Θ_1 till finding the set of parameters that minimize the distance between simulated and observed data, $\widehat{\Theta}_1([\Theta_2]^0)$.

5) Given $\widehat{\Theta}_1([\Theta_2]^0)$, update $[\Theta_2]^1 = \widehat{\Theta}_2(\widehat{\Theta}_1([\Theta_2]^0))$. In particular, fit OLS regressions of Equations (7) and (13) with the simulated aggregate data obtained at the end of Step 3. Iterate solving Steps 2 and 3 using these OLS estimates to compute expectations until reaching the fixed point.[6]

6) If $[\Theta_2]^1 = [\Theta_2]^0$, the algorithm finishes. Otherwise, repeat Steps 2 to 5 with the updated guesses $[\Theta_2]^1$ and $\widehat{\Theta}_1([\Theta_2]^0)$ until convergence is reached.

S3. Data construction, variable definitions, and data sources

A. Aggregate data

Aggregate macro data are used in the solution of the model, as described in the main text. The estimation period is 1967-2007. However, in order to eliminate the influence of initial conditions, I simulate the model starting in 1860. The model is initialized by simulating the first 40 years (1860-1900) using aggregate data for 1900. Then I simulate the remaining years (1900-2007) with actual macro data. As a result, two entire generations go by before the first year of estimation.

[6] This step can also be done before reaching a convergence in Θ_1 given $[\Theta_2]^0$. Periodic updates of expectation parameters can also be programmed after K iterations. Indeed, if $K = 1$, this algorithm coincides with the one described in Lee and Wolpin (2006, 2010).

Output. Output is measured as Gross Domestic Product at chain 2000 U.S. dollars, provided by the Bureau of Economic Analysis (BEA), NIPA Table 1.1.6. Given that the original series starts in 1929, I use the average annual growth rate (1929-2007) to extrapolate backwards to year 1900.

Capital stock. There are two types of capital in the model: structures and equipment capital. Both series are extracted from BEA, combining flow data from Fixed Assets Tables 1.2 ("Chain-Type Quantity Indexes for Net Stock of Fixed Assets") with year 2000 stock data from Fixed Assets Table 1.1 ("Current Cost Net Stock of Fixed Assets"). Resulting series are expressed in chain 2000 U.S. dollars. Series start in 1925, so I extrapolate them backwards to 1900 using average annual growth rates.

Cohort sizes. Cohort sizes are extracted from Integrated Public Use Microdata Series (IPUMS) of the U.S. Census.[7] In particular, I use information from the decennial Censuses from 1900 to 2000, and from the American Community Survey (ACS) 2001-2007. A person is classified as an immigrant if born abroad; individuals born in Puerto Rico and other outlying areas are categorized as natives. Native and immigrant inter-census cohort sizes are estimated following different procedures. For natives, I distribute the cumulative decade cohort size decrease to each year using annual data on mortality rates by age from Vital Statistics of the U.S. (National Center for Health Statistics). For immigrants, I use a similar procedure, using the estimates of the entry age distribution described below instead of mortality rates.

Age at entry. The distribution of entry age of immigrants is estimated using U.S. Census IPUMS. In order to reduce small sample noise, I average out the distributions for immigrants who arrived at $t-1$, $t-2$,..., $t-5$. Since the exact year of immigration is only available in 1900-1930 and 2000 Censuses, and in the ACS (2001-2007), intermediate years are linearly interpolated. Given that the distribution is stable over the years, I estimate a single distribution for each of the following intervals: 1900-1930, 1931-1940, 1941-1950, 1951-1960, 1961-1970, 1971-1980, 1981-1990 and 1991-2007. Finally, in order to obtain the joint distribution of age at entry and initial education, I estimate the entry age distribution conditional on education. Because of data limitations, I approximate it using the "relative" distribution by educational level, i.e. I compute the ratio of conditional and unconditional distributions from the Census 2000, and then I multiply this relative distribution with the time varying unconditional age at entry distribution.[8]

[7] Ruggles, Sobek, Alexander, Fitch, Goeken, Hall, King and Ronnander (2008).

[8] This calculation assumes that the relative distribution is constant over time. Estimates using 1970-1990 Censuses (for which the year of entry is only available by five-year intervals)

Regions of origin. I consider three regions of origin for immigrants: Western Countries, Latin America, and Asia-Africa. Western Countries include Europe, Canada and Atlantic Islands, and Oceania; Latin America include Caribbean Countries, Mexico, and Central and South America; Asia-Africa includes all immigrants from these two continents. The stock of immigrants from each of these regions are drawn from U.S. Census IPUMS 1900-2000 and ACS 2001-2007. Inter-census estimates of the stock of immigrants from each region of origin are obtained by combining a linear interpolation of the share of immigrants from each region and the estimated of cohort sizes described above. The share of the total inflow of immigrants in year t that comes from region i, $s_{i,t}^{flow}$, is then estimated as:

$$s_{i,t}^{flow} = \frac{M_t s_{i,t} - M_{t-1} s_{i,t-1} + s_{65,i,t-1} M_{65,t-1}}{M_t - M_{t-1} + M_{65,t-1}}, \tag{S1}$$

where M_t is the stock of immigrants in year t, $s_{i,t}$ is the share of immigrants that are natural from region i in period t, and $M_{65,t}$ is the stock of 65 years old immigrants in year t. The share $s_{65,i,t-1}$ is approximated with $s_{i,t-35}$ because the average age at entry is around 30 years. The numerator of Equation (S1) is the flow of immigrants from region i in period t, i.e. the observed increase in the stock plus the recovery of those who died (reached age 65); the denominator is the total inflow in period t.

Initial education. Immigrants and natives are assigned initial years or education differently. Initial education of natives is allocated at age 16. The distribution of years of education at this age (by gender) is estimated using U.S. Census IPUMS for 1940-2000 and ACS 2001-2007. Inter-census estimates are linearly interpolated. In censuses before 1940 there is no information on education. Therefore, I use the 1940 Census to infer the initial education of cohorts aged 16 in each of the previous census years, assuming that they concentrate education at the beginning of their lives, and that mortality at these ages is small enough so that it does not induce any bias. Immigrants are assigned education when they enter the United States. To this end, I use U.S. Census IPUMS for years 1970-2000. I assume that immigrants also concentrate their education spells at the beginning of their life; therefore, an individual with a college degree that enters at age 40 is assumed to enter with the college degree, whereas another that entered at age 18 is assumed to enter with a high school diploma.[9] To impute education to earlier cohorts of immigrants, I estimate the distribution of years of education by cohort of entry using U.S. Census of 1970.

support this assumption.

[9] This is supported by the human capital investment literature (e.g. Becker, 1964).

Fertility process. The fertility process is given by the transition probability matrix from 0, 1, or 2+ preschool children at home in period t into 0, 1 or 2+ in $t+1$, conditional on age, education, and gender. Data are drawn from CPS 1964-2007 and U.S. Census 1900-1960. Before 1960, the transition probability matrix is not conditional on education.

Wage adjustments. To avoid biases in parameter estimates, I make three adjustments to wages and/or aggregate skill units. On the one hand, both CPS wages and output data include taxes, but individuals make decisions on a net income basis; to correct for this, I simulate individuals' decisions using net wages, deflating gross simulated wages (fitted to the data) by the ratio of Disposable Personal Income over Personal Income (Bureau of Economic Analysis, NIPA Table 2.1). On the other hand, there are two reasons why total labor compensation produced by the model could be underestimated without further adjustments (and factor shares, biased as a result): first, the focus on intensive margin, and the use of the year as the time unit, generates some discrepancies between aggregated earnings simulated from the model and actual aggregate earnings (some individuals that work a small fraction of the year are considered as not working, whereas others that work a large fraction of the year, but not the entire period, are assumed to work full time for the whole year); and, second, there are some forms of labor compensation that are not wages (e.g. some types of bonuses and in-kind payments). These two discrepancies are corrected by adjusting total wage compensation appropriately. To correct for the first, I adjust the aggregate simulated wage compensation by the ratio of BEA Total Wage and Salary Disbursements (NIPA Table 2.1) over the aggregation of wages obtained from the CPS. The second concern is addressed using the ratio of BEA Total Wage and Salary Disbursements over the Total Compensation of Employees (NIPA Table 2.1).

B. Microdata

All micro-data statistics used in the estimation are constructed with data from two different sources: the March Supplement of Current Population Survey (CPS), and the two cohorts (1979 and 1997) from the National Longitudinal Survey of Youth (NLSY79 and NLSY97).[10]

[10] CPS data are extracted from IPUMS (Ruggles et al., 2008). The CPS interviews households for 8 eight months: when a household enters the sample for the first time is interviewed four consecutive months, then not interviewed during eight months, and finally interviewed four additional consecutive months (the same four calendar months than in the first spell). Thus, a household in the March sample is interviewed in March for two consecutive years. In most of the survey years it is, hence, possible to match a subset of households for two consecutive years obtaining a small panel. IPUMS data has a recoded individual and household identifier that does not allow to match consecutive surveys. I use samples extracted from the NBER to do the matching. Survey years 1971-72, 1972-73, 1976-77, 1985-86 and 1995-96 can not be matched.

Age groups. Individuals are grouped in ten 5-year age groups from 16-20 years old to 61-65. Individuals above 65 and below 16 are not in the model and they are dropped from the samples.

Educational level. I categorize individuals in four education groups: high school dropouts (<12 years of education), high school graduates (12), persons with some college (13-15), and college graduates (16+). In 1992, a methodological change was introduced to CPS regarding education. Before that year, the education variable gives the respondent's highest grade of school or year of college completed; beginning in 1992, the variable classifies high school graduates according to their highest degree or diploma attained. I use the IPUMS recoded educational attainment variable to make it comparable over the years.

Experience. Years of effective experience in blue collar and white collar occupations are calculated from NLSY. The samples are restricted to individuals born from 1962 to 1964 for NLSY79 and from 1980 to 1984 for the NLSY97. I consider only individuals for which their entire path of choices from age 18 to either 1993, 1992, 1991 or 1990 for NLSY79 or to 2006, 2005 or 2004 for NLSY97 is observable. Individual choices are assigned as described below. Experience is counted as the number of years that the individual's choice was to work in the given occupation.

Choices. Individuals are assigned to one of the four mutually exclusive year round alternatives: blue collar or white collar work, attend school, or stay at home. The procedure to assign individuals follows a hierarchical rule. An individual is assigned to school if she reported that school was her main activity during de survey week (CPS) or if she was attending school at survey date (NLSY). She is assigned to work in either of the two occupations if she is not assigned to school and she worked at least 40 weeks during the year before the survey date, and at least 20 hours per week.[11] When an individual is assigned to work, she is assigned to the occupation held during the last year (CPS) or the most recent (NLSY). Blue collar occupations include craftsmen, operatives, service workers, laborers, and farmers, and white collar include professionals, clerks, sales workers, managers, and farm managerial occupations. Finally, those individuals that are not assigned neither to work nor to attend school are assigned to stay at home.

Wages. Hourly wage is computed for individuals that are assigned to either of the work alternatives according to the previous definition. Workers are assumed to earn their wage entirely in the occupation they are assigned to. Earnings include wage and salary income, and self-employment earnings, deflated to year 2000 U.S.$ using the Consumer Price Index. Top-coded annual earnings are multiplied by 1.4;

[11] Hours per week are approximated by the number of hours worked in the previous week.

extreme observations are dropped, i.e., hourly real wage lower than \$2 or larger than \$200 (e.g., as in Lemieux, 2006). Hours worked are calculated combining information on weeks worked last year and hours worked last week.[12,13]

Preschool children. Individuals are allowed to have 0,1, or 2+ preschool children (less than five years old). In the data, households are defined as family units; preschool children living in a two family home are only assigned to their parents. In order to link children with their parents, I use IPUMS-created variables *momloc* and *poploc*, which identify the position of the mother and father in the household respectively. Parent definition includes biological, step- and adoptive parents. Although they fully comparable over years, there are some minor changes that are listed in the database documentation.

Region of origin. The region of birth is assigned as described above for the aggregate data. A small number of individuals for which the country of birth is unknown are dropped from the corresponding samples. CPS started to ask questions related to immigrant status in survey year 1994. Therefore, statistics that include this information are only used from that survey date onwards.

Potential experience abroad. The initial experience endowment for immigrants (experience obtained abroad) is measured as "potential experience" given data availability. In particular, this variable is defined as age at entry minus years of education, minus 6. In the CPS, year of immigration is only available by intervals; additionally, education is also grouped in 0-4, 5-8, 9, 10, 11, 12, 13-15 and 16+ years of education intervals. To construct experience abroad, I use the central point of the corresponding interval both for age at entry and for years of education. Since I do not observe where did the education take place, I assume that individuals concentrate their education spells in the beginning of their lives, regardless of the country in which they were living. Therefore, if an individual's age at entry minus completed education (and minus 6) is zero or negative, I assume that the individual entered in the U.S. with zero experience. The resulting variable is then grouped into the following categories: 0-2, 3-5, 6-8, 9-11 and 12+ years.

Years in the U.S. This variable is constructed in an analogous way to potential experience abroad. It is also grouped in the same categories.

[12] Before 1976, weeks worked last year are only available by intervals; in particular, the relevant intervals are 40-47, 48-49 and 50-52 weeks. Each interval is imputed, respectively 43.1, 48.3 and 51.9 weeks. These figures are obtained from sample means for each interval using data for the five years after 1975.

[13] In the model, individuals are assumed to work 2080 hours per year (40 hours, 52 weeks). Although hours worked by individuals assigned to working categories average a little above this quantity, there is an important concentration of workers in the amount of 40 hours per week (Keane and Moffitt, 1998).

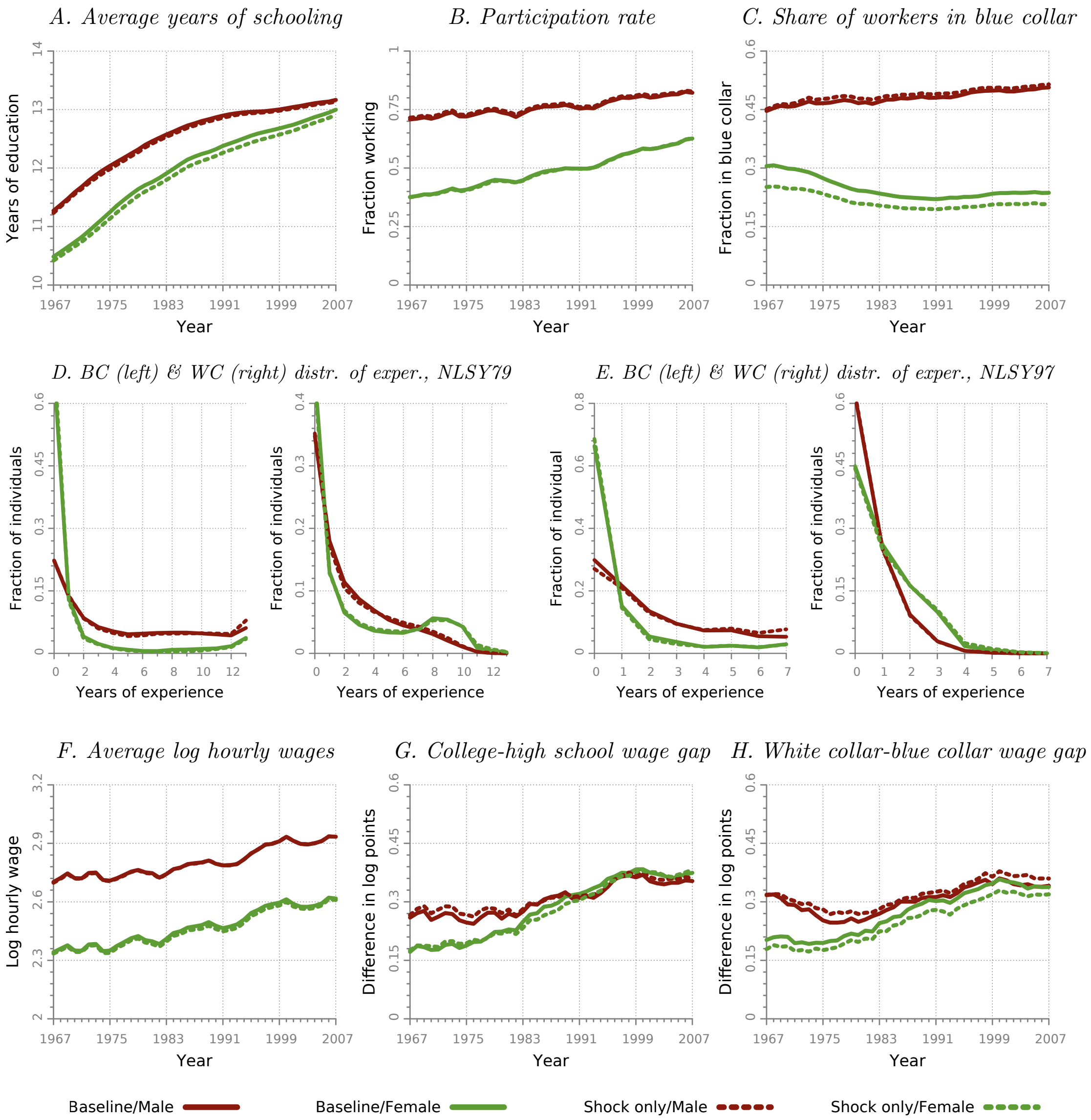
S4. Robustness to the Choice of Expectation Rule: Setting the Autoregressive Coefficients Equal to Zero

Table S3—Expectation Rules for Skill Prices (Autoreg. Coeffs. Set to Zero)

	Blue-collar skill price		White-collar skill price	
Coefficient estimates:				
Constant (η_0)	0.004	(0.002)	0.005	(0.002)
Autoregressive term (η_j)	0.000		0.000	
Δ Aggregate shock (η_z)	0.848	(0.060)	1.122	(0.090)
R-squared goodness of fit measures:				
Differences	0.763		0.717	
Levels	0.999		0.999	
Using predicted shock	0.231		0.224	

Note: The table includes estimates for the coefficients of expectation rules for aggregate skill prices — Equation (13), setting autoregressive coefficients to zero. Goodness of fit measures are reported in the bottom panel. These measures are computed for the prediction of differences and levels for $j = B, W$, where the last one uses the predicted increase in the aggregate shock obtained from Equation (7). Standard errors (in parenthesis) are regression standard errors, and do not account for the error in the estimation of fundamental parameters.

Figure S1. Robustness of Model Fit to Changes in the Expectation Rule



Note: Panels A, B, C, F, G, and H are computed for individuals aged 25-54; actual data for these plots is obtained from March Supplements of the CPS (survey years from 1968 to 2008). In Panels D and E, experience is counted around 1993 (D) and (2006) for individuals in each cohort; sources for actual data in these plots are NLSY79 and NLSY97 as indicated.

S5. Expectation Rules in Counterfactual Scenarios

Table S4—Expectation Rules for Skill Prices In Counterfactuals

	Blue-collar skill price		White-collar skill price	
No capital adjustment ($\partial K/\partial m = 0$):				
Coefficient estimates:				
Constant (η_0)	0.002	(0.001)	0.002	(0.002)
Autoregressive term (η_j)	0.323	(0.047)	0.367	(0.048)
Δ Aggregate shock (η_z)	0.836	(0.046)	1.117	(0.065)
R-squared goodness of fit measures:				
Differences	0.866		0.859	
Levels	0.999		0.999	
Using predicted shock	0.219		0.228	
Full capital adjustment ($\partial r_K/\partial m = 0$):				
Coefficient estimates:				
Constant (η_0)	0.002	(0.001)	0.002	(0.002)
Autoregressive term (η_j)	0.329	(0.046)	0.372	(0.047)
Δ Aggregate shock (η_z)	0.829	(0.045)	1.110	(0.064)
R-squared goodness of fit measures:				
Differences	0.868		0.860	
Levels	0.999		0.999	
Using predicted shock	0.218		0.227	

Note: The table includes estimates for the coefficients of expectation rules for aggregate skill prices —Equation (13)— in counterfactual scenarios. Goodness of fit measures are reported at the bottom of each panel. These measures are computed for the prediction of differences and levels for $j = B, W$. The last one uses the predicted increase in the aggregate shock obtained from Equation (7) instead of the actual increase. Standard errors (in parenthesis) are regression standard errors, and do not account for the error in the estimation of fundamental parameters.

REFERENCES

Aguirregabiria, Victor and Pedro Mira, "Swapping the Nested Fixed Point Algorithm: a Class of Estimators for Discrete Markov Decision Models," *Econometrica*, July 2002, *70* (4), 1519–1543.

Becker, Gary S., *Human Capital. A Theoretical and Empirical Analysis, with Special Reference to Education*, New York: National Bureau of Economic Research, 1964.

Keane, Michael P. and Kenneth I. Wolpin, "The Solution and Estimation of Discrete Choice Dynamic Programming Models by Simulation and Interpolation: Monte Carlo Evidence," *Review of Economics and Statistics*, November 1994, *76* (4), 648–672.

_ **and** _ , "The Career Decisions of Young Men," *Journal of Political Economy*, June 1997, *105* (3), 473–522.

— **and Robert Moffitt**, "A Structural Model of Multiple Welfare Program Participation and Labor Supply," *International Economic Review*, August 1998, *39* (3), 553–89.

Lee, Donghoon and Kenneth I. Wolpin, "Intersectoral Labor Mobility and the Growth of the Service Sector," *Econometrica*, January 2006, *74* (1), 1–46.

— **and** — , "Accounting for Wage and Employment Changes in the US from 1968-2000: A Dynamic Model of Labor Market Equilibrium," *Journal of Econometrics*, May 2010, *156* (1), 68–85.

— **and Matthew Wiswall**, "A Parallel Implementation of the Simplex Function Minimization Routine," *Computational Economics*, September 2007, *30* (2), 171–187.

Lemieux, Thomas, "Increasing Residual Wage Inequality: Composition Effects, Noisy Data, or Rising Demand for Skill?," *American Economic Review*, June 2006, *96* (3), 461–498.

Llull, Joan, "Immigration, Wages, and Education: A Labor Market Equilibrium Structural Model," Barcelona GSE Working papers n. 711, March 2017.

Peri, Giovanni and Chad Sparber, "Task Specialization, Immigration, and Wages," *American Economic Journal: Applied Economics*, July 2009, *1* (3), 135–169.

Ruggles, Steven, Matthew Sobek, Trent Alexander, Catherine A. Fitch, Ronald Goeken, Patricia Kelly Hall, Miriam King, and Chad Ronnander, "Integrated Public Use Microdata Series: Version 4.0," Minneapolis, MN: Minnesota Population Center, 2008.