Llull Cabrer, Joan. «Immigration, wages, and education : a labour market equilibrium structural model». The Review of Economic Studies, Vol. 85 Núm. 3 (2018), p. 1852-1896. DOI 10.1093/restud/rdx053

Postprint of: Review of Economic Studies, 85(3), 1852–1896, July 2018.

Immigration, Wages, and Education: A Labor Market Equilibrium Structural Model

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

This version: June 2017

Recent literature analyzing wage effects of immigration assumes labor supply is fixed across education-experience cells. This paper departs from this assumption estimating a labor market equilibrium dynamic discrete choice model on U.S. micro-data for 1967–2007. Individuals adjust to immigration by changing education, participation, and/or occupation. Adjustments are heterogeneous: 4.2–26.2% of prime-aged native males change their careers; of them, some switch to white collar careers and increase education by about three years; others reduce labor market attachment and reduce education also by about three years. These adjustments mitigate initial effects on wages and inequality. Natives that are more similar to immigrants are the most affected on impact, but also have a larger margin to adjust and differentiate. Adjustments also produce a self-selection bias in the estimation of wage effects at the lower tail of the distribution, which the model corrects.

Keywords: *Immigration, Wages, Human Capital, Labor Supply, Dynamic Discrete Choice, Labor Market Equilibrium*

JEL Codes: *J2, J31, J61.*

During the last forty years, 26 million immigrants of working-age entered the U.S. These immigrants have different skills and work in different occupations than natives, and they changed the composition of the workforce. This change may have affected the skill premium. How do human capital and labor supply decisions react to immigration? Would U.S. natives have spent fewer years in school without

* 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.

§ I would like to thank the editor, Jérôme Adda, and four anonymous referees for thoroughly reading the paper and providing extremely valuable feedback. I am indebted to my thesis adviser, Manuel Arellano, for his encouragement and advice. I am also grateful to Jim Walker for his wonderful sponsorship and comments when I was visiting the University of Wisconsin-Madison. I wish to thank Stéphane Bonhomme, George Borjas, Enzo Cerletti, Sekyu Choi, Giacomo De Giorgi, Juan Dolado, David Dorn, Javier Fernández-Blanco, Jesús Fernández-Huertas Moraga, Christopher Flinn, Carlos González-Aguado, Nils Gottfries, Nezih Guner, Jennifer Hunt, Marcel Jansen, John Kennan, Horacio Larreguy, Sang Yong (Tim) Lee, Pedro Mira, Claudio Michelacci, Robert Miller, Ignacio Monzón, Enrique Moral-Benito, Salvador Navarro, Francesc Obiols, Franco Peracchi, Josep Pijoan-Mas, Roberto Ramos, Pedro Rey-Biel, Robert Sauer, Ricardo Serrano-Padial, Ananth Seshadri, Christopher Taber, Ernesto Villanueva, seminar participants at CEMFI, Bank of Spain, UW-Madison, UAB (Econ), Wash U St. Louis, Bristol, McGill, Uppsala, UPF, Carlos III, Carlo Alberto, IHS-Vienna, Tinbergen, Luca d'Agliano (Milan), Alicante, URV, UAB (Applied Econ), UNav, UdG, UIB, UV, UAM, and conference participants at EALE/SOLE in London, MOOD in Rome, IAB/HWWI in Bamberg, ESWC in Shanghai, EEA in Glasgow, SAEe in Madrid, UCL-Norface in London, IEB Summer School in Barcelona, SED in Cyprus, INSIDE-Norface in Barcelona, and BGSE-Trobada in Barcelona for helpful comments and discussions. Financial support from European Research Council through Starting Grant n. 263600, and the Spanish Ministry of Economy and Competitiveness, through grant ECO2014-59056-JIN, and through the Severo Ochoa Programme for Centers of Excellence in R&D (SEV-2011-0075 and SEV-2015-0563) is gratefully acknowledged.

the massive inflow of foreign workers? Would they have participated more in the labor market? Would they have specialized in different occupations? Providing answers to these questions is crucial to understand the economic consequences of this massive inflow of foreign workers.

Whether and to what extent immigration affected labor market opportunities of native workers has concerned economists and policy makers for years. After an initial strand of the literature exploiting regional differences in immigration, more recent work used cross-skill variation at the national level to identify the effect of immigration on wages.[1] Such analysis considers education-experience cells at a point in time as *closed* labor markets that are differently penetrated by immigrants. As noted by Borjas (2003, p.1337) "the size of the native workforce in each of the skill groups is relatively fixed, so that there is less potential for native flows to contaminate the comparison of outcomes across skill groups". This assumption is present in other papers in the literature (see Card (2009), Ottaviano and Peri (2012), and Llull (2017b) among many others, and Borjas (1999) and Card (2009) for surveys). Even though this cross-skill cell comparison has not brought a consensus on what the effect of immigration on average wages is (which is sensitive to assumptions on elasticities of substitution and on how capital reacts to immigration) most of the papers agree on the existence of asymmetric effects across different workers. As a result, the common assumption of fixed labor supply is not innocuous. Asymmetric effects across different workers change relative wages, and thus generate incentives to adjust human capital and labor supply decisions. Failing to account for these adjustments may lead to a substantial bias in the estimation of wage effects of immigration. Negative effects of immigration on wage levels and inequality would be overstated.

In this paper, I propose and estimate an equilibrium dynamic discrete choice structural model of a labor market with immigration. The model, estimated with U.S. micro-data, is used to identify wage effects of immigration over the last four decades, taking into account labor supply and human capital adjustments by natives and previous generations of immigrants. This approach allows me to address three main points. First, I can quantify and correct the biases in the estimated effects of immigration on wages and on wage inequality introduced by ignoring labor market adjustments to immigration. Second, the dynamic model allows me to identify non-trivial heterogeneous adjustments in education that could not be

[1] Borjas, Freeman and Katz (1992, 1997) introduced the so-called factor proportions approach, which has evolved significantly since then. This methodology compares the supply of workers in a particular skill group to a counterfactual supply in the absence of immigration. Beginning with Card (2001) and Borjas (2003), the elasticities of substitution between different types of labor are estimated. Recent papers implementing this approach on U.S. data include Card (2009), Borjas, Grogger and Hanson (2010), and Ottaviano and Peri (2012) among others.

identified otherwise. And third, I find that labor market detachment produces an additional self-selection bias in the estimation of wage effects of immigration along the native wage distribution, which can be corrected with the model.

The equilibrium framework builds on Altuğ and Miller (1998), Heckman, Lochner and Taber (1998), Lee (2005), and Lee and Wolpin (2006, 2010). The supply side of the model extends the structure of Keane and Wolpin (1994, 1997) to accommodate immigrant and native workers separately. Individuals live from age 16 to 65 and make yearly forward-looking decisions on education, participation and occupation. Immigrants make these decisions as well when they are in the United States. For these immigrants, the model is able to replicate two empirical regularities established in the literature: immigrants downgrade upon entry, that is, they earn lower wages than observationally equivalent natives (Dustmann, Frattini and Preston, 2013); and they assimilate, since between two observationally equivalent immigrants, the one with greater time in the U.S. earns more (LaLonde and Topel, 1992). Human capital accumulates throughout the life-cycle both because of investments in education, and because learning-by-doing on the job leads to accumulation of occupation-specific work experience. In their human capital investment decisions, individuals make forecasts about future wages, which depend on future immigration. Individuals are rational in that they make the best possible forecast given the available information (current labor market conditions and the process describing aggregate uncertainty), but they are unable to perfectly foresee future immigration waves and wages. On the demand side, blue collar and white collar labor is combined with capital to produce a single output. Labor is defined in skill units, which implies that workers have heterogeneous productivity depending on their education, occupation-specific experience, national origin, gender, foreign experience and unobservables. This approach flexibly allows for a continuum of possibilities of imperfect substitution between immigrant and native workers in production. I assume a nested Constant Elasticity of Substitution (CES) production function that accounts for skill-biased technical change through capital-skill complementarity (as in Krusell, Ohanian, Ríos-Rull and Violante, 2000). This is important to correctly estimate native responses to immigration, because it competes with immigration as a source of the increase in wage inequality over the last decades. The equilibrium framework is a crucial feature of the model because it links the immigration-induced supply shift with the changes in incentives for natives through changes in relative wages.

I fit the model to U.S. micro-data data from the Current Population Survey (CPS) and the National Longitudinal Survey of Youth (NLSY) for the period 1967-2007. I then use the estimated parameters to quantify the effect of immigration on labor market outcomes. In order to do so, I define a counterfactual world without large scale immigration in which the immigrant/native ratio is kept constant to 1967 levels. Then, I compare counterfactual wages, human capital, and labor supply with baseline simulations obtained with the estimated parameters.

When I do not allow natives to adjust their human capital and labor supply decisions, results for wages are very much in line with existing papers in the literature, both qualitatively and quantitatively. Overall estimated effects are negative if physical capital does not react to immigration, and virtually zero if capital fully adjusts. Also, the most important effects are on redistribution: less educated, younger, and male individuals are more affected than highly educated, older, and female. When natives and previous generations of immigrants are allowed to respond to immigration by changing their labor supply and human capital decisions, results change in a non-trivial way. Negative effects on wages are mitigated, and redistribution effects are partially arbitraged out, and even reverted in some cases. For example, if capital is not allowed to react to immigration, wages of young male with high school education or less are reduced, on average, by a 4.7% on impact, and those of old college educated female are reduced by a 4%; after human capital and labor supply adjustments take place, wage effects on the former move down to 2.5%, whereas effects on the latter only go down to 3.6%. This is because, even though immigrants are more similar to less educated young men than old educated female, the former have a much larger margin of adjustment (they can increase education, switch occupations, and so on). All this suggests that biases in the estimation of wage effects of immigration are large and ambiguous when labor supply is assumed to be fixed.

In the model, individuals have three adjustment mechanisms: education, occupation, and participation. Results suggest a significant heterogeneity across individuals in optimal reactions to immigration. Between 1.4% and 12.4% of males adjust their education, depending on the assumed counterfactual evolution of capital. Likewise, among those that would work in blue collar jobs in a given cross-section, 1.4% to 4.9% switch to white collar jobs, and 0.6% to 3.4% decide not to work. Overall, 4.2% to 26.2% of the workforce adjust their career paths.

Regarding adjustments on education, the dynamic nature of the model allows me to identify non-trivial heterogeneous responses that have not been identified before in the literature. Some individuals become more likely to pursue a white collar career and, since education is more rewarded there, they extend their stay in school by an average of 3.1–3.2 years.[2] Others, however, become more detached

[2] As a reference, average years of education among individuals aged 25 to 55 increased by about 2.5 years between 1967 and 2007. This suggests that the magnitude of the adjustment is substantial. Nonetheless, this magnitude does not imply that immigration explain the overall increase in the period, as only a small fraction of the population increased their education.

from the labor market and, given the lower return to their investment, they drop out from school earlier, reducing their education by an average of 2.9 years. Which of the two effects dominates is an empirical question, and crucially depends on the assumed counterfactual evolution of capital. When capital fully adjusts, the first channel prevails; the opposite is true if capital does not react to immigration.

Occupation adjustments are also important. As noted above, 1.4% to 4.9% of native male that would work in a blue collar job in a given cross-section switch to a white collar job. This is so because individuals specialize in order to differentiate from immigrants, and they have comparative advantage in white collar jobs.

Finally, the adjustments on participation introduce an additional dimension. Immigration-induced labor market detachment is not randomly distributed over the workforce. Least productive individuals are more likely to drop out from the market as a consequence of immigration. As a result, the comparison of realized wages in baseline and counterfactual scenarios delivers a biased estimate of wage effects of immigration, because of the resulting compositional change (a similar argument to the standard selectivity bias described in Gronau (1974) and Heckman (1979)). The bias is expected to be more severe at the bottom tail of the wage distribution. The structure of the model allows me to identify and correct this bias in a natural way. Potential wages for individuals that decide not to work can be simulated with the estimated model. This allows me to compute the effect of immigration on potential wages along the native distribution comparing the same set of individuals with and without immigration, which avoids the compositional changes that generate the selectivity bias. The comparison of wage effects along the distributions of potential and realized wages allows me to quantify the size of the bias. Results reveal that the bias is only apparent below the median of the native wage distribution, and it is much larger in the scenario without capital adjustment. The size of the bias increases towards the bottom tail. At the 5th percentile, the negative effect of immigration on potential wages of prime-aged native male is 20% to 46% larger than the estimated effect on realized wages, depending on the assumption about counterfactual capital, and 235% to 275% larger in the case of female.

This paper contributes to the extensive literature on labor market effects of immigration. It highlights the importance of relaxing the assumption of fixed labor supply in models like those in Borjas (2003), Card (2009), Ottaviano and Peri (2012), or Llull (2017b). As in these papers, the size of average wage effects depends mostly on the assumption about the counterfactual evolution of capital, and the effects on relative wages are more important than on average wages. However, it contributes by showing that estimates differ substantially depending on whether labor market equilibrium adjustments are accounted for or not. It also departs from these papers in three additional dimensions. The first one is that the production function used here allows for capital-skill complementarity and skill-biased technical change. Lewis (2011) highlights the importance of capital-skill complementarity when analyzing the effect of immigration on wages across local labor markets. Heckman et al. (1998) explore immigration and skill-biased technical change as competing sources for the increase in wage inequality. In the present equilibrium framework, it is very important to include them to avoid biases in the estimation of labor supply responses to immigration. The second one is that it introduces the occupational dimension, which gives micro-structure to the imperfect substitution between natives and immigrants discussed in Ottaviano and Peri (2012) and Manacorda, Manning and Wadsworth (2012) —specialization was already hinted as a potential source for this imperfect substitution in Peri and Sparber (2009) and Ottaviano and Peri (2012). In my model, the extent to which observationally equivalent immigrants and natives are imperfect substitutes in production is endogenously determined by the different choices they make in equilibrium. And the third one is that it departs from the classification of skills in terms of education-experience cells. Dustmann et al. (2013) introduced a more flexible definition of skills: the position along the native wage distribution. The measurement of aggregate labor supply in this paper has a similar flavor.

Another strand of the literature estimates the wage effect of immigration comparing wages across different local labor markets. In that literature, Card (2001), Borjas (2006), and more recently Piyapromdee (2015), model spatial equilibrium responses to immigration in a static framework.[3] Card (2001) additionally introduces occupations and participation to determine how natives and immigrants are competing in the same location. The role of internal migration in these papers is analogous to the one of human capital and labor supply adjustments here: it is the mechanism by which wage effects of immigration in some labor markets are arbitraged out towards the (initially unaffected) others in equilibrium. Piyapromdee (2015) simulates the economy with and without equilibrium adjustments, analogously to what I do below, and also finds that the effects on impact are initially negative, and then they are mitigated by equilibrium adjustments.

A few recent papers use the spatial approach to estimate the effect of immigration on related outcomes like schooling, task specialization, and employment. Hunt (2016) finds that native children, especially native blacks, increase their high school completion rates in order to avoid subsequent competition by unskilled immigrants in the labor market. Peri and Sparber (2009) provide evidence

[3]Kennan and Walker (2010, 2011) and Llull and Miller (2016) on the one hand, and Lessem (2015) on the other pose dynamic models of internal and international migration respectively, with a focus on the migration decision itself rather than on its impact on receiving markets.

that native individuals specialize in language intensive occupations as immigrants have comparative advantage in manual intensive tasks. And Smith (2012) finds that immigration of low educated workers led to a substantial reduction in employment, particularly severe for native youth.

The rest of the paper is organized as follows: Section I provides some descriptive facts about U.S. immigration; Section II presents the labor market equilibrium structural model with immigration; Section III presents the data and discusses identification, solution, and estimation of the model; in Section IV, I present parameter estimates and some exercises that evaluate the goodness of fit of the model; Section V presents the results from the counterfactual simulations, which quantify the labor market effects of immigration. And Section VI concludes.

I. Exploring U.S. mass immigration

According to Census data, the U.S. labor force was enlarged by about 26 millions of working-age immigrants during the last four decades, an increase of almost 0.7 millions per year. This section aims to compare the evolution of the skill composition of immigrants with that of natives, and to establish some correlations between immigration, schooling, and occupational choice. These facts serve as a motivation for the modeling decisions taken in subsequent sections.

Table 1 presents the evolution of the share of immigrants in different subpopulations of the workforce from 1970 to 2008. The share of immigrants among working-age individuals increased from 5.7% to 16.6%. The skill and occupational composition also changed substantially. The share of immigrants among the least educated increased faster than for any other group (6.8% to 33.7%). And immigrants are increasingly more clustered in blue collar jobs: the share of immigrants among blue collar workers increased from 6% to 24% (much more than the overall increase from 5.7% to 16.6%) and, conditional on education, the share of immigrants among dropout blue collar workers increased from 7.2% to 55.5% (compared to the overall increase from 6.8% to 33.7%). In sum, immigrants are increasingly less educated than natives, and they tend to cluster more in blue collar jobs, even conditional on education.

Further exploration of these facts —available at Llull (2017a)— show three additional conclusions. First, the decrease in the relative education of immigrants compared to natives is due to their slower increase educational attainment, not to a decrease in absolute terms. Second, most of it can be explained by the change in the national origin composition of immigrants. And, third, the increasing clustering in blue collar occupations occurred for all two-digit occupations included in the group. In particular, the share of immigrants among farm laborers, among laborers, among service workers, among operatives, and among craftsmen (blue

Table 1—Share of Immigrants in the Population (%)

	1970	1980	1990	2000	2008
A. Working-age population	**5.70**	**7.13**	**10.27**	**14.62**	**16.56**
B. By education:					
High school dropouts	6.84	9.60	17.93	29.02	33.73
High school graduates	4.32	5.14	7.94	12.04	13.27
Some college	5.14	6.63	7.92	9.96	11.65
College graduates	6.48	8.02	10.60	14.59	16.92
C. In blue collar jobs:					
All education levels	*6.03*	*7.83*	*11.21*	*17.53*	*24.07*
High school dropouts	7.18	12.18	23.75	41.03	55.45
High school graduates	4.19	4.94	7.57	12.47	17.30
Some college	5.95	6.14	7.26	9.82	14.07
College graduates	9.53	9.52	12.14	17.89	23.82

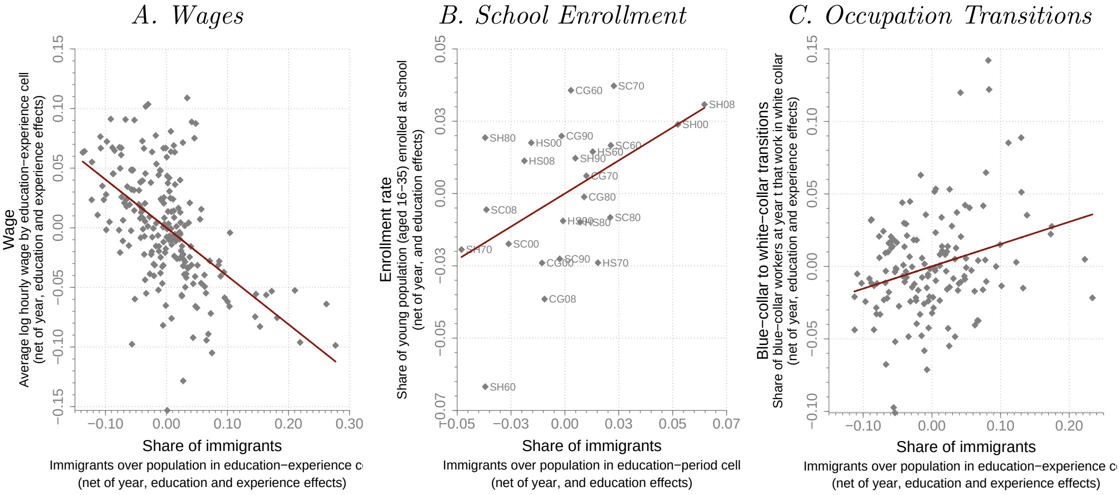
Note: Figures in each panel indicate respectively the percentage of immigrants in the population working-age, in the pool of individuals with each educational level, and among blue collar workers. *Sources*: Census data (1970-2000) and ACS (2008).

collar) increased more rapidly than among professionals, among managers, among clerical and kindred, among sales workers, and among farm managers (white collar). This suggests that the blue collar/white collar classification used in the model below captures reasonably well the differential increase in labor market competition introduced by the new immigrant inflows.

Borjas (2003, Secs. II-VI) compares immigration and wages in different skill cells, defined by education and (potential) experience. He considers four education groups and eight experience categories, defining cells that are then treated as closed labor markets. As immigration varies across skill groups, he uses this variation to identify the effect of immigration on wages in regressions that include different combinations of fixed effects. With this approach, he finds a sizable negative correlation between immigration and wages. I replicate his results using 1960-2000 Censuses and 2008 ACS in Panel A from Figure 1. The figure shows that the correlation between the share of immigrants and the average wage of native males in a cell (net of fixed effects) is negative. In particular, a one percentage point increase in the share of immigrants is associated with a 0.41 (s.e. 0.044) percent decrease in average hourly wages.

Given the research question of this paper, it is worthwhile to look at the correlation between immigration and education. Panel B in Figure 1 compares school enrollment rates and immigrant shares, following an analogous approach to the one described for wages. In particular, I correlate the share of immigrants in a particular education group with enrollment rates of individuals aged 16-35 who exactly achieved that educational level (net of education and time fixed effects). The intuition behind this exercise is as follows: an individual who has just com-

Figure 1. The Correlation of Immigration with Wages, School Enrollment, and Occupational Choice



Note: *Left:* Each point relates average log hourly wage and immigrant share in a given education-experience-year cell. Immigrant shares and average wages are computed for full time workers (20+ hours per week, 40+ weeks per year) aged 16 to 65. Both wages and immigrant shares are net of education, experience, and period fixed effects. The line shows the fitted regression line, with an estimated slope of -0.405 (0.044). *Center:* Each point relates the enrollment rate of individuals with a given completed level of education in a given year and the share of immigrants in that education-year cell. Immigrant shares and enrollment rates are computed with a sample of individuals aged 16-35. Both enrollment rates and immigrant shares are net of education and period fixed effects. The fitted regression line has an estimated slope of 0.458 (0.125). *Right:* Each point relates the fraction of individuals working in blue collar that transit to white collar in the next year and the immigrant share in a given education-experience-year cell. Immigrant shares and transition probabilities are computed with a sample of full time workers aged 16 to 65. Both transition probabilities and immigrant shares are net of education, experience, and period fixed effects. The fitted regression line has an estimated slope of 0.153 (0.045). *General Notes:* Education: high school dropouts, high school graduates, some college, and college graduates); potential experience (age minus education): 9 five-year groups; years: 1960, 1970, 1980, 1990, 2000, and 2008. *Sources:* Census, ACS, and matched March supplements of CPS.

pleted, say, high school, will decide whether to enroll for one additional year or not depending on how tough the labor market competition for high school graduates is. The figure suggests a positive correlation. Specifically, a one percentage point increase in the share of immigrants in a particular group is associated with a 0.46 (s.e. 0.125) points increase in the enrollment rate at that educational level.

Older natives or those who already left education are less likely to go back to school to differentiate themselves from immigrants. A more natural mechanism for them is switching occupations. Panel C in Figure 1 is suggestive of the extent to which this is observed in the data. In this graph, immigrant shares in education-experience cells are related to one year blue collar to white collar transition probabilities in an analogous way to Panel A. The fitted regression suggests that a percentage point increase in the share of immigrants in a cell is associated with a 0.15 (s.e. 0.045) percentage points increase in the one year blue collar to white collar transition probability. This effect is sizable, as it suggests that the increase in immigration of the last decades would explain more than a 10% of the observed increase in blue collar to white collar transitions. The result is indicative of the importance of taking occupational choice into account in the analysis.

The correlations presented in Panels B and C from Figure 1 are suggestive of labor market adjustments to immigration in terms of human capital and labor supply. Career paths and human capital investments are forward-looking decisions that are difficult to assess through reduced form approaches. For this reason, the model below describes the behavior of forward-looking agents making such decisions, within an equilibrium framework that links immigration and labor supply decisions of natives and previous immigrants through changes in relative wages.

II. A labor market equilibrium model with immigration

In this section, I present a labor market equilibrium model with immigration. The model, estimated with U.S. data, is used to quantify the effect of the last four decades of immigration on wages, accounting for human capital and labor supply adjustments by natives and previous generations of immigrants. This approach departs from the literature in that it models the labor supply and human capital decisions explicitly, instead of assuming that labor supply is fixed. It also takes into account skill-biased technical change (considered as an alternative hypothesis for the increase in wage dispersion in the U.S. in recent decades).

A. Career decisions and the labor supply

Native individuals enter in the model at age $a = 16$, and immigrants do so upon arrival in the United States. Both natives and immigrants make yearly decisions until the age of 65 when they die with certainty.[4] Each year, they choose among four mutually exclusive alternatives to maximize their lifetime expected utility. The alternatives are: to work in a blue collar job, $d_a = B$; or in a white collar job, $d_a = W$; to attend school, $d_a = S$; or to stay at home, $d_a = H$.

The decision to migrate to the U.S. is specified outside of the model. Identifying individual migration decisions requires observing immigrants in their home country and in the U.S., and additional micro data on stayers in all countries of origin. There are no data sets that I am aware of that contain all this information. And, even if there were, the extension of the model in this direction would be computationally unfeasible (see Kennan and Walker, 2011; Llull and Miller, 2016). However, in estimation I allow the total inflow of immigrants and the distribution of characteristics with which they enter into the United States to endogenously adjust to U.S. aggregate conditions (aggregate productivity shocks, wage rates,

[4] I abstract from outmigration in this paper. Lessem (2015) shows the importance of circular and return migration in the case of Mexican immigrants. Lubotsky (2007) discusses the implications of selective outmigration in estimating wage profiles for immigrants.

labor supply, and so on).[5] This is so because no orthogonality condition is placed between aggregate quantities and the aggregate shock, as discussed below.

There are L types of individuals that differ in skill endowments and preferences. These types are defined based on national origin and gender. I assume $L = 8$ (male and female for four regions of origin: the United States, Western countries, Latin America, and Asia/Africa). National origin is important for three reasons. First, as noted in Section I, the evolution of the national origin composition of immigration can explain most of the evolution in the educational level of immigrants over the last decades. Second, there are important differences in wages, distribution of occupations, and labor market participation among immigrants coming from these different origins. And, third, these regions differ substantially in the incidence of illegal migration to the United States.

Individual types are, hence, based on observable characteristics. Introducing permanent or persistent unobserved heterogeneity is unfeasible for two reasons. The first one is computational tractability. The second is identification. In particular, since the decision to migrate would be endogenous to these unobservables (Borjas, 1987), identification of their distribution would require modeling individual migration decisions, which is not feasible as discussed above.

At every point in time t, an individual i of type l and age a solves the following dynamic programming problem:

$$V_{a,t,l}(\Omega_{a,t}) = \max_{d_a} U_{a,l}(\Omega_{a,t}, d_a) + \beta\, \mathbb{E}\left[V_{a+1,t+1,l}(\Omega_{a+1,t+1}) \mid \Omega_{a,t}, d_a, l\right], \qquad (1)$$

where $\mathbb{E}[.]$ indicates expectation, β is the subjective discount factor, and Ω_a is the state vector.[6] The list of variables included in Ω_a, as well as the way in which individuals form expectations about future Ω is discussed in Section II.C. The terminal value is $V_{65+1,t,l} = 0\ \forall l, t$. The instantaneous utility function is choice-specific, $U_{a,l}(\Omega_{a,t}, d_a = j) \equiv U^j_{a,t,l}$ for $j = B, W, S, H$. Workers have a linear utility. They are not allowed to save and, hence, they are not able to smooth consumption.

Working utilities are given by:

$$U^j_{a,t,l} = w^j_{a,t,l} + \delta^{BW}_g\, \mathbb{1}\{d_{a-1} = H\} \quad j = B, W, \qquad (2)$$

where $w^j_{a,t,l}$ are individual wages in occupation $j = B, W$, $\mathbb{1}\{A\}$ is an indicator function that takes the value of one if condition A is satisfied and zero otherwise, and δ^{BW}_g is a gender-specific labor market reentry cost that workers pay to get a

[5] Llull (2017b) finds that endogeneity of immigration across skill groups is a major concern. Yet, most papers in the literature (looking at wage effects at the national level using variation across skill cells) assume that immigration is exogenous (e.g. Borjas, 2003).

[6] For notational simplicity, I omit the individual subindex i, which should be present in all individual-specific variables throughout the paper.

job if they were not working (and not in school) in the previous period.[7]

Workers are paid their marginal product. A base rate r_t^j is determined in equilibrium, as discussed below. Individual wages are scaled by the relative productivity of the individual, through a factor $s_{a,l}^j$ that depends on individual characteristics and independent and identically distributed idiosyncratic shocks. Hence, wages $w_{t,a,l}^j \equiv r_t^j \times s_{a,l}^j$ are defined by a fairly standard Mincer equation (Mincer, 1974):

$$w_{a,t,l}^j = r_t^j \exp\{\omega_{0,l}^j + \omega_{1,is}^j E_a + \omega_2^j X_{Ba} + \omega_3^j X_{Ba}^2 + \omega_4^j X_{Wa} + \omega_5^j X_{Wa}^2 + \omega_6^j X_{Fa} + \varepsilon_a^j\}, \quad (3)$$

where:

$$\begin{pmatrix} \varepsilon_a^B \\ \varepsilon_a^W \end{pmatrix} \sim i.i.\mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_{Bg}^2 & \rho_{BW}\sigma_{Bg}\sigma_{Wg} \\ \rho_{BW}\sigma_{Bg}\sigma_{Wg} & \sigma_{Wg}^2 \end{bmatrix} \right).$$

The exponential part of Equation (3) defines an expression for $s_{a,l}^j$ (individual skill units), as a function of individual type l, education E (with $is = nat, immig$), blue collar and white collar domestic experience in the country X_B and X_W, potential experience abroad X_F (age at entry minus education), and a random shock, ε^j, with gender-specific variance σ_{jg}^2 and (gender-invariant) correlation across occupations ρ_{BW}. Idiosyncratic shocks are assumed to be independently and identically distributed across individuals, and uncorrelated with individual and aggregate characteristics. When working in occupation j, individuals accumulate one additional year of occupation j-specific experience, $X_{ja+1} = X_{ja} + \mathbb{1}\{d_a = j\}$, which has a return in the future.

Wages have been modeled extensively in the literature using Mincer equations (see Heckman, Lochner and Todd (2006) for a review). These have been proved to fit life-cycle earnings profiles reasonably well. The Mincer equation approximates the framework of human capital accumulation on the job in Ben-Porath (1967). As noted by Heckman et al. (2006, p. 317), the Mincer equation is consistent with a linearly declining rate of investment on-the-job, which implies that the log-wage is a quadratic function of experience. Formal education is introduced in the model as a special case in which all available time is devoted to skill accumulation. For this reason, education enters linearly in (the log of) Equation (3). Equation (3) also accounts for different rates of human capital accumulation in the different occupations, introducing separate returns to the experience obtained in each.

Observationally equivalent natives and immigrants supply different amounts of skills for several reasons. First, they have different intercepts $\omega_{0,l}^j$, which capture non-fully-portable region-specific skills (e.g. language and culture), and also regional differences in other dimensions, like the prevalence of illegal immigration.

[7] I assume that transitions from school to work are costless. New immigrants pay the reentry cost in all cases except if enrolled in school in the home country in the previous period. These assumptions appear to be the most parsimonious way to fit the relevant transitions in the data.

Second, their returns to education, $\omega^j_{1,is}$, may differ, because immigrants may undertake (at least a part of) their education abroad. As schooling abroad is not necessarily oriented towards the U.S. labor market, foreign education could map into lower wages compared to the education obtained in the United States (e.g. learning Chinese calligraphy vs English grammar).[8] And third, while abroad, immigrants accumulate foreign instead of domestic experience, which potentially have different returns.

These differences generate a good fit of choices and wages of immigrants in the data, which is important to correctly quantify the magnitude of the immigration shock. Moreover, it is also important because they generate two important regularities established in the literature. The first one is *downgrading* of immigrants upon arrival in the United States. Dustmann et al. (2013) define downgrading as the situation in which the position occupied by immigrants along the native wage distribution is below the one they would occupy based on observables. The second regularity is immigrant *assimilation*. LaLonde and Topel (1992) define assimilation as the process whereby, between two observationally equivalent immigrants, the one with greater time in the U.S. earns more. According to this definition, immigrants assimilate as they accumulate some skills in the U.S. that they would not have accumulated in their home country (Borjas, 1999), which in the model is generated by a different (larger) return to domestic compared to foreign experience.[9]

Individuals who decide to attend school face a monetary cost, which is different for undergraduate (τ_1), and graduate students ($\tau_1 + \tau_2$). Additionally, they get a non-pecuniary utility with a permanent component $\delta^S_{0,l}$, a disutility of going back to school if they were not in school in previous period $\delta^S_{1,g}$, and an i.i.d. transitory shock ε^S_a, normally distributed with gender-specific variance σ^2_{Sg}:

$$U^S_{a,l} = \delta^S_{0,l} - \delta^S_{1,g}\mathbb{1}\{d_{a-1} \neq S\} - \tau_1\mathbb{1}\{E_a \geq 12\} - \tau_2\mathbb{1}\{E_a \geq 16\} + \varepsilon^S_a. \tag{4}$$

As a counterpart, they increase their education, $E_{a+1} = E_a + \mathbb{1}\{d_a = S\}$, which provides a return in the future.

Finally, individuals remaining at home enjoy non-pecuniary utility given by:

$$U^H_{a,t,l} = \delta^H_{0,l} + \delta^H_{1,g}n_a + \delta^H_{2,g}t + \varepsilon^H_a. \tag{5}$$

[8] Ideally, I would allow the return to the education *obtained* in the U.S. and abroad to differ. However, the country in which individuals undertake their education is not observable in the data. Therefore, I assume that the return to *all* education is different for natives and immigrants. Likewise, returns to experience are not allowed to vary by national origin or immigrant status because of other data limitations highlighted in Section III.

[9] In line with this, Eckstein and Weiss (2004), using data for Israel, find that foreign experience is almost unvalued upon arrival, and that conditional convergence takes place as the immigrant keeps accumulating local experience.

In this case, on top of its permanent and transitory components $\delta^H_{0,l}$ and ε^H_a (normally distributed with gender-specific variance σ^2_{Hg}), the utility is increased by a gender-specific amount $\delta^H_{1,g}$ for each preschool children living at home, n_a.[10] Finally, the home utility includes gender-specific trend $\delta^H_{2,g}t$. The linear utility assumption implies no income effect in the labor force participation decision, and, hence, participation is driven only by the substitution effect. In a framework with growing wages, everyone would eventually work in the long run. A linear trend in the home utility is a reduced form way of avoiding this problem.

To sum up, the main trade-offs that define the labor supply problem are as follows. Individuals decide whether to enjoy home utility, invest in education, or work in one of the two available occupations. Even though one can enjoy it ($\delta^S_{0,l} + \varepsilon^S_a$ can be positive), attending school typically entails a contemporaneous cost. In return, education provides higher wages in the future. Since returns differ across occupations, the education decision is an important determinant of future career path, and the expected future path will also influence the decision to obtain education. When working, individuals are paid a wage, and they accumulate work experience, which maps into future wages. Forward-looking individuals could be interested in an occupation that pays a lower contemporaneous wage if the experience provides a high enough return in the future. Occupation decisions, hence, affect and are affected by future career prospects. Finally, the wage rate r^j_t is an equilibrium outcome. It channels the effect of immigration towards native choices and wages. If immigrants have comparative advantage in blue collar jobs, immigration may put negative pressure on the blue collar rate, which generates incentives for natives to switch to white collar careers. Likewise, it may also reduce general wage levels, which can make the home option more attractive.

B. *Aggregate production function and the demand for labor*

The economy is represented by an aggregate firm that produces a single output, Y_t, combining labor (blue collar and white collar aggregate skill units, S_{Bt} and S_{Wt}) and capital (structures and equipment, K_{St} and K_{Et}) using a Constant Elasticity of Substitution (CES) technology described by the following production function:

$$Y_t = z_t K^{\lambda}_{St}\{\alpha S^{\rho}_{Bt} + (1-\alpha)[\theta S^{\gamma}_{Wt} + (1-\theta)K^{\gamma}_{Et}]^{\rho/\gamma}\}^{(1-\lambda)/\rho}. \tag{6}$$

Equation (6) is a Cobb-Douglas combination of structures and a composite of labor and equipment capital. This composite is a CES aggregate of blue collar

[10] The variable n_a is assumed to take one of the following values: 0, 1 or 2 (the latter for 2 or more children). Fertility is exogenous to individual shocks. The transition probability matrix for the number of children is taken from the data, it is potentially correlated with aggregate shocks, and it depends on gender, education, age and cohort.

labor and another CES aggregate, which combines equipment capital and white collar labor. Parameters α, θ, and λ are connected to the factor shares, and ρ and γ are related to the elasticities of substitution between the different inputs. The elasticity of substitution between equipment capital and white collar labor is given by $1/(1-\gamma)$, and the elasticity of substitution between equipment capital or white collar labor and blue collar labor is $1/(1-\rho)$. Neutral technological progress is provided by the aggregate productivity shock z_t, whose evolution is described by:

$$\ln z_{t+1} - \ln z_t = \phi_0 + \phi_1(\ln z_t - \ln z_{t-1}) + \varepsilon^z_{t+1}, \quad \varepsilon^z_{t+1} \sim \mathcal{N}(0, \sigma^2_z). \tag{7}$$

This process allows for a linear trend in levels, with slope ϕ_0, and business cycle fluctuations around it. The aggregate shock is assumed to be independent of idiosyncratic shocks, but it is allowed to be correlated with aggregate supplies of capital and labor (including immigrant inflows and the distribution of characteristics with which immigrants enter into the U.S.). Skill units are supplied by workers according to the exponential part of Equation (3). Even though, as noted above, I abstract from explicitly modeling individual migration decisions, in estimation I assume that immigrant inflows (and its distribution of skills) are determined endogenously following a known process (independent of idiosyncratic shocks but endogenous to aggregate fluctuations). Likewise, I also abstract from modeling individual savings decisions, and I proceed analogously with the aggregate capital supply.[11] In the counterfactual experiments I simulate alternative scenarios with different assumptions about these processes, as discussed below.

Equation (6) is somewhat different from the three-level nested CES proposed by Card and Lemieux (2001) and popularized in the immigration literature by Borjas (2003).[12] In particular, it differs in three aspects: (i) it allows for capital-skill complementarity and skill-biased technical change, (ii) it takes occupations into account and (iii) instead of classifying individuals in skill cells based on educational level and age, it defines skills in a flexible way, accounting for observed and unobserved heterogeneity, in a similar spirit as in Dustmann et al. (2013).

Capital-skill complementarity is important to account for skill-biased technical change. Krusell et al. (2000) show that, technical change reduced the relative

[11] These known processes are irrelevant for estimation because: (i) observed immigrant inflows and capital stocks are assumed to reflect equilibrium values; (ii) as discussed in Section III.A, identification is achieved from individual wages and choices, which avoids imposing orthogonality conditions between the aggregate productivity shock z_t and capital stocks or immigration processes; and (iii) the approximation to rational expectations described in Section II.C does not require solving the model for counterfactual values of capital.

[12] Borjas (2003) specifies a production function that is a Cobb-Douglas combination of capital and a labor aggregate; this labor aggregate is a CES combination of labor in four education cells, each being defined a CES aggregate of workers over eight experience cells. Ottaviano and Peri (2012) experiment with alternative nesting structures, and also define a CES combination of native and immigrant workers within each skill cell.

price of equipment capital dramatically starting in early 1970s. Using a production function that resembles the one in Equation (6), they find that this technical change is skill-biased because $\rho > \gamma$ (meaning that equipment capital is more complementary to skilled labor than to unskilled labor). In particular, the increasing speed of accumulation of equipment capital increases the demand for white collar workers. These authors find that this mechanism alone can explain most of the variation in the skill premium over the subsequent decades. In an equilibrium framework, not accounting for the increase in the demand of white collar workers induced by the accumulation of equipment capital would lead to an overestimation of the reaction of natives to the inflow of immigrant blue collar workers.

Allowing for different occupations in the production function is also important. Natives and immigrants may be imperfect substitutes in production because their different skills may lead them to different choices of occupations (Ottaviano and Peri, 2012, p. 175).[13] The evidence provided by Ottaviano and Peri (2012) suggests that it is important to account for this imperfect substitution. In the present paper, occupational choice endogenizes the extent to which this imperfect substitution between natives and immigrants shows in the data. Additionally, occupation specialization is an important adjustment mechanism employed by natives to react to the labor market competition induced by immigrants (Peri and Sparber, 2009). Finally, occupational switching is also an important determinant of the increase in wage inequality (Kambourov and Manovskii, 2009).[14]

Finally, Dustmann et al. (2013) discuss the importance of defining skill groups in a flexible way, departing from the skill-cell approach in Borjas (2003) and Ottaviano and Peri (2012). In particular these authors note that immigrants downgrade upon entry into the destination country. As a result, they do not compete with the natives that share the same observable skills, but, instead, with those that work in the same jobs. Skill units, defined in Equation (3), determine, together with occupation, a more accurate measurement of labor market competition. They also generate further wage heterogeneity, which allows to quantify heterogeneous effects along the native wage distribution. And, importantly, despite all the extra flexibility, this approach is more parsimonious.

[13] In a partial equilibrium framework, introducing occupations may be misleading as the supplies in each occupation are clearly endogenous objects. In the spatial approach, Card (2001) did so using an instrumental variables approach. Finding an instrument in the skill-cell approach seems more difficult.

[14] In my model, I only allow for two occupations: blue collar and white collar. Adding extra occupations is very costly. Each additional occupation implies an extra choice, an additional experience variable in the state space, more than 20 additional parameters to estimate, and an additional skill price to solve in equilibrium. As noted in Section I, the share of immigrants in all two-digit occupations included in the blue collar group increased by more than the share of immigrants in any of the two-digit occupations included in the white collar group. Thus, this classification seems enough to capture the differential supply shift generated by immigration.

C. Expectations

In order to make their decisions, individuals need to forecast the future path of the state variables, including future skill prices. The state vector at year t, $\Omega_{a,t}$, includes the following state variables: age, education, blue collar and white collar effective work experience, foreign potential experience, previous year decision, calendar year, number of children, idiosyncratic shocks, current skill prices, and the necessary information to forecast future skill prices. The first eight evolve deterministically given choices. Thus, workers only face uncertainty about future skill prices, number of children, and idiosyncratic shocks.

Let $F(n_{a+1}, \varepsilon_{a+1}, r_t^B, r_t^W | \Omega_{a,t})$ denote the distribution of these variables in the next period conditional on the current state, with $\varepsilon_a \equiv (\varepsilon_a^B, \varepsilon_a^W, \varepsilon_a^S, \varepsilon_a^H)'$. I assume:

$$F(n_{a+1}, \varepsilon_{a+1}, r_t^B, r_t^W | \Omega_{a,t}) = F^\varepsilon(\varepsilon_{a+1}) F^n(n_{a+1} | n_a, E_a, a, t) F^r(r_t^B, r_t^W | \Omega_{a,t}). \quad (8)$$

Equation (8) implies that the processes for the idiosyncratic shock, number of children, and skill prices are independent. As noted above, $F^\varepsilon(.)$ is a multivariate normal with gender-specific parameters $\Sigma_g \equiv (\sigma_{Bg}, \sigma_{Wg}, \sigma_{Sg}, \sigma_{Hg}, \rho_{BW})'$, independent of individual-specific and aggregate variables. The assumption of independence with respect to aggregate supplies of skills and skill prices is consistent with assuming that individuals are atomistic (Altuğ and Miller, 1998). The fertility process $F^n(.)$ is conditional on education, age, calendar year, and current number of children, and it is independent of any other state variable, including current and future idiosyncratic and aggregate shocks (conditional on calendar year).

Forecasting skill prices is more complicated. Future skill prices depend on future aggregate supplies of labor and capital, and on the aggregate productivity shock. The process for the aggregate shock is described by Equation (7). Future capital stocks in equilibrium depend on the future aggregate shock and labor supplies, and on the (unspecified) capital supply process. Future labor supply depends on future aggregate shock, capital stocks, and cohort sizes (including the future stock of immigrants), and on the future distribution of individual-specific state variables in the economy (Krusell and Smith, 1998). Under rational expectations, $F^r(.)$ should be such that individuals make the best possible forecast conditional on the available information in period t. Thus, to specify $F^r(.)$, one should specify a process for all the above, including the endogenous responses of immigration and capital stocks to the aggregate shock and to labor supply, and, importantly, the entire distribution of individual state variables in the economy. This is unfeasible. Alternatively, to make the problem tractable, I follow an approach that combines the approximation algorithm in Krusell and Smith (1998) and the framework in Altuğ and Miller (1998), in the same spirit as Lee and Wolpin (2006, 2010). Specifically, I approximate future skill prices by the following autoregressive process:

$$\Delta \ln r_{t+1}^{j} = \eta_0^{j} + \eta_j \Delta \ln r_t^{j} + \eta_z^{j} \Delta \ln z_{t+1}. \tag{9}$$

This rule is a good approximation to rational expectations if the parameter vector $\eta \equiv (\eta_0^B, \eta_B, \eta_z^B, \eta_0^W, \eta_W, \eta_z^W)'$ is such that Equation (9) provides a good fit to the process for skill prices. As shown in Section IV, this seems to be the case: conditional on z_{t+1}, the process explains 99.9% of the variation in skill prices. Providing an almost perfect fit, however, does not mean that individuals perfectly foresee future skill prices, because z_{it+1} is not observable at time t.

Equation (9) is a reduced form of the model structure that individuals use to predict future skill prices. Thus, the vector η is not really a vector of parameters, but, instead, an implicit function of the *fundamental* parameters of the model, and, hence, part of the solution. The fact that Equation (9) provides an extremely good fit of the process for skill prices indicates that current skill prices and especially the evolution of the aggregate shock are (almost) sufficient statistics to predict future skill prices for a given η.[15] This is reasonable given that all aggregate processes are assumed to endogenously react to the aggregate productivity shock. For example, expectations about future immigration and its effect future wages are determined by expectations about the evolution of the aggregate shock (which determines future levels of immigration directly and indirectly through equilibrium adjustments), and its mapping into future wages (which includes wage effects of immigration). If a positive aggregate productivity shock is expected to lead an increase unskilled immigration, η_z^B could be relatively small compared to η_z^W (if unskilled immigration drives blue collar relative wages down). Likewise, an unconditional expectation of low skilled immigration in the future could imply that η_0^B is relatively small compared to η_0^W (again, if unskilled immigration puts downward pressure to blue collar wages).

D. *Equilibrium*

The market structure of this economy is as follows. Supplies of capital and immigrants are given by processes determined outside of the model that endogenously react to aggregate conditions in the economy, but that are independent of individual unobservable characteristics.[16] The aggregate supply of skill units in

[15] A counterfactual evolution of state variables would determine a different vector η.

[16] An assumption that would be consistent with this framework is that workers and capitalists are different groups of individuals, so that workers are not allowed to save, but their decisions affect interest rates through aggregate labor supply, and capitalists are not allowed to work, but their decisions affect wages through aggregate capital supply. For immigration, this assumption requires that unobserved idiosyncratic shocks are not persistent, as noted in Section II.A.

occupation $j = B, W$ is given by:

$$S_{jt} = \sum_{a=16}^{65} \sum_{i=1}^{N_{a,t}} s^j_{a,i} \mathbb{1}\{d_{a,i} = j\}, \tag{10}$$

where $N_{a,t}$ is the cohort size. The aggregate demands are derived from the aggregate firm's profit maximization, which equalizes marginal returns to rental prices. In particular, labor demands are given by:

$$r^B_t = (1-\lambda)\alpha \left(\frac{S_{Bt}}{KBW_t}\right)^\rho \frac{Y_t}{S_{Bt}}, \tag{11}$$

for blue collar skill units, where $KBW_t \equiv \{\alpha S^\rho_{Bt} + (1-\alpha)[\theta S^\gamma_{Wt} + (1-\theta)K^\gamma_{Et}]^{\rho/\gamma}\}^{1/\rho}$ is the CES aggregate labor and equipment capital in Equation (6) and:

$$r^W_t = (1-\lambda)(1-\alpha)\theta \left(\frac{KW_t}{KBW_t}\right)^\rho \left(\frac{S_{Wt}}{KW_t}\right)^\gamma \frac{Y_t}{S_{Wt}}, \tag{12}$$

for white collar skill units, where $KW_t \equiv [\theta S^\gamma_{Wt} + (1-\theta)K^\gamma_{Et}]^{1/\gamma}$ is the equipment capital-white collar labor CES aggregate. Demands of structures and equipment capital are given by analogous expressions. The equilibrium is given by market clearing conditions. Equilibrium prices r^B_t, r^W_t, r^S_t and r^E_t (the last two are the rental prices of structures and equipment capital respectively) are such that supply and demand of immigration, of skill units in the United States, and of capital are equalized. Empirically, baseline levels of immigration and capital in equilibrium are observed in the data. In counterfactuals, immigration levels are implied by the design of the policy experiment, and different scenarios for capital supply adjustment are simulated, as noted in Section V. Baseline and counterfactual labor supplies are obtained by solving the equilibrium.

Given equations (11) and (12) we can write the (log of the) relative white collar to blue collar skill price as:

$$\ln \frac{r^W_t}{r^B_t} = \ln \frac{(1-\alpha)\theta}{\alpha} + (\rho - 1)\ln \frac{S_{Wt}}{S_{Bt}} + (\rho - \gamma)\ln \frac{KW_t}{S_{Wt}}. \tag{13}$$

Equation (13) can be interpreted as a reformulation of Tinbergen's race between technology and the supply of skills (Tinbergen, 1975).[17] The second term of this equation is the negative contribution of the relative supply of skills (if $\rho < 1$) and the last term captures the biased technical change through the increase in the speed of accumulation of equipment capital (whenever $\rho > \gamma$). Immigration

[17] Tinbergen (1975) suggests that the overall change in the gap between skilled and unskilled wages is driven by two contrasting forces: the relative increase in the supply of skills, which tends to close the gap, and a skill-biased technical change, which opens it. Acemoglu (2002), and Acemoglu and Autor (2011) survey the literature that tests this hypothesis.

changes the relative supplies of skills. Not allowing for capital-skill complementarity (imposing $\gamma = \rho$) would put all the burden of the increase in the relative wages observed in the last decades to the change in relative labor supplies. Since immigration pushed S_{Wt}/S_{Bt} steadily down over the last decades coinciding with a period of important skilled-biased technical change, wrongly imposing $\gamma = \rho$ would induce a negative bias in the estimation of $(\rho - 1)$, leading an over-prediction of the effect of immigration on relative wages.

III. Identification, data, and estimation

This section gives an overview of the main identification arguments, describes the most important features of data construction, and introduces a sketch of the algorithm used for the solution and estimation of the model. A more thorough discussion of model identification is presented in Appendix A. Additionally, detailed descriptions of the solution/estimation algorithm, and of variable definitions and sample selection are available in the Online Supplement (Llull, 2017a).

A. Identification

The following discussion builds on previous work by Hotz and Miller (1993), Altuğ and Miller (1998), Magnac and Thesmar (2002), Arcidiacono and Miller (2011, 2015), and Kristensen, Nesheim and de Paula (2015). The main assumption exploited for identification is that the idiosyncratic shocks are independent of all other state variables. Identification also relies on the additional assumption that conditional choice probabilities (CCPs) are identified nonparametrically from observed decisions in the data. The latter is not trivial in practice, because the aggregate shock and the skill prices, which are not observable, are state variables, and because only partitions of the state vector are included in each of the datasets used in estimation. To simplify the argument, I assume hereinafter that CCPs are identified (identification of the CCPs is discussed in Appendix A).

The wage equations are identified following standard arguments in the self-selection literature (Heckman, 1974, 1979). In particular, we can use a control function approach that corrects for the bias induced by the fact that the disturbance is not zero-mean conditional on $d_a = j$. This control function is a mapping on the CCPs (Heckman and Robb, 1986; Hotz and Miller, 1993). Given the parametric assumption for the distribution of ε_a, this mapping is known. But even if it was not, it would be nonparametrically identified since the model provides exclusion restrictions, like the fact that the number of children affects participation but not wages (Ahn and Powell, 1993; Das, Newey and Vella, 2003). Aggregate skill prices, which are not observable, are identified as the coefficients of calendar time dummies. This requires a normalization of one of the intercepts, $w^j_{0,l}$ for some l, in each wage equation; I normalize native-male intercepts in both equations to zero.

Identification of skill prices leads to identification of the production function parameters and the aggregate shock. Individual skill units are identified as $s^j_{a,i} = \exp\{\ln w^j_{a,t,l} - \ln r^j_t\}$, and aggregate skill units are identified integrating $s^j_{a,i}$ over the sample of individuals working in occupation j. Using data on output Y_t and capital K_{Et} and K_{St}, the production function parameters are identified from the first order conditions of the firm's problem —if at least three periods are available— without imposing any additional orthogonality condition (see Appendix A). Given them, aggregate shocks $\{z_t\}^T_{t=t_0}$ are identified as residuals in Equation (6), and the AR(1) coefficients for the shock process (ϕ_0 and ϕ_1 in Equation (7)) are identified as regression coefficients. Likewise, combining $\{z_t\}^T_{t=t_0}$ and $\{r^B_t, r^W_t\}^T_{t=t_0}$, the equilibrium value of η is identified from Equation (9).

The identification of the remaining parameters of the model follows standard arguments in the literature. I fix the discount factor β, which is proved to be identified only through the functional form assumptions of the model (Magnac and Thesmar, 2002). The parameters that remain to be identified are δ^{BW}_g, $\delta^S_{0,l}$, $\delta^S_{1,g}$, τ_1, τ_2, σ_{Sg}, $\delta^H_{0,l}$, $\delta^H_{1,g}$, $\delta^H_{2,g}$, and σ_{Hg}. Proposition 1 in Hotz and Miller (1993) establishes that the mapping between value functions and CCPs can be inverted so that we can express continuation values as a function of the CCPs. Kristensen et al. (2015) prove that this result still holds in the case in which utility functions do not satisfy additive separability, as it is the case here. Identification of the dynamic model is thus provided by observed choices and CCPs for future periods. Even though the discrete choices are made based on the difference between the utility obtained from each alternative and the one obtained from a base alternative, there is no need for further normalizations. The reason is that δ^{BW}_g is common for blue collar and white collar alternatives, and that the parameters from the wage equation, including variance-covariance parameters, have already been identified above, which is ultimately the result of observing wages.

Three remarks about identification are worth highlighting. First, the parameters of the production function are identified from the variation in the micro data on wages and choices. This permits identification without imposing orthogonality conditions between aggregate variables and the aggregate shock, allowing capital stocks and immigration to be endogenous to aggregate fluctuations.[18,19] If spec-

[18] Nonlinear least squares estimation of the production function (or of the labor demand equations) would impose orthogonality conditions between the residual (aggregate shock) and the observed supplies of capital and labor, which would be violated if immigration or capital stocks are endogenous. Instead, identifying the production function parameters from individual wages and choices as I do here avoids this problem and allows me to recover consistent estimates.

[19] An additional advantage of this approach is that it avoids the incidental parameters problem in the estimation of the production function when the total number of calendar years T is fixed. Yet, this problem is still present in the estimation of the parameters of the expectation rules

ified, the endogenous processes for the supply of immigrants and capital would predict equilibrium quantities as a function of the aggregate shock. Observed capital and immigrant stocks are assumed to be realizations of these equilibrium outcomes in the baseline economy, and thus, they are sufficient statistics for the processes that generated them. As a result, identification is achieved without specifying these processes. However, observed values of the same aggregates are no longer equilibrium values in counterfactual scenarios. Thus, counterfactual simulations require additional assumptions about these processes as discussed below.

Second, permanent (or even persistent) unobserved heterogeneity cannot be identified in this model because immigrants are not observed in their home countries prior to migration. As discussed in Aguirregabiria and Mira (2010, p. 55), in the model with unobserved heterogeneity, the initial condition would potentially be endogenous because it would not be independent of the individual's unobservable. More specifically, there would be a self-selection of who migrates and when based on these unobservables (Borjas, 1987). In order to recover the parameters of such model one would need to specify migration decisions which, as discussed above, is not feasible. Alternatively, one could try to identify the distribution of types for each possible value of observables at entry, which include education, age, and region of origin. However, that would increase the computational burden and the parameter space so much that it would also be unfeasible.

And, third, education decisions (and individual decisions in general) are identified in the model through the exclusion restrictions, the observation of wages conditional on education, and the conditional choice probabilities, which embed individuals' expectations about future returns to education. Identification relies on the assumption that the current level of education is uncorrelated with the idiosyncratic shock. It also depends on expectations about future equilibrium wages, which are approximated by the rule presented above and identified with quite unrestrictive assumptions. As it is shown below, the model is able to replicate the observed evolution of individual choices, distribution of state variables, one-year transitions, wage profiles, and returns to education, which is reassuring.

B. Model solution and estimation

For estimation, it is convenient to differentiate two types of parameters: *expectation parameters*, Θ_2, which include the forecasting rule described in Equation (9), and the process for the aggregate shock in Equation (7), and the *fundamental parameters* of the model, Θ_1, which include the remaining parameters. Parameters

and the aggregate shock process: they remain an estimation problem even if N tends to infinite. An analogous situation occurs in Lee and Wolpin (2006, 2010). Altuğ and Miller (1998) avoid this problem by using the Euler condition for consumption and information on assets.

Θ_2 are part of the solution of the model, and hence can be expressed as $\Theta_2(\Theta_1)$.[20]

Parameters Θ_1 are estimated by Simulated Minimum Distance (SMD). The SMD estimator minimizes the distance between a set of statistics obtained from micro-data, listed in Section III.C, and their counterparts predicted by the model. The expectation parameters $\Theta_2(\Theta_1)$ are obtained as a fixed point in an algorithm that obtains equilibrium skill prices and aggregate shocks simulating the behavior of individuals who form their expectations using a guess of Θ_2, and updates guess fitting Equations (7) and (9) to the simulated data. Thus, the estimation process requires a nested algorithm that estimates Θ_1, and solves for Θ_2 given Θ_1.

Lee and Wolpin (2006, 2010) describe a nested algorithm with an inner procedure that finds the fixed point in Θ_2 for every guess of Θ_1, and an outer loop that finds Θ_1 using a polytope minimization algorithm. The main problem with this procedure is that it requires solving the fixed point problem in every evaluation of Θ_1, and this increases the computational burden significantly.[21] Alternatively, I propose an algorithm that avoids this problem by swapping the two procedures, in the spirit of Aguirregabiria and Mira (2002). In particular, Θ_1 is estimated for every guess of Θ_2, which is updated at a lower frequency. In other words, I estimate $\Theta_1(\Theta_2)$ for every guess of Θ_2. This algorithm is described in detail in the Online Supplement (Llull, 2017a).

C. Data

To estimate the model I fit 27,636 statistics computed with micro-data from 1967 to 2007 obtained from the March Supplement of the Current Population Survey (CPS), and the two cohorts of the National Longitudinal Survey of Youth (NLSY79 and NLSY97). These statistics, listed in Table 2, include information on choice probabilities, one-year transitions, distributions of education and experience, and mean, log-first difference, and variance of wages, all this conditional on observable characteristics. Additionally, aggregate data on output and the stocks of structures and equipment capital from the Bureau of Economic Analysis, and on cohort sizes (by gender and immigrant status), distribution of entry age for immigrants, distribution of initial schooling (at age 16 for natives and at entry for immigrants), and distribution of immigrants by region of birth from the Census are used in the solution and estimation. The transition probability process for fertility (number of preschool children) is directly estimated from observed tran-

[20] Parameters from the aggregate shock process are fundamental parameters, but since the aggregate shock is estimated as a residual, which requires recovering equilibrium supplies of skills, they are obtained from the solution of the model, and are treated as expectation parameters.

[21] This problem is exacerbated in parallelized minimization routines (like the algorithm by Lee and Wiswall (2007), used here). In particular, if one of the threads takes more iterations to find the fixed point $\Theta_2(\Theta_1)$, the other threads remain idle waiting for it to converge.

TABLE 2—DATA

Group of statistics	Source		Number of statistics
TOTAL			**27,636**
Proportion of individuals choosing each alternative..			**5,074**
By year, sex, and 5-year age group	CPS	$41 \times 2 \times 10 \times (4-1)$	2,460
By year, sex, and educational level	CPS	$41 \times 2 \times 4 \times (4-1)$	984
By year, sex, and preschool children	CPS	$41 \times 2 \times 3 \times (4-1)$	738
By year, sex, and region of origin	CPS	$15 \times 2 \times 4 \times (4-1)$	360
Immigr., by year, sex, and foreign potential exp.	CPS	$15 \times 2 \times 5 \times (4-1)$	450
By sex and experience in each occupation	NLSY	$2 \times (5 \times 5 + 4 \times 4) \times (2-1)$	82
Wages:			**6,044**
Mean log hourly real wage...			*3,000*
By year, sex, 5-year age group, and occupation	CPS	$41 \times 2 \times 10 \times 2$	1,640
By year, sex, educational level, and occupation	CPS	$41 \times 2 \times 4 \times 2$	656
By year, sex, region of origin, and occupation	CPS	$15 \times 2 \times 4 \times 2$	240
Immigrants, by year, sex, fpx, and occupation	CPS	$15 \times 2 \times 5 \times 2$	300
By sex, experience in each occupation, and occ.	NLSY	$2 \times (5 \times 5 + 4 \times 4) \times 2$	164
Mean 1-year growth rates in log hourly real wage...			*2,148*
By year, sex, previous, and current occupation	CPS§	$35 \times 2 \times 2 \times 2$	280
By year, sex, 5-year age group, and current occ.	CPS§	$35 \times 2 \times 10 \times 2$	1,400
By year, sex, region of origin, and current occ.	CPS§	$13 \times 2 \times 4 \times 2$	208
Immigr., by year, sex, years in the U.S., and occ.	CPS§	$13 \times 2 \times 5 \times 2$	260
Variance in the log hourly real wages...			*896*
By year, sex, educational level, and occupation	CPS	$41 \times 2 \times 4 \times 2$	656
By year, sex, region of origin, and occupation	CPS	$15 \times 2 \times 4 \times 2$	240
Career transitions...			**12,138**
By year and sex	CPS§	$35 \times 2 \times 4 \times (4-1)$	840
By year, sex, and age	CPS§	$35 \times 2 \times 10 \times 4 \times (4-1)$	8,400
By year, sex, and region of origin	CPS§	$13 \times 2 \times 4 \times 4 \times (4-1)$	1,248
New entrants taking each choice by year and sex	CPS	$15 \times 2 \times (4-1)$	90
Immigrants, by year, sex, and years in the U.S.	CPS§	$13 \times 2 \times 5 \times 4 \times (4-1)$	1,560
Distribution of highest grade completed...			**4,260**
By year, sex, and 5-year age group	CPS	$41 \times 2 \times 10 \times (4-1)$	2,460
By year, sex, 5-year age group, and immigr./native	CPS	$15 \times 2 \times 10 \times 2 \times (4-1)$	1,800
Distribution of experience...			**120**
Blue collar, by sex	NLSY	$2 \times (13+7)$	40
White collar, by sex	NLSY	$2 \times (13+7)$	40
Home, by sex	NLSY	$2 \times (13+7)$	40

Note: Data are drawn from March Supplements of the Current Population Surveys for survey years 1968 to 2008 (CPS), the National Longitudinal Survey of Youth both for 1979 and 1997 cohorts (NLSY), and the CPS matched over two consecutive years —survey years 1971–72, 1972–73, 1976–77, 1985–86 and 1995–96 can not be matched— (CPS§). Statistics from the CPS that distinguish between natives and immigrants can only be computed for surveys from 1994 on. There are 10 five-year age groups (ages 16–65), two genders (male and female), two immigrant status (native and immigrant), four regions of origin (U.S. (natives), Western countries, Latin America, and Asia/Africa), four educational levels (<12,12,13–15 and 16+ years of education), three categories of preschool children living at home (0, 1 and 2+), and five foreign potential experience (fpx) and years in the country groups (0–2,3–5,6–8,9–11 and 12+ years). Redundant statistics that are linear combinations of others (e.g. probabilities add up to one) are not included (neither in the table, nor in the estimation).

sitions in the data (Census and CPS). Data sources, sample construction, and variable definitions are detail in the Online Supplement (Llull, 2017a).

There are four aspects of the data construction that are worth describing in more detail here. First, not all necessary information is contained in a single dataset that can be used in estimation. The CPS has a short panel dimension that allows me to compute transition probabilities, but does not include information on effective experience in blue collar and white collar. The NLSY is a long panel, and allows me to compute effective experience, but it only follows two cohorts over time (which would make the identification of equilibrium quantities and production function parameters less credible) and it is not refreshed with new cohorts of immigrants. The need of combining these two datasets prevents direct estimation of choice probabilities conditional on all the observable state variables at a time, and, hence, the implementation of non-full solution methods (CCP estimation).

Second, individuals need to be assigned to mutually exclusive annual choices. I do so following a similar approach to Lee and Wolpin (2006). Individuals are assigned to school if attending school was their main activity at the time of the survey. Else, they are assigned to work if they worked at least 40 weeks during the year preceding the survey, and at least 20 hours per week. If working, they are assigned an occupation based on the main occupation held in the previous year (CPS) or most recent one (NLSY). Blue collar occupations include craftsmen, operatives, service workers, laborers, and farmers, and the white collar group includes professionals, clerks, sales workers, managers and farm managerial occupations. The remaining individuals are assigned to the home alternative.

Third, it is important to provide some notes about the measurement of immigration. Immigrants are defined as individuals born abroad. In the CPS, immigrants are only identifiable starting in 1993. For this reason, all statistics in Table 2 that are conditional on immigrant status are computed only for the period 1993-2007. This is used below to check the goodness of the model in fitting choices and wages of immigrants for the period before 1993, which constitutes a sort of out-of-the-sample fit check. Data from the CPS and the U.S. Census are assumed to include both legal and undocumented aliens. In fact, these datasets are used by the literature and by policy makers to quantify the importance of illegal immigration using the residual method (Hanson, 2006; Baker and Rytina, 2013). However, it is generally accepted that they undercount the total stock of undocumented immigrants to some extent (Hanson, 2006). Some papers, like Lessem (2015) and others surveyed in Hanson (2006) use the Mexican Migration Project, which includes information about legal status, but their focus is on Mexican migration.

And, fourth, in estimation I compare data statistics with simulated counterparts. Simulated statistics are obtained from simulating the behavior of cohorts of 5,000 natives and 5,000 immigrants (some starting life abroad and only making decisions once in the U.S.). Thus, each simulated cross-section includes up to $(5,000 + 5,000) \times 50 = 500,000$ observations, which are weighted using data on cohort sizes.

IV. Estimation results

A. Parameter estimates

This Section discusses parameter estimates, listed in Tables 3 and 4. I first present estimates for the fundamental parameters of the model, which are those in Equations (2) through Equation (7). Then I present the recovered equilibrium values for η in Equation (9), along with some goodness of fit statistics.

Fundamental parameters of the model. Table 3 presents estimates for the fundamental parameters of the model. Standard errors, in parentheses, account for both sampling and simulation error, as detailed in Appendix C.

Panel A presents estimates for the gender×origin constants for each alternative. There are substantial differences in preferences and productivity between immigrants and natives, and, among immigrants, by national origin. Latin American immigrants have a comparative advantage in blue collar jobs, whereas Asian/African and Western immigrants have comparative advantage in white collar. Given the change in the national origin composition of immigrant inflows in recent decades, these differences can explain the increasing concentration of immigrants in blue collar occupations. Western and Asian/African immigrants also have a stronger preference for education, which makes them more likely to enroll.

Returns to education in blue collar jobs, in Panel B, are smaller for immigrants than for natives (5.8% versus 7.2% per extra year of education), while they are remarkably similar for white collar (10.9% versus 11%). Foreign potential experience is positively rewarded in blue-collar occupations (1.7%) and negatively rewarded in white collar jobs (−5.9%). All these differences make similar natives and immigrants to work in different occupations. Ottaviano and Peri (2012) find that observationally equivalent natives and immigrants are imperfect substitutes in production because they are employed in different occupations. Thus, the model endogenously generates such imperfect substitutability.

These estimates have implications for the ability of the model in predicting different regularities about immigration that have been established in the literature. LaLonde and Topel (1992), Borjas (1985, 1995, 2015), and Lubotsky (2007) for the United States, Dustmann and Preston (2012) for the United Kingdom and the United States, and Eckstein and Weiss (2004) for Israel show that immigrants *assimilate* as they spend time in the destination country. As returns to foreign potential experience are smaller than those to domestic experience both in blue

Table 3—Parameters Estimates

A. Origin×gender constants:	Nat. male	Nat. female	Western countries	Latin America	Asia/ Africa
Blue collar	0	-0.341	0.087	-0.032	-0.021
		(0.0010)	(0.0292)	(0.0169)	(0.0060)
White collar	0	-0.291	0.135	-0.161	0.152
		(0.0015)	(0.0355)	(0.0169)	(0.0140)
School	2,186	5,866	8,291	2,302	28,207
	(68)	(84)	(249)	(864)	(382)
Home	16,420	11,333	16,957	11,259	15,162
	(53)	(29)	(764)	(240)	(143)

B. Wage equations:	Blue collar		White collar	
Education—Natives ($\omega_{1,nat}$)	0.072	(0.0001)	0.110	(0.0001)
Education—Immigr. ($\omega_{1,imm}$)	0.058	(0.0005)	0.109	(0.0004)
BC Experience (ω_2)	0.094	(0.0001)	0.001	(0.0002)
BC Experience squared (ω_3)	-0.0023	(0.00001)	-0.0006	(0.00002)
WC Experience (ω_4)	0.028	(0.0002)	0.106	(0.0002)
WC Experience squared (ω_5)	-0.0013	(0.00001)	-0.0030	(0.00001)
Foreign Experience (ω_6)	0.017	(0.0005)	-0.059	(0.0012)
Variance-covariance matrix of i.i.d. shocks:				
Std. dev. male (σ_{male})	0.452	(0.0059)	0.589	(0.0024)
Std. dev. female (σ_{female})	0.389	(0.0042)	0.476	(0.0033)
Correlation coefficient (ρ_{BW})		0.048	(0.0043)	

C. Utility parameters:	Male		Female	
Labor market reentry cost (δ_1^{BW})	8,968	(77)	12,400	(180)
School utility parameters:				
Undergraduate Tuition (τ_1)		13,841	(85)	
Graduate Tuition ($\tau_1 + \tau_2$)		70,970	(869)	
Disutility of school reentry (δ_1^S)	29,009	(207)	37,357	(597)
Home utility parameters:				
Children (δ_1^H)	-1,799	(47)	3,626	(75)
Trend (δ_2^H)	62.92	(0.83)	53.77	(0.55)
Standard dev. of i.i.d. shocks:				
School (σ^S)	1,150	(8)	215	(2)
Home (σ^H)	10,227	(638)	5,316	(243)

	Elast. of substit. param.		Factor share parameters		
D. Production function:	BC vs Eq. (ρ)	WC vs Eq. (γ)	Struct. (λ)	BC (α)	WC (θ)
	0.288	-0.059	0.073	0.555	0.452
	(0.006)	(0.005)	(0.011)	(0.007)	(0.010)

E. Aggregate shock process:	Constant (ϕ_0)	Autoregressive term (ϕ_1)	St. dev. of innovations (σ_z)
	0.001	0.328	0.026
	(0.003)	(0.114)	(0.021)

Note: The table presents parameter estimates for Equations (2) to (7). Native male constants for wage equations are normalized to zero. Immigrant male and native female constants are estimated. The constant for a female immigrant from region i is obtained as the sum of the constant for a region i male immigrant and the difference between the constant for native females and native males. The individual subjective discount factor, β, is set to 0.95. Standard errors (calculated as described in Appendix C) are in parentheses. Standard errors for the aggregate shock process are regression standard errors.

collar and in white collar jobs, immigrants assimilate in this model, because between two observationally equivalent immigrants, the one that spent more years in the U.S. earns more. Borjas (1985, 1995, 2015) noted that the relative quality of immigrant cohorts decreased in recent years. In the model, the change in the relative importance of Latin American aliens in the recent cohorts fosters immigrant concentration in blue collar jobs, and leads to a decrease in average immigrant productivity, which is consistent with these findings. Finally, Dustmann et al. (2013) and Dustmann and Preston (2012) find evidence of immigrants downgrading upon arrival in the destination country. The comparative advantage of some groups immigrants in blue collar jobs and the differences in productivity with respect to natives generate this phenomenon. This is the case, for example, of Latin American immigrants. Other groups like the highly educated Western immigrants that arrive in the United States straight after attending school in their home country are more likely enroll in school, work in white collar jobs, and make higher wages in the U.S. than comparable natives, which would be a form of *upgrading*.

One of the most important differences between the production function in Equation (6) and the nested-CES production function used in the immigration literature (Borjas, 2003; Ottaviano and Peri, 2012) is that Equation (6) allows for capital-skill complementarity and skill-biased technical change. Elasticities of substitution implied by the estimates of ρ and γ are respectively 1.40 and 0.94, very much in line with Krusell et al. (2000).[22] These elasticities indicate that equipment capital and white collar labor are relative complements. This capital-skill complementarity links the fast accumulation of equipment capital and the increase in the white collar/blue collar wage gap, as shown in Equation (13). Several papers have tested capital-skill complementarity with different data since the seminal work by Griliches (1969). Most of them agree in the existence of some degree of complementarity between capital and skilled labor, even though there is a variety of estimates for the elasticities of substitution (Hamermesh, 1986).

The remaining parameters of the model, which are crucial for the model to fit choices, wages, and transitions observed in the data, are also reasonable and in line with the literature. Women are less productive than men in both occupations (to a larger extent in blue collar), obtain a larger utility from attending school, and a smaller utility from staying at home. This is consistent with the observed wage gap, enrollment rates, and female concentration in white collar occupations. They also obtain a higher boost in home utility when having preschool children,

[22] Results for the other parameters of Equation (6) are also in line with the literature. The estimates of α and θ imply a share of structures of 7.3% (similar to Krusell et al. (2000) and Greenwood, Hercowitz and Krusell (1997)), a roughly constant capital share (Kaldor, 1957), and an increasing importance of white collar labor within the labor share. The aggregate shock process includes a small trend and an important cyclical component.

which fits their larger propensity to drop from the labor market for childbearing. And their reentry costs both to labor market and to school are larger than male counterparts, which, together with a lower variance in the home decision, makes them less likely to transit in and out from non-employment (compared to men).

Estimated returns to education fit within the variety of results surveyed by Card (1999), which range from 5 to 15% with most of the estimated causal effects clustering between 9% and 11%. Results are also qualitatively in line with (although somewhat larger than) Keane and Wolpin (1997), Lee (2005), and Lee and Wolpin (2006). Compared to own experience, returns to cross experience are much lower, flatter, and turn negative after a certain level. Along with the positive correlation between blue collar and white collar idiosyncratic shocks, this is important to fit observed transitions across occupations. This leaves some degrees of freedom for the variance of the idiosyncratic shock to capture the observed variance in wages. Likewise, the estimated cost for reentering school is quite large ($29,009$US$ and $37,357$US$ for male and female), which is in line with the observation that very few individuals return to school after leaving it. And the estimated labor market reentry cost is close to nine thousand US$ for males, and above twelve thousand for females, one quarter and almost one half of the average full-time equivalent annual wage for males and females respectively. All this provides a good fit of wages and transitions across alternatives, as discussed in the next section. Yet, in reality there could be some permanent unobserved factor (ability or taste heterogeneity), not included in the model, that makes individuals more likely to persist in their choices. Omitting such heterogeneity could lead to an overstatement of the other factors that drive persistence in this model. In terms of the estimated parameters, this would potentially imply underestimating cross-experience effects, overestimating transition costs, and/or overestimating returns to own experience. Furthermore, if these unobservables include ability, and high ability individuals are more likely to educate and to earn higher wages for a given educational level, this omission could also induce an overestimation of the returns to education.

Expectation rules. Table 4 presents the equilibrium values of η in Equation (9). The growth rate of the aggregate shock seems the most important piece to explain variation in skill prices. White collar skill prices react more to shocks than blue collar prices. Estimates also show some state-dependence, and a small positive trend that adds to the one included in the aggregate shock. The selection of these particular rules as an approximation to rational expectations balanced a trade-off between simplicity and goodness of the approximation.[23] The bottom panel of the

[23] In the Online Appendix (Llull, 2017a), I check the sensitivity of the results to restricting the autoregressive coefficients to zero. Predictions of the model are quite similar. The restricted approximation rule still explains more than 70% of the variation of first differences in skill prices.

Table 4—Expectation Rules for Skill Prices

	Blue-collar skill price		White-collar skill price	
Coefficient estimates:				
Constant (η_0)	0.002	(0.001)	0.002	(0.002)
Autoregressive term (η_j)	0.324	(0.046)	0.367	(0.048)
Δ Aggregate shock (η_z)	0.835	(0.046)	1.118	(0.065)
R-squared goodness of fit measures:				
Differences	0.870		0.858	
Levels	0.999		0.999	
Using predicted shock	0.221		0.222	

Note: The table includes estimates for the coefficients of expectation rules for aggregate skill prices —Equation (9). Goodness of fit measures are reported in the bottom panel. These measures are computed for the prediction of differences and levels for $j = B, W$. 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.

table includes three different R^2 measures that summarize the explanatory power of these rules. The first one, a standard R^2 for the model in differences, indicates that the rules are able to predict more than 85% of the variation in growth rates of skill prices. The second one measures the goodness of the rules in fitting the variation in levels, displaying almost a perfect fit. This large explanatory power, however, does not imply that individuals have perfect foresight of future skill prices, as they do not observe z_{t+1} in period t. Accounting for that, the third measure replaces $\Delta \ln z_{t+1}$ by its prediction from Equation (7). Results suggest that individuals are only able to forecast around a 22% of the variation in (the growth rate of) skill prices one period ahead, which is far from perfect foresight.

B. Model fit

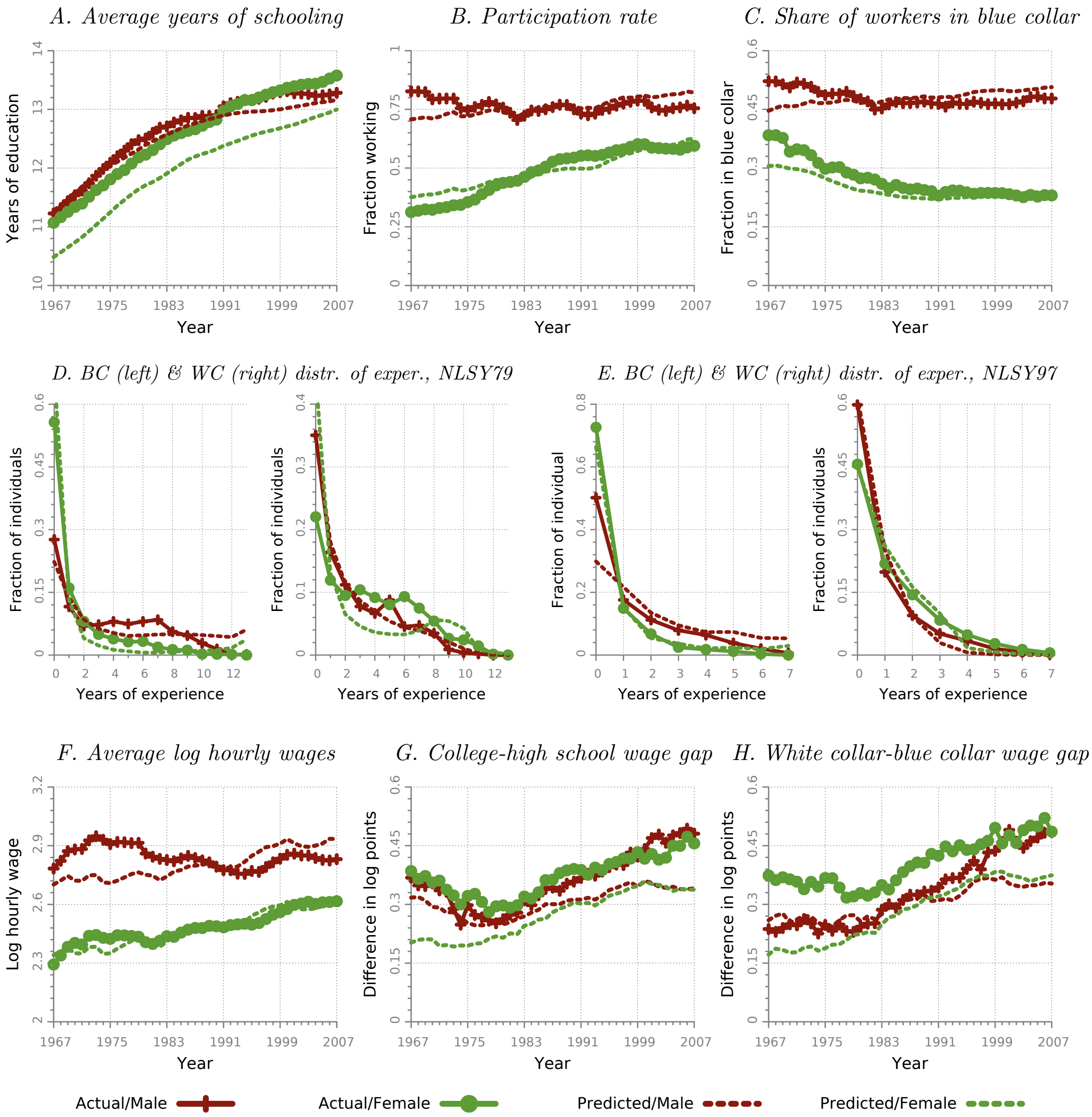
In this section, I compare predicted and actual values of the most relevant aggregates for individuals aged 25 to 54 in order to evaluate the goodness of fit of the estimated model. I focus on this age range because it is the one for which I compare baseline and counterfactual outcomes in Section V.

Figure 2 compares actual and predicted aggregate statistics, both for male and female. Panel A includes average years of schooling.[24] The model predicts well both the level and the change in education over the sample period for males. For females, it also fits very accurately the increase observed in the data (around 2.5 years), but somewhat under-predicts the level (by around a third of a year).[25]

[24] To calculate them, I average the following values: 0 for no education, preschool or kindergarten, 2.5 for 1st to 4rth grade, 6.5 for 5th to 8th grade, 9, 10, 11, and 12 for the corresponding grades, 14 for some college, and 16 for bachelors degree or more.

[25] This under-prediction of female education goes along with an underestimation of the college to high school and white collar to blue collar wage gaps (Panels G and H). More generally, the structure of the model has some difficulty in fitting the similarity of male and female in

Figure 2. Actual and Predicted Aggregates



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.

---

wage gaps and education along with the substantial difference in labor market participation and occupational choice. For example, increasing $\omega^W_{0female}$ would increase the white collar-blue collar wage gap and, most likely, education and the college-high school wage gap, but it would also increase female participation and the fraction of female working in white collar, which are currently well matched. The model would have a better chance in fitting female aggregates if fertility decisions were modeled endogenously, and/or if I modeled endogenous household formation. In the current model, if a female is hit by a "children shock", she may dropout out from school because of the increased utility of the home alternative when a preschool child is present. Later on, when the child is not preschool anymore, this female would face an extremely large cost of reentering into education, so she might decide not to go back to school. In reality, prospective high educated women may endogenously postpone their fertility decision (e.g., see Adda, Dustmann and Stevens, 2017, for a discussion). Regarding endogenous household formation, many papers have documented assortative mating on different labor-market-relevant traits (e.g. Greenwood, Guner, Kocharkov and Santos, 2014; Guner, Kulikova and Llull, 2017), and

Table 5—Actual vs Predicted Transition Probability Matrix

	Choice in t							
	Blue collar		White collar		School		Home	
Choice in $t-1$	Act.	Pred.	Act.	Pred.	Act.	Pred.	Act.	Pred.
Blue collar	0.75	0.77	0.11	0.10	0.00	0.00	0.14	0.13
White collar	0.06	0.07	0.83	0.83	0.00	0.00	0.10	0.10
Home	0.11	0.08	0.13	0.13	0.01	0.01	0.76	0.79

Note: The table includes actual and predicted one-year transition probability matrix from blue collar, white collar, and home (rows) into blue collar, white collar, school, and home (columns) for individuals aged 25-54. Actual and predicted probabilities in each row add up to one. Actual data is obtained from one-year matched March Supplements of the CPS (survey years from 1968 to 2008).

Panels B and C compare actual and predicted labor force participation and fraction of employees working in blue collar occupations respectively. The model fit of these dimensions is remarkable. It accurately reproduces the participation level, the increase in female labor force participation, the fraction of individuals working in blue collar occupations, and the gender gaps in the two variables. Panels D and E evaluate the goodness of the model in fitting the distribution of experience in the NLSY samples. For individuals in the NLSY79 (Panel D), experience is measured around 1993, when individuals are aged around 30. For the NLSY97 sample (Panel E), it is measured around 2006, with individuals aged around 25. In general, the model provides a good fit of these distributions. Panels F through H show the model fit for wages. The model displays a remarkable fit of female average wages (trend and level), the level of male wages and, hence, the gender gap, the college to high school (except the trend in the last few years, and the level for women in early years), and the trend and level of white collar to blue collar wage gaps for male, and the trend in these gaps for female. It is unable to replicate the hump shape in the evolution of male wages observed between 1970 and 1990. This could be because of the rather parsimonious parametrization of the aggregate production function, or of not allowing the returns to skills to vary over such a long period. The model also under-predicts the level of the college and white collar wage gaps for female, as noted in Footnote 25.

The fit of the model in terms of transition probabilities is evaluated in Table 5. The table presents actual and predicted transition probability matrix from blue collar, white collar, and home alternatives into blue collar, white collar, school, and home.[26] Transitions from the three alternatives are extremely well replicated

have discussed its relevance for inequality and labor supply, along with the importance of modeling two-earner households (see Greenwood, Guner and Vandenbroucke, 2017 for a review). All this is absent in the model, and even though an extension in this direction would be unfeasible, it could help in improving the fit of the aforementioned aggregates for females.

[26] Transitions from school into each of the four categories is omitted from the table because very few people is in school in the relevant age group (25-54).

Table 6—Out of Sample Fit: Act. vs Pred. Statistics for Immigrants

	Out-of-sample						In-sample	
	1970		1980		1990		1993-2007	
	Act.	Pred.	Act.	Pred.	Act.	Pred.	Act.	Pred.
A. Male								
Share with high school or less	0.67	0.69	0.57	0.61	0.52	0.55	0.55	0.56
Average years of education	10.8	11.1	11.4	11.8	11.7	12.1	11.9	12.1
Participation rate	0.77	0.56	0.68	0.61	0.63	0.66	0.75	0.72
Share of workers in blue collar	0.57	0.57	0.55	0.54	0.53	0.51	0.58	0.51
B. Female								
Share with high school or less	0.78	0.78	0.68	0.69	0.56	0.58	0.54	0.53
Average years of education	10.3	10.8	10.9	11.5	11.5	12.1	12.0	12.5
Participation rate	0.32	0.25	0.36	0.31	0.41	0.40	0.49	0.52
Share of workers in blue collar	0.46	0.45	0.45	0.44	0.39	0.43	0.41	0.43

Note: The table presents actual and predicted values of the listed aggregates for immigrants. Statistics for 1993-2007 are obtained from March Supplements of the CPS, and are used in the estimation. Data for 1970, 1980, and 1990 are from U.S. Census microdata samples and not used in the estimation.

by the model. In particular, the model captures very well the persistence in each of the alternatives, occupational switches, the fact that individuals rarely go back to school after leaving it, and transitions back and forth from working to home.

The formal discussion on identification in Section III.A, the discussion of the parameter estimates in Section IV.A, and the results in Figure 2 and Table 5 provide some evidence that the model presented in this paper and the variation used to identify its parameters are meaningful. Yet, it is reassuring to explore further evidence in the same direction. The remainder of this section presents six additional exercises that provide further validation of this conclusion. First, Table 6 analyzes the goodness of the model in predicting immigrant choices out-of-sample. As noted in Section III.C, whether a person is an immigrant or not is only identifiable in the CPS starting in 1993. Thus, no separate information for natives and immigrants before 1993 is used in the estimation. Given that the immigrant group is too small to drive the main aggregate trends (the percentage of immigrants in the population of working-age is below 10%), and that natives and immigrants had very different trends in education and choices over the period, correctly fitting these trends would provide evidence that individual choices are well identified, at least for immigrants. Table 6 evaluates the goodness of the model on fitting education, participation, and occupational choice of immigrants in census years 1970, 1980, and 1990. To do so, it compares predicted values from the model to data from the U.S. Census microdata samples, which are not used in the estimation. As it emerges from the table, the model does a good job in predicting levels, trends, and gender gaps for the different aggregates.

Second, Figure B1 in Appendix B gives a sense of whether the variation in the data is enough to identify the parameters. While the curvature of the minimum distance criterion function is difficult to represent in the multidimensional space,

Table 7—Estimated and Simulated Returns to Education

	Data		Simulation	
Least Squares (OLS)	0.096	(0.000)	0.096	(0.002)
Selection-corrected (Heckman, 1979)	0.123	(0.001)	0.114	(0.005)

Note: The table presents coefficients for years of education in OLS and Heckman (1979) selection-corrected regressions fitted on actual and simulated data. All regressions include dummies for potential experience (age minus education), gender, and year. In the selection-correction model, dummies for the number of children are included as exclusion restrictions. Actual data are obtained from the CPS. The sample period is 1967 to 2007. Random subsamples of 500,000 observations are drawn for both actual and simulated data. Nationally representative weights are used in the regressions. Standard errors, in parentheses, are calculated in the standard way in the left column, and are obtained from redrawing 100 times from the asymptotic distribution of the parameter estimates in the right column.

one can plot sections of it moving one of the parameters and leaving others fixed at the estimated values. Figure B1 provides these sections for each of the parameters of the model. As it emerges from the figure, all parameters move the criterion function substantially and have a clear minimum at the estimated value.

Third, to further evaluate what variation identifies education decisions in practice, in Table 7 I present estimates of returns to education obtained from fitting OLS and Heckman (1979) selection-corrected regressions on actual and simulated data. All regressions control for gender, year, and potential experience (age minus education) dummies. For the selection correction models, I use dummies for the number of children as exclusion restrictions. Overall, estimates in actual and simulated data are remarkably similar, which suggests that parameters are effectively identified from the variation discussed in Section III.A.

Fourth, as discussed above, the model presented in this paper is able to endogenously generate imperfect substitutability between observationally equivalent natives and immigrants, in the sense described in Ottaviano and Peri (2012) and Peri and Sparber (2009) (i.e. through natives and immigrants specializing in different occupations). Table 8 presents reduced form estimates of such elasticities of substitution, estimated following the approach described in Ottaviano and Peri (2012), and compares them with the results obtained by these authors. Point estimates are very much in line, which provides further confirmation of the ability of the model in fitting the data, and of the credibility of parameter identification.

Fifth, even though specified outside of the model, immigration is assumed to be endogenous to aggregate conditions and the aggregate shock. This implies that the aggregate shock is potentially correlated with immigrant inflows and immigrant composition. Whether empirically this is the case or not is checked in Figure 3. Taking the period 1993-2007, in which the CPS distinguishes immigrants from natives, I plot the predicted aggregate shock against the inflow rate of new immigrants, and their distribution of education, age, and region of origin. Results confirm the importance of not imposing orthogonality cconditions between the

Table 8—Predicted Elasticity of Substitution between Immigrants and Natives

	Ottaviano and Peri (2012)		Simulations: Census years: 1970-2006		Simulations: Anual frequency: 1967-2007	
Baseline regression:						
Men	-0.048	(0.010)	-0.054	(0.011)	-0.050	(0.009)
Pooled Men and Women	-0.037	(0.012)	-0.065	(0.017)	-0.073	(0.014)
Men, Labor Supply is Employment	-0.040	(0.012)	-0.022	(0.012)	-0.008	(0.010)
Regression without cell and year dummies:						
Men	-0.063	(0.005)	-0.084	(0.015)	-0.083	(0.017)
Pooled Men and Women	-0.044	(0.006)	-0.137	(0.019)	-0.150	(0.020)
Men, Labor Supply is Employment	-0.066	(0.006)	-0.063	(0.022)	-0.060	(0.026)

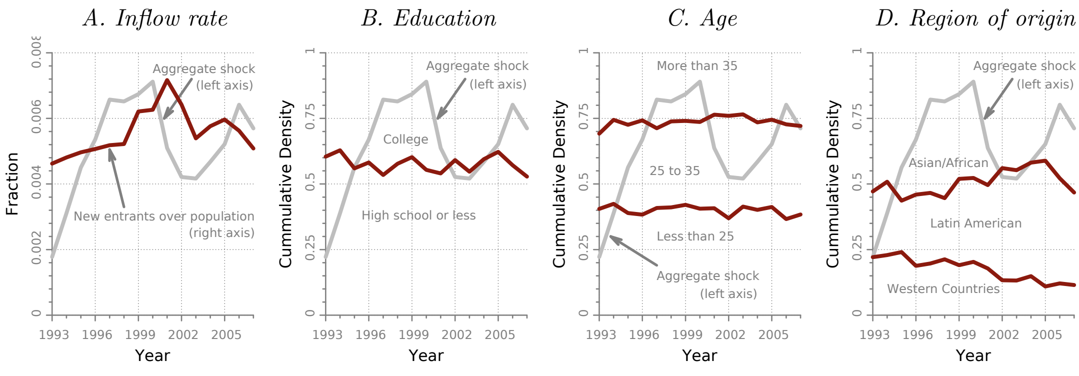
Note: The table presents OLS estimates of $-1/\sigma_N$ from the following regression:

$$\ln\left(w_{Fkt}/w_{Dkt}\right) = \varphi_k + \varphi_t - 1/\sigma_N \ln\left(L_{Fkt}/L_{Dkt}\right) + u_{kt},$$

where $\{F, D\}$ indicate immigrants and natives respectively, k indicates education-experience cells, t indicates calendar year, w indicates average wages of skill cell k in year t, and L is labor supply in the corresponding cell. This regression corresponds to Equation (8) in Ottaviano and Peri (2012). The first column of the simulation results uses the same frequency as in Ottaviano and Peri (2012), excluding 1960; The second one include years 1967-2007 with annual frequency. Standard errors, in parentheses, are obtained drawing 100 times from the asymptotic distribution of parameter estimates.

immigration process and the aggregate shock. The aggregate shock is markedly increasing between 1993 and 1997, stagnated between 1997 and 2000, decreasing between 2000 and 2004, and increasing again from 2005 to 2007. Pro-cyclically with some lag, immigrant inflows increased between 1993 and 2001 and decreased until 2003, increasing again in 2004 and 2005, and decreasing afterwards. Unlike aggregate inflows, the skill composition of new entrants seems quite stable along the business cycle, and even constant in the case of education and age at entry.[27]

Figure 3. Predicted Aggregate Shock and Recent Immigration



Note: Gray lines (right axis) plot the predicted values for the aggregate shock. Red lines (left axis) plot the share of new entrants over population (A), and, for them, the distribution of education (B), age (C), and country of origin (D). Data figures are smoothed with a 3-year moving average. Inflow rate is computed dividing the observed immigrants that entered over the preceding two (three) years divided by the number of years they refer to. *Source:* Current Population Survey, 1994-2008.

[27] If anything, only immigration from Latin America shows a slightly counter-cyclical evolution, at the expense of immigration from Asia/Africa, which would be somewhat pro-cyclical. Llull (2016) suggests that immigration from countries located at different distances from the destination can react differently to income shocks.

And, finally, I check the robustness of the model to setting the autoregressive coefficients of the expectation rules for skill prices equal to zero. Results for this exercise are included in the Online Supplement (Llull, 2017a). Estimates of the other parameters of the expectation rules, goodness of fit measures, and the model fit of the aggregates included in Figure 2 are virtually unchanged.

V. Understanding the consequences of immigration: counterfactual exercises

To evaluate the labor market effects of immigration, I compare baseline simulations with simulations of a counterfactual U.S. economy without the last four decades of mass immigration. Specifically, in the counterfactual economy, the share of immigrants among individuals of working-age is kept constant at the pre-immigration levels (by age and gender). In order to compensate for native population growth and retirement/death of previous immigrant cohorts, some immigration is allowed. I consider 1967 as the pre-immigration year, and 2007 as a final year.[28] Simulation results are presented for natives aged 25-54, separately for male and female. As discussed above, some of the aggregate processes, including immigration and capital stocks, are not modeled explicitly, but they are allowed to be endogenous to aggregate fluctuations. This implies that they are not necessarily policy invariant. Thus, the policy experiment needs to specify their counterfactual evolution.

Being the policy variable, the counterfactual sizes of the immigrant inflows are determined by design, as described in the previous paragraph. I also keep the composition of immigrant inflows invariant with respect to the baseline.[29] The choice of this particular policy experiment is motivated by three main reasons. First, it allows me to focus on size instead of composition effects of immigration, which gives a "wage elasticity to immigration" that is comparable to the one discussed in the literature. The predictions of this exercise thus provide an assessment of what the literature is missing by not taking into account human capital and labor supply adjustments by natives and previous immigrants. Second, this particular policy is consistent with implementing a lottery system for immigrant visas (which is how green cards and H1-B visas are assigned in practice). And third, in accordance to the results presented in Figure 3, composition of immigrant inflows seem to be mostly driven by long run trends rather than fluctuating in the short

[28] The share of immigrants in the workforce in 1967 is 5.1%. The choice of this particular period is based on three points. First, during mid 1960s, the fraction of foreign born individuals in the U.S. population reached its minimum level of the century. Second, one of the largest changes in U.S. immigration policy, the Amendments to the Immigration and Nationality Act, was passed in 1965. And third, it coincides with the estimation period.

[29] See Llull (2017c) for an analysis of immigration policies that are selective in terms of skills.

run, which suggests that demand factors (e.g. immigration policy) are less likely to drive the skill composition of immigration. The effects of changing immigrant composition and the consequences of selective immigration policies are left for future research.

Finally, capital adjustments to immigration have been subject to debate in the literature. Borjas (2003) assumes that the capital is the same in baseline and counterfactual scenarios. Ottaviano and Peri (2012) keep the return to capital fixed. In a one-time-inflow steady state economy, the first scenario defines the short run impact of immigration, the second provides the long run effect, and the elasticity of capital supply determines the length of the transition between steady states. However, in a non-stationary framework like the one used here, for a big economy like the U.S. that can influence world interest rates, and in which new waves of immigrants arrive each year, the *true* impact of immigration depends on the speed of adjustment of capital and on the world's supply of capital and labor. In this case, the no adjustment/full adjustment scenarios are only bounds for the *true* effect. In my simulation exercises, I focus on these two bounds.[30]

A. The effect of immigration on wages

The way in which immigration affects wages of natives is threefold. First, there is a direct impact through the changes in labor supply produced by the inflow of new workers, which affects skill prices in equilibrium. Second, natives may adjust their skills and choices, and affect their wages as a result. And, finally, as a result of these adjustments, there are equilibrium feedback effects on skill prices that lead to a different equilibrium. In this section, I disentangle these different channels, and I put them in perspective with respect to the exercises implemented by previous papers in the literature.

Table 9 presents separate evidence of the first and third channels. The top row presents counterfactuals in which natives and immigrants are not allowed to adjust their choices and skills with respect to the baseline. Thus, it provides a good measure of the effect on impact described as the first channel. When capital is not allowed to adjust, effects on impact are large (-4.92% for blue collar skill prices, and -3.90% for white collar). This results are in line with the literature. For the 1980-2000 period, Borjas (2003) finds that, holding physical and human capital fixed, immigration reduced wages by 3.2%. According to Table 1 above, the share of immigrants in the workforce increased from 5.7 to 16.56 percent (10.86 points)

[30] In the full capital adjustment scenario, the counterfactual returns to structures and equipment capital are assumed to be the same as in the baseline so that the elasticity of the capital supply is infinity. The fact that the model delivers different (gross) returns to each of the two types of capital is not necessarily inconsistent with the presence of a single interest rate, as depreciation rates of structures and equipment capital may differ.

Table 9—Effects on Skill Prices and the Role of Equilibrium

	No capital adjustment ($\partial K/\partial m = 0$):				Full capital adjustment ($\partial r_K/\partial m = 0$):			
	Blue collar		White collar		Blue collar		White collar	
No labor market adjustment	-4.92	(0.95)	-3.90	(0.60)	-1.76	(0.99)	0.86	(0.46)
Equilibrium effect	2.36	(0.78)	0.58	(0.72)	1.63	(1.00)	-0.86	(0.46)
Total effect	-2.56	(0.36)	-3.33	(0.39)	-0.13	(0.49)	-0.00	(0.15)

Note: The table compares baseline and counterfactual skill prices. Left and right panels correspond to different assumptions on counterfactual capital as indicated. "No labor market adjustment" indicates a scenario in which individuals are not allowed to adjust their human capital, occupational choice, and labor supply in response to immigration. "Equilibrium effect" is the difference between the total effect and the effect without labor market adjustment. Standard errors, in parentheses, are obtained drawing 100 times from the asymptotic distribution of parameter estimates.

between 1970 and 2008, and from 7.13 to 14.62 (7.49 points) in the 1980-2000 period. Taking into account that about 37.5% of employees work in blue collar jobs, a back of the envelope calculation suggests that results in Table 9 (for 1967-2007) are consistent with a decrease of $(4.92 \times 0.375 + 3.90 \times 0.625) \div 10.86 \times 7.49 = 3.0\%$ for 1980-2000, almost exactly the effect obtained by Borjas (2003). The same calculation for the full capital adjustment scenario delivers virtually no change on average, as in Ottaviano and Peri (2012). Similar to what occurs in these papers, however, immigration produces redistribution, as different "types" of labor are affected differently. Results in Table 9 suggest that, in both scenarios, blue collar skill prices are the most negatively affected by immigration on impact, which is consistent with the fact that immigrants tend to cluster in blue collar jobs.

The bottom row of Table 9 shows the total effect of immigration on skill prices, accounting for equilibrium feedback effects after adjustments in choices and human capital take place. The central row includes the difference between top and bottom rows, and, therefore, provides a quantification of the equilibrium feedback effects (third channel). Results suggest that equilibrium forces substantially mitigate the initial impact of immigration. When capital is not allowed to adjust, about one half and one third of the initial effects on blue collar and white collar skill prices disappear. When capital fully adjusts, initial effects are arbitraged out completely. Moreover, in both scenarios, differential effects across occupations are partially arbitraged out, and in the first case relative effects are even reverted. Thus, equilibrium feedback effects are central to understand wage effects of immigration, and ignoring them induces substantial biases in estimation.

Table 10 shows wage effects of immigration on wages for different groups of natives, based on age, education, and gender. As noted by Borjas (2003), Ottaviano and Peri (2012) and many others, immigrants tend to affect more the wages of native workers that are closer substitutes. Table 10 shows that, on impact, this is also the case here. Results in the top row of each panel indicate that, when natives are not allowed to adjust their choices and skills, younger individuals are more affected than older, individuals with high school education or less are more affected than individuals with college education, and male are more negatively affected than female. For example, if capital is not allowed to react to immigration, wages of young male with high school education or less are reduced, on average, by 4.74% on impact, and those of old college educated female are reduced by only 4.02%; when capital fully reacts, average wage effects on these groups are -1.32% and $+0.57\%$ respectively. However, a different story occurs as a consequence of human capital and labor supply adjustments. In this case, wages of high school, of older individuals, and of males, are the ones that adjust the most. These individuals are more likely to work in blue collar jobs, and, hence, they have the possibility of switching to white collar. Additionally, there is selection out of the labor market, which is also more prevalent in these groups (except for male vs female). And finally, while less educated individuals can increase their education (and switch to a white collar career), highly educated individuals could decide to obtain less education if the expected returns are lower. In our earlier example, wage effects are adjusted to -2.49% and to -3.59% when capital does not react, and -0.06% and 1.16% when it fully adjusts.

In sum, Tables 9 and 10 provide evidence of the importance of accounting for labor supply and human capital in the estimation of wage effects of immigration. Effects on impact are straightforward: wages of closer competitors to immigrants are more severely affected. And results are numerically consistent with existing results in the literature. However, when natives alter their human capital and labor supply decisions, they introduce non-trivial adjustments. Natives that are more similar to immigrants are more affected on impact, but they have a larger margin of adjustment, which makes them ultimately less affected in some cases.

B. Human capital and labor supply adjustments

The equilibrium effects on wages described above summarize a variety of heterogeneous adjustments. Individuals affect average wages by changing their labor supply decisions, and by accumulating different levels of human capital. Tables 11 and 12 respectively summarize these two.

Table 11 compares baseline and counterfactual choices in the terminal year 2007. The left column indicates the fraction of individuals in blue collar, white collar, or home in the absence of mass immigration that make a different choice as a consequence of immigration. The right panel presents the distribution of choices conditional on adjusting. When capital is not allowed to adjust, 8.8% of male and 15.1% of female in blue collar occupations change their decisions. For male, the majority of them 55.8% switch occupations, and around 38.5% decide not to work.

Table 10—Wage Effects for Different Groups

	High school:				College:			
Age group:	25–39		40–54		25–39		40–54	
	No capital adjustment ($\partial K/\partial m = 0$):							
A. Male								
No labor market adjustment	-4.74	(0.72)	-4.55	(0.49)	-4.33	(0.22)	-4.18	(0.27)
Equilibrium effect	2.26	(0.82)	3.22	(0.90)	1.47	(0.70)	0.80	(0.44)
Total effect	-2.49	(0.28)	-1.33	(0.49)	-2.86	(0.66)	-3.38	(0.54)
B. Female								
No labor market adjustment	-4.47	(0.39)	-4.12	(0.38)	-4.08	(0.44)	-4.02	(0.48)
Equilibrium effect	1.33	(0.61)	2.48	(0.63)	0.43	(1.24)	0.43	(1.24)
Total effect	-3.14	(0.49)	-1.64	(0.50)	-3.65	(0.96)	-3.59	(1.00)
	Full capital adjustment ($\partial r_K/\partial m = 0$):							
A. Male								
No labor market adjustment	-1.32	(0.86)	-0.82	(0.64)	-0.24	(0.20)	0.15	(0.10)
Equilibrium effect	1.38	(0.93)	0.91	(0.73)	0.50	(0.47)	-0.01	(0.22)
Total effect	0.06	(0.17)	0.10	(0.20)	0.26	(0.30)	0.14	(0.22)
B. Female								
No labor market adjustment	-0.60	(0.58)	0.29	(0.42)	0.41	(0.54)	0.57	(0.49)
Equilibrium effect	0.69	(0.58)	-0.15	(0.40)	0.81	(1.03)	0.59	(0.94)
Total effect	0.09	(0.30)	0.14	(0.20)	1.22	(0.83)	1.16	(0.78)

Note: The table compares baseline and counterfactual average log wages for native males and females in different groups. In each panel, results are presented for different assumptions on counterfactual capital as indicated. "No labor market adjustment" indicates a scenario in which individuals are not allowed to adjust their human capital, occupation, and participation decisions. "Equilibrium effect" is the difference between the total effect and the effect without labor market adjustment. Standard errors, in parentheses, are obtained drawing 100 times from the asymptotic distribution of parameter estimates.

A small fraction (5.6%) switch blue collar for school. In the case of blue collar female, the most frequent adjustment is towards the home alternative (58.3%), followed by switching to white collar (40.9%).

Overall, the fraction of blue collar individuals that switch to white collar is similar across genders ($8.8 \times 55.8\% = 4.9\%$ for male, and $15.1 \times 40.9\% = 6.2\%$ for female). The key difference is that a much larger fraction of blue collar female switches to the home alternative ($15.1 \times 58.3\% = 8.8\%$ versus $8.8 \times 38.5\% = 3.4\%$ of males). This is not surprising, given that labor market attachment of female is typically lower. Another important adjustment is white collar female towards the home alternative ($5.2 \times 87.5\% = 4.6\%$). Other adjustments are smaller.

When capital fully reacts, labor market adjustments are less severe. The majority of adjustments still occur to individuals who would be working in blue collar, but now the detachment from the labor market is less frequent, in favor of more occupation switching. This result is natural given that in the no capital adjustment scenario, on top of the changes in relative wages on impact, there is an important reduction of the overall wage levels, whereas in the full capital adjustment case, relative wages change, but the average level is virtually unchanged

Table 11—Labor Supply Adjustments

Choice w/o immigr.	Fraction adjusting		Of which: Switch occ.		Stay home		Go to school	
				No capital adjustment ($\partial K/\partial m = 0$):				
A. Male								
Blue collar	8.8	(1.2)	55.8	(4.0)	38.5	(3.2)	5.6	(1.4)
White collar	7.7	(1.6)	52.3	(3.9)	40.1	(3.4)	7.6	(1.0)
Home	4.1	(0.7)	——		——		14.4	(2.9)
B. Female								
Blue collar	15.1	(4.2)	58.3	(7.9)	40.9	(7.3)	0.8	(1.0)
White collar	5.2	(2.0)	11.2	(7.1)	87.5	(6.4)	1.3	(1.4)
Home	4.0	(1.1)	——		——		1.7	(3.6)
				Full capital adjustment ($\partial r_K/\partial m = 0$):				
A. Male								
Blue collar	2.1	(1.4)	68.6	(5.3)	27.4	(5.2)	4.1	(1.3)
White collar	0.3	(0.6)	58.2	(9.7)	25.2	(8.4)	16.6	(4.3)
Home	1.6	(1.0)	——		——		9.1	(2.9)
B. Female								
Blue collar	6.1	(3.9)	68.7	(12.4)	30.7	(11.9)	0.6	(1.1)
White collar	0.7	(1.6)	46.6	(17.3)	52.1	(15.9)	1.3	(3.0)
Home	2.6	(1.5)	——		——		11.6	(5.6)

Note: The left column presents the percentage of native male and female individuals aged 25-54 that, in the cross-section of 2007, change their decisions in baseline and counterfactual simulations. The three remaining colums show the percentage of these individuals that do each of the adjustments indicated in the top row. Percentages are presented conditional on the choice made in the absence of immigration (counterfactual). Top and bottom panels make different assumptions a the counterfactual evolution of capital as indicated. Standard errors, in parentheses, are obtained drawing 100 times from the asymptotic distribution of parameter estimates.

on impact. Adjustments from individuals working in white collar become almost inexistent, and also those for female who stay at home. Male staying at home in the absence of immigration are more likely to reenter the labor market, mostly into white collar jobs.

Human capital adjustments are driven by two confronting forces. On one side, white collar occupations become more attractive and, given that returns to education are larger in those occupations, education also becomes more attractive. On the other hand, the increased competition puts downward pressure on wages (especially blue collar) and reduces labor market attachment, which disincentives the accumulation of skills (education and experience). Which of these two prevails in each context is an empirical question that is addressed in Table 12.

The first column of the table summarizes the average career and education adjustments for male and female under both capital scenarios. In particular, it presents the number of years education, each of the experiences, and the time spent at home are increased. By construction, the sum of all increases equals zero. In the case of education, the discouragement effect seems to dominate both for male

Table 12—Education and Career Adjustments

	All		Increase Education		Reduce Education		Keep educ. & change exp.	
			No capital adjustment ($\partial K/\partial m = 0$):					
A. Male								
Share of total	100.0		1.3	(0.4)	11.1	(2.4)	13.8	(1.6)
Average change in years of:								
Education	-0.28	(0.09)	3.20	(0.19)	-2.91	(0.12)	0.00	(0.00)
Experience in blue collar	-0.23	(0.16)	-7.65	(1.37)	2.33	(0.40)	-2.54	(0.37)
Experience in white collar	-0.07	(0.14)	3.32	(0.79)	-2.41	(0.31)	0.82	(0.34)
Time spent at home	0.58	(0.12)	1.13	(0.55)	2.99	(0.37)	1.72	(0.09)
B. Female								
Share of total	100.0		3.9	(1.4)	3.6	(3.0)	6.6	(0.7)
Average change in years of:								
Education	-0.00	(0.11)	2.71	(0.53)	-3.04	(0.30)	0.00	(0.00)
Experience in blue collar	-0.29	(0.15)	-4.26	(2.07)	1.04	(1.05)	-2.33	(1.01)
Experience in white collar	0.14	(0.21)	7.27	(2.37)	-4.17	(0.92)	-0.03	(0.77)
Time spent at home	0.15	(0.27)	-5.72	(2.09)	6.17	(0.80)	2.36	(0.72)
			Full capital adjustment ($\partial r_K/\partial m = 0$):					
A. Male								
Share of total	100.0		1.2	(0.7)	0.2	(0.8)	2.8	(1.6)
Average change in years of:								
Education	0.03	(0.03)	3.06	(0.17)	-2.90	(0.29)	0.00	(0.00)
Experience in blue collar	-0.13	(0.12)	-4.77	(1.55)	3.98	(1.88)	-2.40	(1.22)
Experience in white collar	0.08	(0.08)	3.20	(0.82)	-2.63	(1.28)	1.20	(0.69)
Time spent at home	0.02	(0.06)	-1.49	(0.95)	1.55	(0.79)	1.20	(0.67)
B. Female								
Share of total	100.0		3.8	(1.9)	0.4	(2.1)	1.7	(0.8)
Average change in years of:								
Education	0.08	(0.09)	2.34	(0.49)	-2.69	(0.55)	0.00	(0.00)
Experience in blue collar	-0.12	(0.06)	-1.71	(1.51)	-0.09	(1.57)	-2.91	(1.65)
Experience in white collar	0.14	(0.18)	3.81	(1.51)	-6.03	(1.49)	1.25	(1.18)
Time spent at home	-0.11	(0.26)	-4.44	(1.29)	8.81	(1.29)	1.66	(1.39)

Note: The top row of each panel indicates the fraction of individuals in each of the groups listed in the top row. The four rows at the bottom indicate the average change in the number of years in each of the alternatives accumulated by 2007. By construction, the sum of changes across alternatives in a given panel adds to zero. Different panels provide simulation results for the two genders in different capital scenarios as indicated. Standard errors, in parentheses, are obtained drawing 100 times from the asymptotic distribution of parameter estimates.

and female when capital is not allowed to adjust. However, this effect is reversed when capital fully adjusts, as education is increased on average. This difference is the result, once again, of the different effect on wage levels in the two scenarios. Furthermore, natives move out of blue collar careers towards white collar careers and out of the labor market. When capital fully adjusts, occupation switching prevails. In the no capital adjustment scenario, labor market detachment is the alternative that increases the most, both for male and female.

These aggregate effects mask much larger heterogeneous effects that compensate with each other. Between 1.2 and 1.3% of male and 3.8–3.9% of female increase their education, respectively by more than 3 and 2.3–2.7 years on average (second column). These individuals tend to replace their blue collar career for a white collar one, and increase or roughly keep their labor market attachment. However, another group reduce their education (third column). This group represents 11% of male, but only 3.6% of female when capital does not adjust, and less than one percent of them when capital fully adjusts.[31] For them, the disincentive effect dominates, and they spend 2.7 to 3.0 years less at school on average. They substantially detach from the labor market: they spend between 1.6 and 3.0 additional years at home in the case of male, and between 6.1 and 8.8 additional years in the case of female. There is third group of individuals who, even though they do not adjust their education, change their career profile (fourth column). They represent 13.8% and 6.6% of male and female when capital does not adjust, and 2.8% and 1.7% when it fully adjusts. These individuals typically reduce their participation as blue collar workers at the expense of increasing their white collar experience in some cases, and their time at home in others.

In sum, Tables 11 and 12 provide evidence of heterogeneous adjustments for different individuals, which highlights the different forces at place. Some individuals switch from a blue collar career to a white collar career. When they are in school, they tend to extend their education. Others detach from the labor market, and they tend to drop out from school earlier as a result. When capital fully adjusts, the first effect dominates, and the overall level of education is increased. However, if capital does not adjust, the second effect prevails, and education is reduced on average. The magnitude of education adjustments for individuals who change their education could potentially be increased if the model included permanent unobserved heterogeneity. In my model, individuals learn whether they will follow a particular career path based mostly on the sequence of shocks they experience. With permanent unobserved heterogeneity, these paths would be somewhat more predetermined, and individuals potentially at the margin between two different career paths, which are the ones who adjust education, would know so with higher certainty. Thus, their expected gains from adjustment would be increased, and the adjustments they do would be more sizable. Whether fewer or more individuals would adjust is less clear.

C. Self-selection and the effect of immigration along the wage distribution

Immigration does not affect all individuals in the same way. Section I shows that immigrants are less skilled than natives, and increasingly more concentrated in blue collar occupations, and Sections V.A and V.B provide evidence of heterogeneous effects on wages for different workers, and of heterogeneous human capital and labor supply adjustments. Using data for the U.K., Dustmann et al. (2013)

[31] This difference between male and female could be driven by the under-prediction of female education discussed in Footnote 25. Since female education is under-predicted in the baseline, the predicted share of female that do this adjustment could also be under-predicted.

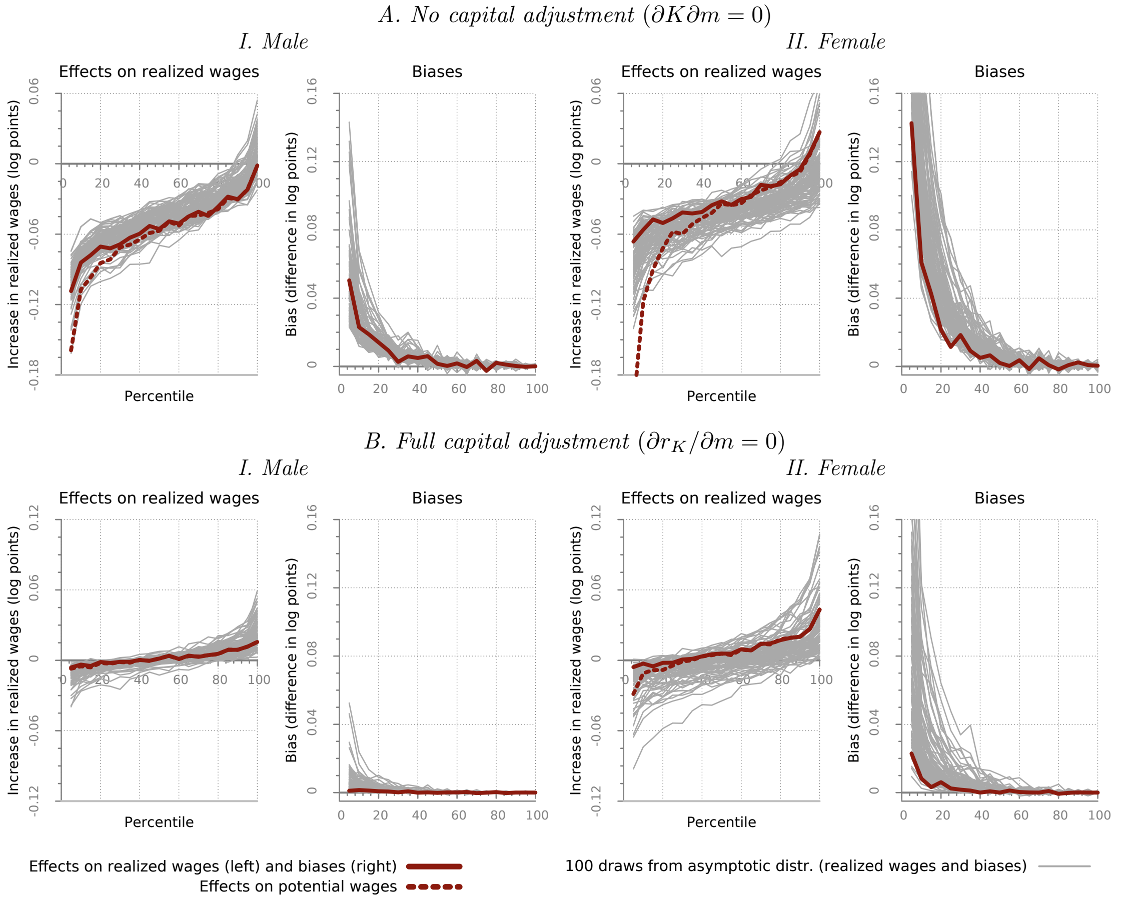
find heterogeneous wage effects along the native wage distribution. They find that, in positions of the wage distribution where immigrants tend to be overrepresented, wage effects are more negative than in positions with fewer immigrants. A relevant question is whether the model is able to predict similar results.

This question is addressed in Figure 4. Black solid lines in the left plot of each panel present results for an exercise that is analogous to Dustmann et al. (2013). In particular, these lines plot the effect of immigration along the distribution of wages of U.S.-born male and female aged 25-54. They compare individual wages in baseline and counterfactual simulations for individuals that work in both cases. Thus, they show the increase in realized wages induced by immigration in each percentile. Top and bottom panels differ in the assumed counterfactual evolution of capital. Gray lines depict analogous simulation results obtained from 100 random draws of the asymptotic distribution of the parameter estimates.

Results differ across scenarios but have a common feature: individuals at the bottom of the wage distribution are more negatively affected than individuals at the top. When capital is not allowed to adjust, both level and redistribution effects are stronger. Wage effects go down to -10.8% and -6.6% for male and female respectively in the bottom percentiles. They are less negative at the top of the distribution: at the very top, they get close to zero for males, and even positive for females. When capital fully adjusts, negative effects at the lower half of the wage distribution are relatively small (less than 1% for both male and female), and at the top of the distribution they turn positive and go up to 1.6% and 4.3% for male and female respectively. This is in line with Dustmann et al. (2013).

As it occurs in Dustmann et al. (2013), black lines in Figure 4 compare wages in baseline and counterfactual scenarios for individuals that are working in each case. However, results in Section V.B suggest that an important fraction of individuals adjust to immigration by dropping out from the labor market. This labor market detachment is unlikely to be random. Instead, individuals at the bottom of the wage distribution are more likely to be deterred than individuals at the top. This non-random selection out of the labor force may generate biases in the estimation of wage effects of immigration (a similar argument to the standard selectivity bias described in Gronau (1974) and Heckman (1979)). One of the main advantages of the structural model is that it provides a way to correct for this bias. This is so because wage functions are explicitly specified, and idiosyncratic shocks are simulated for all alternatives. Thus, we can simulate the potential wage of all individuals in baseline and in counterfactual scenarios, whether they work (in which case it coincides with realized wage) or not. As a result, we can compute the effect of immigration on potential wages for all individuals that work in absence of immigration. This comparison is depicted by dashed lines in the left plot

Figure 4. Wage Effects Along the Wage Distribution and Selection Biases



Note: The figure plots the average differences in log hourly wages in baseline and counterfactual scenarios along the baseline wage distribution of native male and female aged 25-54 in 2007. The left figure in each pair represents wage effects on realized wages (solid black) and on potential wages (dashed black), and the right figure plots the difference between the two. Gray lines plot the effects on realized wages and the biases obtained for 100 random draws from the asymptotic distribution.

of each panel in Figure 4. Results show that the selectivity bias substantially affects estimates of wage effects below the median, and is particularly severe at the bottom of the distribution. At the 5th percentile, the drop in wages induced by immigration goes from 10.8% to 15.9% and from 0.6% to 0.7% for males, without and with capital adjustment respectively, and from 6.6% to 20.9% and from 0.6% to 2.9% for females in the two scenarios.

Black solid lines in the right plots for each panel show the difference between the effects on realized and potential wages, thus providing a quantification of this bias. Gray lines are again obtained from 100 random draws from the empirical distribution. In all cases both the main simulations and the random draws show positive biases at the bottom half of the wage distribution, and are particularly large at the bottom, confirming the results in the left plots. This suggests that the estimated biases are not only large, but also statistically significant.

These biases are quite large for both genders. Yet, they are larger for female.

They represent 46% and 20% of the initial effect for males, and 275% and 235% for females. This is not surprising, given that female attachment to the labor market is typically weaker, especially at the bottom of the wage distribution. Importantly, they reinforce the importance of accounting for labor supply and human capital adjustments when analyzing wage effects of immigration.

VI. Conclusion

This paper estimates a labor market equilibrium dynamic discrete choice model to quantify wage effects of immigration taking into account labor market adjustments by natives and previous generations of immigrants. The model, estimated using micro-data from CPS and NLSY for 1967-2007, is used to simulate a counterfactual U.S. economy without the last four decades of mass immigration. The exercise delivers three main conclusions. First, labor market adjustments are crucial to understand the effect of immigration on wages and inequality, and not allowing for them generates sizable biases in the estimated effects. Second, as a result of immigration some individuals (those switching to a white collar career and/or increasing their attachment to the labor market) increase their education substantially while others (those reducing their labor market attachment) reduce it, and a dynamic model is essential to identify these effects. And third, non-random detachment from the labor market, which mostly affects individuals at the bottom tail of the wage distribution, introduces additional self-selection biases to the estimation of wage effects of immigration.

REFERENCES

Acemoglu, Daron, "Technical Change, Inequality, and the Labor Market," *Journal of Economic Literature*, March 2002, *40* (1), 7–72.

_ **and David H. Autor**, "Skills, Tasks and Technologies: Implications for Employment and Earnings," in Orley C. Ashenfelter and David E. Card, eds., *Handbook of Labor Economics*, Vol. 4B, Amsterdam: North-Holland Publishing Company, 2011, chapter 12, pp. 1043–1171.

Adda, Jérôme, Christian Dustmann, and Katrien Stevens, "The Career Costs of Children," *Journal of Political Economy*, April 2017, *125* (2), 293–337.

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.

_ **and** _ , "Dynamic Discrete Choice Structural Models: A Survey," *Journal of Econometrics*, May 2010, *156* (1), 38–67.

Ahn, Hyungtaik and James L. Powell, "Semiparametric Estimation of Censored Selection Models with Nonparametric Selection Mechanism," *Journal of Econometrics*, July 1993, *58* (1-2), 3–29.

Altuğ, Sumru and Robert A. Miller, "The Effect of Work Experience on Female Wages and Labor Supply," *Review of Economic Studies*, January 1998, *65* (1), 45–85.

Arcidiacono, Peter and Robert A. Miller, "Conditional Choice Probability Estimation of Dynamic Discrete Choice Models with Unobserved Heterogeneity," *Econometrica*, November 2011, *79* (6), 1823–1867.

_ **and** _ , "Identifying Dynamic Discrete Choice Models Off Short Panels," mineo Carnegie Mellon University, June 2015.

Baker, Bryan and Nancy Rytina, "Estimates of the Unauthorized Immigrant Population Residing in the United States: January 2012," Technical Report, Office of Immigration Statistics, Department of Homeland Security August 2013.

Ben-Porath, Yoram, "The Production of Human Capital and the Life Cycle of Earnings," *Journal of Political Economy*, August 1967, *75* (4), 352–365.

Borjas, George J., "Assimilation, Changes in Cohort Quality, and the Earnings of Immigrants," *Journal of Labor Economics*, October 1985, *3* (4), 463–489.

_ , "Self-Selection and the Earnings of Immigrants," *American Economic Review*, September 1987, *77* (4), 531–553.

_ , "Assimilation and Changes in Cohort Quality Revisited: What Happened to Immigrant Earnings in the 1980s?," *Journal of Labor Economics*, April 1995, *13* (2), 201–245.

_ , "The Economic Analysis of Immigration," in Orley C. Ashenfelter and David E. Card, eds., *Handbook of Labor Economics*, Vol. 3A, Amsterdam: North-Holland Publishing Company, 1999, chapter 28, pp. 1697–1760.

_ , "The Labor Demand Curve Is Downward Sloping: Reexamining the Impact of Immigration on the Labor Market," *Quarterly Journal of Economics*, November 2003, *118* (4), 1335–1374.

_ , "Native Internal Migration and the Labor Market Impact of Immigration," *Journal of Human Resources*, Spring 2006, *41* (2), 221–258.

_ , "Immigration and Globalization: A Review Essay," *Journal of Economic Literature*, December 2015, *53* (4), 961–974.

_ **, Jeffrey T. Grogger, and Gordon H. Hanson**, "Immigration and the Economic Status of African-American Men," *Economica*, April 2010, *77* (306), 255–282.

_ **, Richard B. Freeman, and Lawrence F. Katz**, "On the Labor Market Impacts of Immigration and Trade," in George J. Borjas and Richard B. Freeman, eds., *Immigration and the Work Force: Economic Consequences for the United States and Source Areas*, Chicago: The University of Chicago Press, 1992, chapter 7, pp. 213–244.

_ , _ , and _ , "How Much Do Immigration and Trade Affect Labor Market Outcomes?," *Brookings Papers on Economic Activity*, Spring 1997, *1997* (1), 1–67.

Card, David E., "The Causal Effect of Education on Earnings," in Orley C. Ashenfelter and David E. Card, eds., *Handbook of Labor Economics*, Vol. 3A, Amsterdam: North-Holland Publishing Company, 1999, chapter 30, pp. 1801–1863.

_ , "Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration," *Journal of Labor Economics*, January 2001, *19* (1), 22–64.

_ , "Immigration and Inequality," *American Economic Review: Papers and Proceedings*, May 2009, *99* (2), 1–21.

_ and Thomas Lemieux, "Can Falling Supply Explain The Rising Return To College For Younger Men? A Cohort-Based Analysis," *Quarterly Journal of Economics*, May 2001, *116* (2), 705–745.

Das, Mitali, Whitney K. Newey, and Francis Vella, "Nonparametric Estimation of Sample Selection Models," *Review of Economic Studies*, January 2003, *80* (1), 33–58.

Dustmann, Christian and Ian Preston, "Comment: Estimating the Effect of Immigration on Wage," *Journal of the European Economic Association*, February 2012, *10* (1), 216–223.

_ , Tommaso Frattini, and Ian Preston, "The Effect of Immigration along the Distribution of Wages," *Review of Economic Studies*, January 2013, *80* (1), 145–173.

Eckstein, Zvi and Yoram Weiss, "On the Wage Growth of Immigrants: Israel 1990-2000," *Journal of the European Economic Association*, June 2004, *2* (4), 665–695.

Greenwood, Jeremy, Nezih Guner, and Guillaume Vandenbroucke, "Family Economics Writ Large," *Journal of Economic Literature*, 2017, *forthcoming.*

_ , _ , Georgi Kocharkov, and Cezar Santos, "Marry Your Like: Assortative Mating and Income Inequality," *American Economic Review: Papers and Proceedings*, May 2014, *104* (5), 348353.

_ , Zvi Hercowitz, and Per Krusell, "Long-Run Implications of Investment-Specific Technological Change," *American Economic Review*, June 1997, *87* (1), 342–362.

Griliches, Zvi, "Capital-Skill Complementarity," *Review of Economics and Statistics*, November 1969, *51* (4), 465–468.

Gronau, Reuben, "Wage Comparisons – A Selectivity Bias," *Journal of Political Economy*, November-December 1974, *82* (6), 1119–1143.

Guner, Nezih, Yuliya A. Kulikova, and Joan Llull, "Marriage and Health: Selection, Protection, and Assortative Mating," Barcelona GSE Working Paper 795, March 2017.

Hamermesh, Daniel S., "The Demand for Labor in the Long Run," in Orley C. Ashenfelter and David E. Card, eds., *Handbook of Labor Economics*, Vol. 1, Amsterdam: North-Holland Publishing Company, 1986, chapter 8, pp. 429–471.

Hanson, Gordon H., "Illegal Migration from Mexico to the United States," *Journal of Economic Literature*, December 2006, *44* (4), 869–924.

Heckman, James J., "Shadow Prices, Market Wages, and Labor Supply," *Econometrica*, July 1974, *42* (4), 679–694.

_ , "Sample Selection Bias as a Specification Error," *Econometrica*, January 1979, *47* (1), 153–161.

_ **and Richard Robb**, "Alternative Methods for Solving the Problem of Selection Bias in Evaluating the Impact of Treatments on Outcomes," in Howard Wainer, ed., *Drawing Inferences from Self-Selected Samples*, New York: Springer-Verlag. (Reprinted in 2000, Mahwah, NJ: Lawrence Erlbaum Associates.), 1986, chapter 4, pp. 63–107.

_ **, Lance Lochner, and Christopher Taber**, "Explaining Rising Wage Inequality: Explorations with a Dynamic General Equilibrium Model of Labor Earnings with Heterogeneous Agents," *Review of Economic Dynamics*, January 1998, *1* (1), 1–58.

_ **,** _ **, and Petra E. Todd**, "Earnings Functions, Rates of Return, and Treatment Effects: The Mincer Equation and Beyond," in Eric A. Hanushek and Finis Welch, eds., *Handbook of the Economics of Education*, Vol. 1, Amsterdam: Elsevier Science, 2006, chapter 7, pp. 307–458.

Hotz, V. Joseph and Robert A. Miller, "Conditional Choice Probabilities and the Estimation of Dynamic Models," *Review of Economic Studies*, July 1993, *60* (3), 497–529.

Hunt, Jennifer, "The Impact of Immigration on the Educational Attainment of Natives," *Journal of Human Resources*, Published online before print, September 2016.

Kaldor, Nicholas, "A Model of Economic Growth," *Economic Journal*, December 1957, *67* (268), 591–624.

Kambourov, Gueorgui and Iourii Manovskii, "Occupational Mobility and Wage Inequality," *Review of Economic Studies*, April 2009, *76* (2), 731–759.

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.

Kennan, John and James R. Walker, "Wages, Welfare Benefits, and Migration," *Journal of Econometrics*, May 2010, *156* (1), 229–238.

_ **and** _ , "The Effect of Expected Income on Individual Migration Decisions," *Econometrica*, January 2011, *79* (1), 211–251.

Kristensen, Dennis, Lars Nesheim, and Áureo de Paula, "CCP and the Estimation of Nonseparable Dynamic Models," mimeo, University College London, November 2015.

Krusell, Per and Anthony A. Smith Jr., "Income and Wealth Heterogeneity in the Macroeconomy," *Journal of Political Economy*, October 1998, *106* (5), 867–896.

_ **, Lee E. Ohanian, José-Víctor Ríos-Rull, and Giovanni L. Violante**, "Capital-Skill Complementarity and Inequality: A Macroeconomic Analysis," *Econometrica*, September 2000, *68* (5), 1029–1054.

LaLonde, Robert J. and Robert H. Topel, "The Assimilation of Immigrants in the U.S. Labor Market," in George J. Borjas and Richard B. Freeman, eds., *Immigration and the Work Force: Economic Consequences for the United States and Source Areas*, Chicago: The University of Chicago Press, 1992, chapter 3, pp. 67–92.

Lee, Donghoon, "An Estimable Dynamic General Equilibrium Model of Work, Schooling, and Occupational Choice," *International Economic Review*, February 2005, *46* (1), 1–34.

_ **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.

Lessem, Rebecca H., "Mexico-U.S. Immigration: Effects of Wages and Border Enforcement," mimeo, Carnegie Mellon University, October 2015.

Lewis, Ethan G., "Immigration, Skill Mix, and Capital-Skill Complementarity," *Quarterly Journal of Economics*, May 2011, *126* (2), 1029–1069.

Llull, Joan, "Understanding International Migration: Evidence from a New Dataset of Bilateral Stocks (1960-2000)," *SERIEs — The Journal of the Spanish Economic Association*, June 2016, *7* (2), 221–255.

_ , "Online Supplement to "Immigration, Wages, and Education: a Labor Market Equilibrium Structural Model"," available online at `http://pareto.uab.cat/jllull/Papers/Supplementary_material_REStud2016.pdf`, March 2017.

_ , "The Effect of Immigration on Wages: Exploiting Exogenous Variation at the National Level," *Journal of Human Resources*, Published online before print, April 2017.

_ , "Selective Immigration Policies and the U.S. Labor Market," mimeo, Universitat Autónoma de Barcelona, June 2017.

_ **and Robert A. Miller**, "Internal Migration in Dual Labor Markets," mimeo, Universitat Autònoma de Barcelona, December 2016.

Lubotsky, Darren, "Chutes or Ladders? A Longitudinal Analysis of Immigrant Earnings," *Journal of Political Economy*, October 2007, *115* (5), 820–867.

Magnac, Thierry and David Thesmar, "Identifying Dynamic Discrete Decision Processes," *Econometrica*, March 2002, *70* (2), 801–816.

Manacorda, Marco, Alan Manning, and Jonathan Wadsworth, "The Impact of Immigration on the Structure of Wages: Theory and Evidence from Britain," *Journal of the European Economic Association*, February 2012, *10* (1), 120–151.

McFadden, Daniel L., "A Method of Simulated Moments for Estimation of Discrete Response Models Without Numerical Integration," *Econometrica*, September 1989, *57* (5), 995–1026.

Mincer, Jacob, *Schooling, Experience and Earnings*, New York: Columbia University Press, 1974.

Ottaviano, Gianmarco I. P. and Giovanni Peri, "Rethinking the Effect of Immigration on Wages," *Journal of the European Economic Association*, February 2012, *10* (1), 152–197.

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

Piyapromdee, Suphanit, "The Impact of Immigration on Wages, Internal Migration, and Welfare," mimeo, University College London, June 2015.

Smith, Christopher L., "The Impact of Low-Skilled Immigration on the Youth Labor Market," *Journal of Labor Economics*, January 2012, *30* (1), 55–89.

Tinbergen, Jan, *Income Difference: Recent Research*, Amsterdam: North-Holland Publishing Company, 1975.

Vella, Francis, "Estimating Models with Sample Selection Bias: A Survey," *Journal of Human Resources*, Winter 1998, *33* (1), 127–169.

Appendix A: Detailed discussion of identification

The following discussion of model identification builds on previous work by Hotz and Miller (1993), Altuğ and Miller (1998), Magnac and Thesmar (2002), Arcidiacono and Miller (2011, 2015), and Kristensen, Nesheim and de Paula (2015). The arguments rely on the assumption of conditional independence of the idiosyncratic shocks. They also rely on the assumption that conditional choice probabilities (CCPs) are identified nonparametrically from observed choices in the data. The latter is not trivial in practice, mainly because the aggregate shock z_t and the skill prices $r_t^j(z_t)$, which are not observable, are state variables. For the sake of the argument, assume initially that CCPs are identified. I discuss this assumption below. In what follows, I denote by $\tilde{\Omega}_{a,t}$ the partition of the state vector $\Omega_{a,t}$ that excludes the idiosyncratic shocks ε_a, and $\mathbb{P}(\tilde{\Omega}_{a,t})$ denotes the vector of CCPs.

A1. Wages

Identification of the wage equations follows standard arguments from the self-selection literature (e.g. Heckman (1974, 1979); see Vella (1998) for a review). As in Heckman and Robb (1986), we can write Equation (3) as:

$$\begin{aligned}\ln w_{a,t,l}^j = \ln r_t^j + \omega_{0,l}^j + \omega_{1,is}^j E_a + \omega_2^j X_{Ba} + \omega_3^j X_{Ba}^2 + \omega_4^j X_{Wa} + \omega_5^j X_{Wa}^2 \\ + \omega_6^j X_{Fa} + \lambda^j(\mathbb{P}(\tilde{\Omega}_{a,t}); \Sigma_g) + \epsilon_a^j,\end{aligned} \tag{A1}$$

where $\lambda^j(\mathbb{P}(\tilde{\Omega}_{a,t}); \Sigma_g) \equiv \mathbb{E}[\varepsilon_a^j | \tilde{\Omega}_{a,t}, d_a = j]$ is the standard selection correction, and ϵ_a is independent of $\tilde{\Omega}_{a,t}$ and d_a. The functional form of $\lambda^j(.)$ is implied by the assumption that F^ε is a zero mean multivariate normal distribution with a variance-covariance matrix characterized by the gender-specific parameters Σ_g. Despite its derivation is cumbersome, given normality and absence of additive separability assumptions, Kristensen et al. (2015) prove that this representation is still valid in this context. Thus, the vector of return parameters ω^j is identified off wage data for individuals working in occupation j using the orthogonality conditions of the form $\mathbb{E}[\epsilon_a^j | \tilde{\Omega}_{a,t}, d_a = j] = 0$. I assume that σ_{Bg}, σ_{Wg}, and ρ_{BW} are also identified from these moment conditions, and those for the variance in wages, given the specification of $\lambda^j(.)$. Aggregate skill prices r_t^j, which are not observable, are identified as the coefficients of calendar time dummies. This requires a normalization of one of the intercepts in each wage equation, $w_{0,l}^j$ for some l; I normalize native-male intercepts in both equations to zero.

Ahn and Powell (1993) prove that the parameters of the wage equation would be identified even if $\lambda^j(.)$ was not known, as long as the model includes exclusion restrictions that affect the participation decision but not wages. The model includes such restrictions. For example, the number of children affects the utility of the home alternative (and hence participation), but not wages. Under the appropriate exclusion restrictions, Das, Newey and Vella (2003) prove that the wage equation would be identified even nonparametrically, up to an additive constant.

A2. Production function, aggregate shock process, and expectation rule

Identification of skill prices in Equation (A1) leads to identification of the production function parameters, and the aggregate shock. Given skill prices, individual skill units are identified as $s^j_{a,i} = \exp\{\ln w^j_{a,t,l} - \ln r^j_t\}$, and aggregate skill units are identified aggregating those over the sample of individuals working in occupation j. Using data on output Y_t, blue collar and white collar labor shares, $\varsigma_{Bt} \equiv \frac{r^B_t S_{Bt}}{Y_t}$ and $\varsigma_{Wt} \equiv \frac{r^W_t S_{Wt}}{Y_t}$, are also identified. Noting that the share of output devoted to pay labor and equipment capital is equal to $1 - \lambda$, we can write:

$$\varsigma_{Bt} + \varsigma_{Wt} + \varsigma_{Et} = 1 - \lambda, \tag{A2}$$

where $\varsigma_{Et} \equiv \frac{r^E_t K_{Et}}{Y_t}$ is the equipment capital share in period t. Hence:

$$\varsigma_{Et} = 1 - \lambda - (\varsigma_{Bt} + \varsigma_{Wt}) \equiv 1 - \lambda - \varsigma_{Lt}, \tag{A3}$$

where ς_{Lt} is the labor share in period t.

The demand for capital equipment is analogous to that for white collar labor in Equation (12). Rewriting it in terms of the capital equipment share yields:

$$\varsigma_{Et} = (1-\lambda)(1-\alpha)(1-\theta)\left(\frac{KW_t}{KBW_t}\right)^{\rho}\left(\frac{K_{Et}}{KW_t}\right)^{\gamma}, \tag{A4}$$

where $KBW_t \equiv \{\alpha S^{\rho}_{Bt} + (1-\alpha)[\theta S^{\gamma}_{Wt} + (1-\theta)K^{\gamma}_{Et}]^{\rho/\gamma}\}^{1/\rho}$ is the CES aggregate labor and equipment capital in the production function, and $KW_t \equiv [\theta S^{\gamma}_{Wt} + (1-\theta)K^{\gamma}_{Et}]^{1/\gamma}$ is the equipment capital-white collar labor CES aggregate. Dividing this expression by the demand equation for white collar skill units in Equation (12) gives, upon rearrangement:

$$\varsigma_{Et} = \varsigma_{Wt}\frac{1-\theta}{\theta}\left(\frac{K_{Et}}{S_{Wt}}\right)^{\gamma}. \tag{A5}$$

Taking logs and first differencing we obtain:

$$\Delta \ln \varsigma_{Et} = \Delta \ln \varsigma_{Wt} + \gamma \Delta \ln \frac{K_{Et}}{S_{Wt}}. \tag{A6}$$

Substituting ς_{Et} by its expression in Equation (A3) gives

$$\Delta \ln \varsigma_{Et} = \Delta \ln(1 - \lambda - \varsigma_{Lt}) = -\ln\left(1 + \frac{\Delta\varsigma_{Lt}}{1-\lambda-\varsigma_{Lt}}\right) \approx -\frac{\Delta\varsigma_{Lt}}{1-\lambda-\varsigma_{Lt}}. \tag{A7}$$

This approximation should be quite accurate, as the changes in the labor share are known to be small. Deriving the equivalent expression for $\Delta_2 \ln \varsigma_{Et}$, where Δ_2 indicates the second difference, substituting back into Equation (A6), and dividing by the first difference we obtain:

$$\frac{\Delta \varsigma_{Lt}}{\Delta_2 \varsigma_{Lt}} = \frac{\Delta \ln \varsigma_{Wt} + \gamma \Delta \ln(K_{Et}/S_{Wt})}{\Delta_2 \ln \varsigma_{Wt} + \gamma \Delta_2 \ln(K_{Et}/S_{Wt})}. \tag{A8}$$

Solving for γ in the above expression we obtain:

$$\gamma = \frac{(\Delta_2 \varsigma_{Lt})\, \Delta \ln \varsigma_{Wt} - (\Delta \varsigma_{Lt})\, \Delta_2 \ln \varsigma_{Wt}}{(\Delta \varsigma_{Lt})\, \Delta_2 \ln(K_{Et}/S_{Wt}) - (\Delta_2 \varsigma_{Lt})\, \Delta \ln(K_{Et}/S_{Wt})}, \tag{A9}$$

where all elements in the right-hand-side are identified functions from the data. Substituting this expression into Equation (A6), replacing $\Delta \ln \varsigma_{Et}$ by its approximation in Equation (A7), and solving for λ we obtain:

$$\lambda = 1 - \varsigma_{Lt} - \frac{\Delta \varsigma_{Lt}}{\Delta \ln \varsigma_{Wt} + \gamma \Delta \ln \frac{K_{Et}}{S_{Wt}}}, \tag{A10}$$

where all the right-hand-side elements are identified.

Having identified λ, the sequence of equipment capital prices (and shares) is identified from Equation (A3). Parameter θ is identified from Equation (A5), and, hence the CES aggregate KW_t is also identified. The parameters ρ and α are similarly identified using the ratio of white collar and blue collar shares:

$$\frac{\varsigma_{Wt}}{\varsigma_{Bt}} = \frac{(1-\alpha)\theta}{\alpha} \left(\frac{KW_t}{S_{Bt}}\right)^{\rho} \left(\frac{S_{Wt}}{KW_t}\right)^{\gamma}. \tag{A11}$$

Taking logs, first differencing, and solving for ρ we obtain:

$$\rho = \frac{\Delta \ln(\varsigma_{Wt}/\varsigma_{Bt}) + \gamma \Delta \ln(KW_t/S_{Wt})}{\Delta \ln(KW_t/S_{Bt})}, \tag{A12}$$

where all the elements in the right hand side are identified. Since ρ is identified, α is identified from Equation (A11).

Finally, having identified λ, γ, θ, ρ, and α, the sequence of aggregate shocks is obtained as a residual in Equation (6). Given them, the AR(1) coefficients for the shock process (ϕ_0 and ϕ_1 in Equation (7)) are identified as regression coefficients, following standard time series arguments. Likewise, combining $\{z_t\}_{t=t_0}^{T}$ with the recovered sequence of skill prices we identify the parameter vector η from the expectation rule in Equation (9) as regression coefficients.

A3. Utility parameters

The identification of the remaining parameters of the model follows standard arguments in the literature. I fix the discount factor β, which is proved to be identified only through the functional form assumptions of the model (Magnac and Thesmar, 2002). The transition probability function for the preschool children process is identified from observed transitions in the data. The parameters that remain to be identified are δ_g^{BW}, $\delta_{0,l}^S$, $\delta_{1,g}^S$, τ_1, τ_2, σ_{Sg}, $\delta_{0,l}^H$, $\delta_{1,g}^H$, $\delta_{2,g}^H$, and σ_{Hg}. Given that the wage equation parameters are identified, and since δ_g^{BW} is common for alternatives $j = B, W$, there is no need for further normalizations.

Proposition 1 in Hotz and Miller (1993) establish that the mapping between value functions and CCPs can be inverted so that we can express continuation values as a function of the CCPs. Kristensen et al. (2015) prove that this result still holds in the case in which utility functions do not satisfy additive separability, as it is the case here. The argument is facilitated from the fact that wage equations are already identified as noted above.

Let $d_{a+\ell}^*$ denote the optimal choice at age $a+\ell$ when the state vector before observing the idiosyncratic shocks is $\tilde{\Omega}_{a+\ell,t+\ell}$. In each of the periods following age a through age 65, individuals obtain an expected flow utility given by:

$$\begin{aligned}\mathbb{E}[U_{a+\ell,l}|\Omega_{a,t}, d_a, l] = {} & F(\tilde{\Omega}_{a+\ell,t+\ell}|\Omega_{a,t}, d_a, l) \\ & \times \left\{ \textstyle\sum_{j\in\{B,W,S,H\}} \mathbb{E}[U_{a+\ell,l}^j|\tilde{\Omega}_{a+\ell,t+\ell}, l, d_{a+\ell}^* = j] P(d_{a+\ell}^* = j|\tilde{\Omega}_{a+\ell,t+\ell}, l) \right\}.\end{aligned} \tag{A13}$$

The transition function $F(\tilde{\Omega}_{a+\ell,t+\ell}|\Omega_{a,t}, d_a, l)$ is a function of the CCPs, and of the transition processes given by Equations (7), (8), and (9). The evolution of the state variables that do not transit deterministically given choices (i.i.d. idiosyncratic shock, number of children, and skill prices) is determined by the Markovian functions $F^r(.)$, $F^n(.)$, and $F^\varepsilon(.)$, which are identified as noted above —the latter up to parameters σ_{Sg} and σ_{Hg}. Assuming that the CCPs $\mathbb{P}(\tilde{\Omega}_{a+\ell,t+\ell})$ are nonparametrically identified, the expectation term in Equation (A13) can be written as a function of the CCPs, the state variables $\tilde{\Omega}_{a+\ell,t+\ell}$, and parameters. Thus, the continuation value:

$$\mathbb{E}\left[V_{a+1,t+1,l}(\Omega_{a+1,t+1}) \mid \Omega_{a,t}, d_a, l\right] = \sum_{\ell=1}^{65-a} \mathbb{E}[U_{a+\ell,l}|\Omega_{a,t}, d_a, l] \equiv \bar{V}_j(\mathbb{P}), \tag{A14}$$

can also be expressed as a function of the CCPs, state variables, and parameters.

To complete the discussion, we need to note that the discrete choices are made based on the difference between the utilities of the given alternatives and that of a base one. Given that the parameters from the wage equation are identified from Equation (A1), it is natural to fix one of the working alternatives as the base. For example, fix the blue collar option for this role. This utility can be expressed as:

$$\mathbb{w}^B(\tilde{\Omega}_{a,t}) - \delta^{BW}\, \mathbb{1}\{d_{a-1} = H\} + \tilde{\varepsilon}_a^B + \beta V_B(\mathbb{P}), \tag{A15}$$

where $\mathbb{w}^B(\tilde{\Omega}_{a,t})$ is the exponential of all the elements in Equation (A1) except for the last two multiplied by $\exp\left(\frac{1}{2}\sigma^2_{Bg}\right)$, and $\tilde{\varepsilon}^B_a \equiv \mathbb{w}^B(\tilde{\Omega}_{a,t})\left[\exp\left(\varepsilon^B_a - \frac{1}{2}\sigma^2_{Bg}\right) - 1\right]$ is a conditionally-heteroskedastic zero-mean shifted log-normal shock. Both $\mathbb{w}^B(\tilde{\Omega}_{a,t})$ and σ_{Bg} are identified from the wage data, as discussed above. Likewise, since δ^{BW}_g is common to the blue collar and the white collar alternatives there is no need to make further normalizations. Finally, the remaining variances are also identified without further normalizations since the variances of $\tilde{\varepsilon}^B_a$ and the analogous $\tilde{\varepsilon}^W_a$ are identified, as so is the covariance.[32] Additionally, all covariances that involve ε^S_g or ε^H_g are set to zero. Thus, the only two variance parameters that are left to be identified are σ_{Sg} and σ_{Hg}, which are, hence, identified.

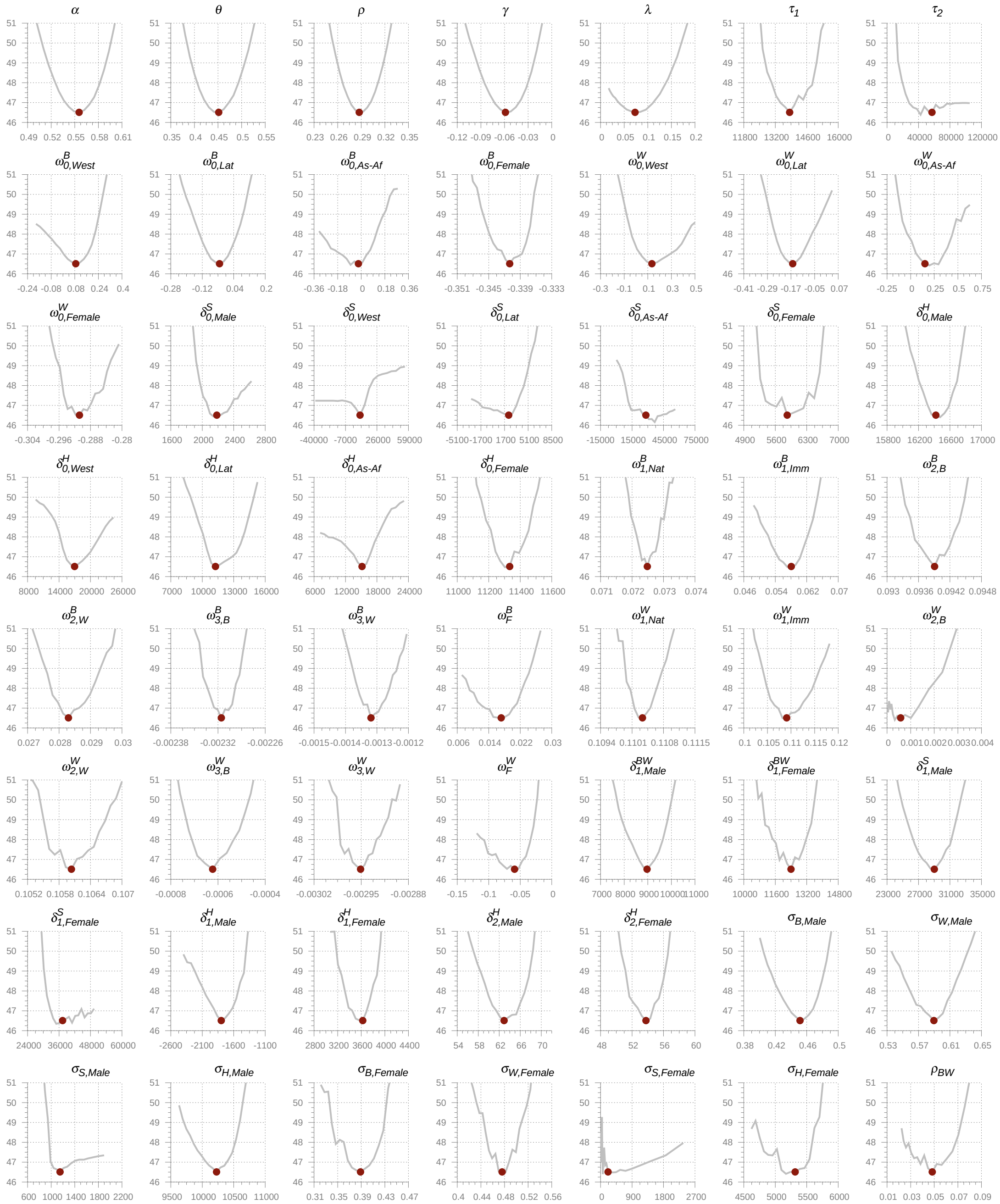
A4. Conditional choice probabilities

All the discussion so far assumes that CCPs are identified. However, two complications impede the use of this assumption in practice. The first one is that the CCPs are actually not identified because they depend on the aggregate shock z_t and the skill prices $r^j_t(z_t)$, which are not observed. Even if they were observed we would only have as many realizations of them as periods are in the data. A solution for this complication is given by Arcidiacono and Miller (2015). The other complication is that, as discussed in Section III.C, we do not observe all state variables and outcomes in the same data set. I assume that CCPs and conditional wages are identified off the long list of statistics used in the estimation.

[32] $\text{Var}(\tilde{\varepsilon}^B_a|\tilde{\Omega}_{a,t}) = \mathbb{w}^B(\tilde{\Omega}_{a,t})^2[\exp(\sigma^2_{Bg}) - 1]$ and $\text{Var}(\tilde{\varepsilon}^W_a|\tilde{\Omega}_{a,t}) = \mathbb{w}^W(\tilde{\Omega}_{a,t})^2[\exp(\sigma^2_{Wg}) - 1]$, as well as $\text{Cov}(\tilde{\varepsilon}^B_a, \tilde{\varepsilon}^W_a|\tilde{\Omega}_{a,t}) = \mathbb{w}^B(\tilde{\Omega}_{a,t})\mathbb{w}^W(\tilde{\Omega}_{a,t})[\exp(\rho_{BW}\sigma_{Bg}\sigma_{Wg}) - 1]$, are functions of the data and identified parameters.

Appendix B: Curvature of the objective function

Figure B1. Sections of the objective function



Note: Solid lines plot the evolution of the objective function when changing the corresponding parameter and leaving others constant at the estimated values. Red dots indicate point estimates.

Appendix C: Standard errors

Parameters estimates solve the following minimum distance problem:

$$\begin{aligned}\hat{\theta} &= \arg\min_{\theta} ||\hat{\pi}(x) - \tilde{\pi}(x_S(\theta))|| = \\ &= \arg\min_{\theta} [\hat{\pi}(x) - \tilde{\pi}(x_S(\theta))]' W [\hat{\pi}(x) - \tilde{\pi}(x_S(\theta))]. \end{aligned} \tag{C1}$$

Weights are proportional to the sample size used to calculate each statistic. Specifically, W is a diagonal matrix with the (weighted) sample size of each element.[33]

The asymptotic distribution of parameters is obtained by applying the delta method to the sample statistics. In particular:

$$\text{Var}(\hat{\theta}) = (G'WG)^{-1} G'WV_0WG(G'WG)^{-1}, \tag{C2}$$

where G is the $P \times R$ matrix of partial derivatives of the R statistics included in π with respect to the P parameters included in θ.

In the estimation problem defined by Equation (C1) there are two sources of error. First, data statistics $\hat{\pi}(x)$ are estimated with sampling error. And second, the function that maps parameters into statistics, $\tilde{\pi}(x_S(\theta))$, does not have a closed form solution, and I need to simulate it, introducing a simulation error.

The remainder of this Appendix is devoted to provide an estimator of V_0. It is important to notice that, given the two sources of error, asymptotic theory should be applied two-way: taking the sample size and the number of simulations to infinity. To handle it, the problem can be split in the difference between the following two elements: $\sqrt{N}\left(\hat{\pi}(x) - \pi(\theta_0)\right)$ and $\sqrt{M}\left(\tilde{\pi}(x_S(\theta_0)) - \pi(\theta_0)\right)$, where N is the sample size and M is the number of simulations.

C1. Minimum distance asymptotic results

Consider R statistics from the data such that:

$$\mathbb{E}[Y_K] = \pi_k(\theta_0), \quad k = 1, ..., R. \tag{C3}$$

Without loss of generality, we assume that these statistics are means. These means are estimated with k different samples S_k, each of them of size N_k. Some of these samples may overlap (e.g., the sample used to estimate the share of 16-20 years old males choosing to work in blue collar in year 1967 may include some individuals

[33] Weighted sample size is defined in this context as $\left(\sum_i p_i^2 / \left(\sum_i p_i\right)^2\right)^{-1}$, where p_i is the individual weight in the sample. If $p_i = p \ \forall i$, this sum is equal to the sample size. The weighted sample size is inverse of the precision of the variance of the weighted sample mean: $Var(\bar{x}) = \sigma_x^2 \sum_i p_i^2 / \left(\sum_i p_i\right)^2$.

that are also used to estimate the share of high school dropout males choosing blue collar in that year). Sample counterparts of these statistics are given by:

$$\hat{\pi}_k = \frac{1}{N_k} \sum_{i \in S_k} Y_{ki}. \tag{C4}$$

Therefore, if the functional form of $\pi(\theta)$ was known, we could write:

$$\hat{\theta} = \arg\min_{\theta \in \Theta} ||\hat{\pi} - \pi(\theta)||. \tag{C5}$$

Let us introduce some additional notation:

$$d_{ki} \equiv \mathbb{1}\{i \in S_k\}, \tag{C6}$$

$$S_{ij} \equiv S_i \cap S_j, \tag{C7}$$

$$S \equiv S_1 \cup ... \cup S_R, \tag{C8}$$

$$N \equiv \sum_{i \in S} \left(\sum_k d_{ki} - \sum_k \sum_j d_{ki} d_{ji} \right), \tag{C9}$$

$$\lambda_{kN} \equiv \frac{N_k}{N} \xrightarrow{N \to \infty} \lambda_k, \tag{C10}$$

$$\psi_{ki} \equiv Y_{ki} - \pi_k(\theta_0). \tag{C11}$$

Now we can write:

$$\begin{pmatrix} \sqrt{N_1}(\hat{\pi}_1 - \pi_1) \\ \sqrt{N_2}(\hat{\pi}_2 - \pi_2) \\ ... \\ \sqrt{N_R}(\hat{\pi}_R - \pi_R) \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{N_1}} \sum_{i \in S_1} \psi_{1i} \\ \frac{1}{\sqrt{N_2}} \sum_{i \in S_2} \psi_{2i} \\ ... \\ \frac{1}{\sqrt{N_R}} \sum_{i \in S_R} \psi_{Ri} \end{pmatrix} = \begin{pmatrix} \sqrt{\lambda_{1N}} & 0 & \dots & 0 \\ 0 & \sqrt{\lambda_{2N}} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sqrt{\lambda_{RN}} \end{pmatrix}^{-1} \times$$

$$\times \frac{1}{\sqrt{N}} \sum_{i \in S} \begin{pmatrix} d_{1i}\psi_{1i} \\ d_{2i}\psi_{2i} \\ ... \\ d_{Ri}\psi_{Ri} \end{pmatrix} \equiv \Lambda \frac{1}{\sqrt{N}} \sum_{i \in S} d_i \circ \psi_i, \tag{C12}$$

where $\circ$ denotes the Hadamard or element-by-element product. Due to the central limit theorem (CLT), and Cramer's theorem, as $N \to \infty$:

$$\Lambda \frac{1}{\sqrt{N}} \sum_{i \in S} d_i \circ \psi_i \underset{d}{\to} \mathcal{N}\left(0, \Lambda\, \mathbb{E}[(d_i \circ \psi_i)(\psi_i \circ d_i)']\Lambda\right). \tag{C13}$$

Therefore, by the analogy principle we can define an estimator of the variance-covariance matrix of the R sample statistics as:

$$\hat{\Omega} = \begin{pmatrix} \frac{1}{N_1}\hat{\sigma}_1^2 & \frac{N_{12}}{N_1 N_2}\hat{\sigma}_{12} & \dots & \frac{N_{1R}}{N_1 N_R}\hat{\sigma}_{1R} \\ \frac{N_{12}}{N_1 N_2}\hat{\sigma}_{12} & \frac{1}{N_2}\hat{\sigma}_2^2 & \dots & \frac{N_{2R}}{N_2 N_R}\hat{\sigma}_{12} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{N_{1R}}{N_1 N_R}\hat{\sigma}_{1R} & \frac{N_{2R}}{N_2 N_R}\hat{\sigma}_{2R} & \dots & \frac{1}{N_R}\hat{\sigma}_R^2 \end{pmatrix} \tag{C14}$$

where $\hat{\sigma_{ij}} = \frac{1}{N_{ij}} \sum_{k \in S_{ij}} \psi_{ki}\psi'_{kj}$, and $\hat{\sigma_i^2} = \frac{1}{N_i} \sum_{k \in S_i} \psi_{ki}\psi'_{ki}$.

C2. Simulated minimum distance asymptotic results

Suppose $\hat{\pi}$ is an estimator of some characteristic π of the distribution of Y based on the sample $\{Y_i\}_{i=1}^N$ such that:

$$\sqrt{N}[\hat{\pi} - \pi(\theta_0)] \underset{d}{\rightarrow} \mathcal{N}(0, \Omega). \tag{C15}$$

Let $g(.,.)$ and $F(.)$ be known functions. Assume $Y = g(U, \theta_0)$ with $U \sim F$. Also let $\tilde{\pi}(\theta_0, U^M)$ be the same estimating formula as $\hat{\pi}$ but based on the artificial sample $\{g(U_j, \theta_0)\}_{j=1}^M$ constructed from a simulated sample U^M. Since $M \rightarrow \infty$:

$$\sqrt{M}[\tilde{\pi}(\theta_0, U^M) - \pi(\theta_0)] \underset{d}{\rightarrow} \mathcal{N}(0, \Omega) \tag{C16}$$

independently of $\hat{\pi}$. Therefore, as long as $0 < \lim_{N,M\rightarrow\infty}(N/M) \equiv \kappa < \infty$

$$\sqrt{N}[\hat{\pi} - \tilde{\pi}(\theta_0, U^M)] =$$

$$= \sqrt{N}[\hat{\pi} - \pi(\theta_0)] - \sqrt{\frac{N}{M}}\sqrt{M}[\tilde{\pi}(\theta_0, U^M) - \pi(\theta_0)] \underset{d}{\rightarrow} \mathcal{N}\left(0, (1+\kappa)\,\Omega\right) \tag{C17}$$

Note that this result includes the case in which we can simulate a sample of size m for every observation $i = 1, ..., N$, so that $M = mN$, and $\kappa = 1/M$, which is the case analyzed in McFadden (1989).

Finally, to generalize the result to multiple statistics with overlapping samples as in Section C1, let $(l_{1i}, ..., l_{Ri})$ and $(\delta_1, ..., \delta_R)$ play the role of $(d_{1i}, ..., d_{Ri})$ and $(\lambda_1, ..., \lambda_R)$ in the simulated samples, we similarly have that:

$$\begin{pmatrix} \sqrt{M_1}(\tilde{\pi}_1(\theta_0, U^{M_1}) - \pi_1) \\ \sqrt{M_2}(\tilde{\pi}_2(\theta_0, U^{M_2}) - \pi_2) \\ ... \\ \sqrt{M_R}(\tilde{\pi}_R(\theta_0, U^{M_R}) - \pi_R) \end{pmatrix} \equiv \Delta \frac{1}{\sqrt{M}} \textstyle\sum_{i \in U^M} l_i \circ \psi_i \underset{d}{\rightarrow} \mathcal{N}\left(0, \Delta\, \mathbb{E}[(l_i \circ \psi_i)(\psi_i \circ l_i)']\Delta\right). \tag{C18}$$

Therefore:

$$\begin{pmatrix} \sqrt{N_1}(\hat{\pi}_1 - \tilde{\pi}_1(\theta_0, U^{M_1})) \\ \sqrt{N_2}(\hat{\pi}_2 - \tilde{\pi}_2(\theta_0, U^{M_2})) \\ ... \\ \sqrt{N_R}(\hat{\pi}_R - \tilde{\pi}_R(\theta_0, U^{M_R})) \end{pmatrix} \underset{d}{\rightarrow} \mathcal{N}\left(0, V_0\right), \tag{C19}$$

and:

$$\hat{V} = \begin{pmatrix} \left(\frac{1}{N_1} + \frac{1}{M_1}\right)\hat{\sigma}_1^2 & \left(\frac{N_{12}}{N_1 N_2} + \frac{M_{12}}{M_1 M_2}\right)\hat{\sigma}_{12} & \cdots & \left(\frac{N_{1R}}{N_1 N_R} + \frac{M_{1R}}{M_1 M_R}\right)\hat{\sigma}_{1R} \\ \left(\frac{N_{12}}{N_1 N_2} + \frac{M_{12}}{M_1 M_2}\right)\hat{\sigma}_{12} & \left(\frac{1}{N_2} + \frac{1}{M_1}\right)\hat{\sigma}_2^2 & \cdots & \left(\frac{N_{2R}}{N_2 N_R} + \frac{M_{2R}}{M_2 M_R}\right)\hat{\sigma}_{12} \\ \vdots & \vdots & \ddots & \vdots \\ \left(\frac{N_{1R}}{N_1 N_R} + \frac{M_{1R}}{M_1 M_R}\right)\hat{\sigma}_{1R} & \left(\frac{N_{2R}}{N_2 N_R} + \frac{M_{2R}}{M_2 M_R}\right)\hat{\sigma}_{2R} & \cdots & \left(\frac{1}{N_R} + \frac{1}{M_R}\right)\hat{\sigma}_R^2 \end{pmatrix}. \tag{C20}$$