The Value of a Job and the Structure of Labor Market Surplus

Davide Alonzo* Giovanni Gallipoli*

October 18, 2020

Preliminary and Incomplete Draft

Abstract
We characterize the distribution of occupation-specific worker surplus for different demographic groups in the US labor market. We combine information about earnings and the distribution of workers across jobs to separately identify systematic pecuniary and non-pecuniary components of worker surplus. We then use variation in estimated worker-job match values to (i) recover labor factor productivity across occupation and demographic groups, and (ii) estimate substitution intensity between heterogeneous job inputs. Through counterfactual exercises, we quantify the extent to which technological progress, as opposed to shifts in the heterogeneous valuations of identical jobs, account for structural transformation in the labor market.

JEL Codes: J62, J24.

Keywords: occupations; workers; returns; heterogeneity, technology, equilibrium.

*Vancouver School of Economics, University of British Columbia, 6000 Iona Drive, Vancouver, BC V6T 1L4, Canada. E-mail: davidealonzo03@gmail.com and gallipol@mail.ubc.ca.
1 Introduction

The structure of employment and wages has changed significantly over recent decades. Considerable labor market shifts have been documented in the literature. These include the increase in the employment and wages of skilled workers (Katz and Murphy, 1992; Katz and Autor, 1999; Beaudry et al., 2016; Valletta, 2017), the decline of middle-paying occupations (Acemoglu and Autor, 2011), the emergence of IT-intensive occupations (Gallipoli and Makridis, 2018), the growing presence of women in high paying occupations (Cortes et al., 2018), the growing reward to soft and non-cognitive skills in the labor market (Deming, 2017), the shrinking labor supply of young men (Aguiar et al., 2017) and the convergence of the occupational distributions of different demographic groups (Hsieh et al., 2019).

While we generally recognize that both values and distributions of worker-job matches have experienced profound and lasting changes, we know much less about the nature of these changes. Can they be explained by productivity dynamics alone? Or are they the by-product of adjustments in the non-pecuniary value of different occupations? To what extent are similar jobs valued differently by different workers? And do these different forces reinforce or offset each other in equilibrium?

The observation that occupational sorting and job characteristics have evolved over time suggests the possibility that workers’ surplus from observationally identical matches (i.e. workers with certain characteristics matched to specific occupations) may have also changed, in both pecuniary and non-pecuniary returns.\(^1\) In this paper we characterize the distribution of occupation-specific returns for different demographic groups in the US labor market. We document how these returns can be decomposed into pecuniary and non-pecuniary components, and how these have evolved over time. We then quantify the extent to which different aspects of technological change, as opposed to shifts in the non-pecuniary value of jobs, account for the rapidly changing structure of employment and wages.

We begin by presenting a model that delivers a simple measure of the heterogeneous values that different workers associate with identical occupations. The procedure is sufficiently general to decompose each worker’s surplus into different components, and further relate them to observable characteristics of both job and worker. We examine equilibrium outcomes in a competitive labor market with a heterogeneous supply side (workers) and a representative firm demanding labor in a set of heterogeneous occupations. Sorting is driven by the size of the match-specific surplus, which includes a systematic component (common to all matches of a particular worker-job type) and an individual random component. The systematic surplus comprises pecuniary and non-pecuniary returns, both of which vary over time. The model delivers a tractable empirical counterpart that can be estimated using observations from

\(^1\)For example, in addition to experiencing changes in wages, different occupations may have experienced changes in their time demands (Erosa et al., 2017; Cubas et al., 2019) and, therefore, their value to different types of workers beyond what is captured by standard pecuniary measures.
repeated cross-sections of the distribution of earnings and workers across jobs. The selection of individual workers into specific occupations hinges on match-quality draws, as in classical models of random job selection described in the taxonomy of earnings functions of Willis (1986). Unlike those models, systematic match quality explicitly reflects measurable non-pecuniary components that drive total surplus from employment.

Our approach imposes minimal assumptions on the mechanics driving equilibrium outcomes, apart from the low-level requirements of a standard Roy-model in which relative surplus comparisons drive sorting. Moreover, the approach is flexible enough to allow for the identification of changes in the distributions of heterogeneous rewards (pecuniary and non-pecuniary). Hence, the model can be used to characterize if, and how, these returns have changed for specific demographic subsets of the working-age population. That is, one can transparently draw inference about which types of workers have been positively (or negatively) impacted by the changes in the occupational structure of the labor market over recent decades, and the extent to which these impacts occur through changes in pecuniary or non-pecuniary returns.

Estimates reveal that large and persistent shifts have occurred in the distribution of job-specific worker surplus since the 1980s. Some of these changes are surprising, insofar they suggest that gender-specific surplus may be different from what wages alone may convey. Our findings are consistent with the view that the impact of changes in employment composition (due to technological change and globalization) has been very uneven across demographic groups (e.g. Autor et al., 2014; Cortes et al., 2017) and with the observation that certain groups may have suffered while others thrived.

In the second part of the paper, we nest our estimates of the distribution of worker surplus within a simple production structure and we draw inference about relative productivities and substitutability of heterogeneous worker-job inputs. To estimate production technology parameters we employ variation in our match-specific surplus estimates to instrument for changes in labor factors. In this way, we recover production technology parameters over different time intervals and, then, use them to perform counterfactual exercises that quantify the relative contribution of demand and supply shifts to the distribution of match-specific worker surplus and employment.

2 A Labor Market with Heterogeneous Workers and Jobs

We posit a competitive model of the labor market with two-sided heterogeneity (workers and jobs). Equilibrium sorting reflects the distribution of relative returns, both pecuniary and non-pecuniary.
**Environment.** There are $T$ time periods indexed by $t$ and $M > 1$ separate labor markets indexed by $m$. Each market each year is an independent labor market with its own labor supply and representative firm.

**Workers.** In each market $m$ at time $t$, there is a continuum of workers of size $S_{mt}$. Each worker belongs to one of $I$ demographic groups (types) indexed by $i$. Let $\mu_{imt}$ be the exogenous mass of workers of type $i$, such that $\sum_i \mu_{imt} = S_{mt}$. Workers choose their occupation $j = 1, ..., J$. They can also choose not to work, in which case $j = 0$.

The utility that the workers obtain from each possible choice $j = 0, ..., J$ is the sum of a systematic component that depends on their type, their occupational choice, and the labor market they belong to, $U_{ijmt}$, and an individual unobserved component $\theta_j$ which captures the idiosyncratic preferences for different occupations.

Each worker supplies optimally supplies $h_{ijmt}$ hours of labor paid at the hourly rate $w_{ijmt}$. Workers are hand-to-mouth and consume their income which is given by the sum of their labor income (if any) and their type-specific non-labor income $y_{imt}$ (exogenous). The total systematic utility from working in occupation $j$ is given by

$$U_{ijmt} = \max_{h_{ijmt}} \left( u_c(c_{ijmt}) - u_h^i(h_{ijmt}) + b_{ijt} \right) \quad \text{s.t.} \quad c_{ijmt} = w_{ijmt}h_{ijmt} + y_{imt}$$

where $u_c(\cdot)$ is a standard utility function, $u_h^i(\cdot)$ captures the disutility from working, which we allow to differ across types, and $b_{ijt}$ captures non-pecuniary benefits accruing to a type $i$ worker working in occupation $j$ during period $t$. Conditioning on the specific match, the systematic component of utility from working varies across markets, and within each time period, only by the pecuniary payments (wage and non-labor income), while the non-pecuniary component of utility can change freely across occupations and demographic groups, and over time.

The systematic utility from not working ($j = 0$) is given by $U_{i0mt}(y_{imt}) = u_c(y_{imt}) - u_h(0)$. Hence, there is no equivalent non-pecuniary systematic benefit from not working, namely there is no $b_{i0t}$. Such an element could not be separately identified from the other $b_{ijt}$'s. Not including it is equivalent to normalizing $b_{i0t} = 0$ for all $t$ and all $i$, without loss of generality. Given this normalization, the value of non-employment does not explicitly account for the value of home production. This implies that the value of home production enters as a negative non-pecuniary benefit through $b_{ijt}$. That is, the additively separable $b_{ijt}$ term subsumes, among other things, the value of home production. It is therefore possible to relate variation in estimated $b_{ijt}$ with changes in home productivity.

In addition to the systematic component of utility, workers choosing occupation $j$ receive an individual unobserved component $\theta_j$ which captures the idiosyncratic preferences for dif-
different occupations. We assume that $\theta_j$ is randomly distributed as a type I extreme value with zero location parameter and scale parameter $\sigma_\theta$. Notice that the distribution of the preference shock is independent of time and market.

Given a sequence of realized preference shocks, each worker solves

$$\max_{j=0,1,...,J} U_{ijmt} + \theta_j$$

By the properties of the extreme value distribution, the fraction of workers supplying labor to occupation $j$ (including choosing not to work $j=0$) is then given by

$$\mu_{ijmt} = \frac{\exp \left( \frac{U_{ijmt}}{\sigma_\theta} \right)}{\sum_{j'=0}^J \exp \left( \frac{U_{ij'mt}}{\sigma_\theta} \right)} \mu_{imt}$$

**Firms.** We assume that in each market and time period, there is a representative firm that demands labor in each occupation. We assume that the production technology is the same across markets but changes in time. Let $F_t(\{L_{ijmt}\})$ be the production function, where $L_{ijmt}$ is the total amount of hours demanded in occupation $j$, market $m$, from demographic types $i$, in year $t$. The representative firm solves

$$\max_{\{L_{ijmt}\}_{i=1,...,I; j=1,...,J}} F_t(\{L_{ijmt}\}) - \sum_{i=1}^I \sum_{j=1}^J w_{ijmt} L_{ijmt}$$

this implies that

$$w_{ijmt} = \frac{\partial F_t(\{L_{ijmt}\})}{\partial L_{ijmt}}$$

Let

$$F_t(\{L_{ijmt}\}) = A_t \left[ \sum_j \alpha_{ij} \bar{L}_{jmt}^\rho \right]^{\frac{1}{\rho}}$$

where

$$\bar{L}_{jmt} = \sum_i \beta_{ijt} L_{ijmt}$$

and

$$\sum_j \alpha_{ij} = 1 \quad \forall t$$
This production function implies that within an occupation workers of different types are perfect substitutes (once adjusted for their productivity parameters), and the rate of substitution depends on types and occupations.\textsuperscript{2}

The firm’s first order condition, i.e the inverse labor demand function, is

\[
    w_{ijmt} = \left. \frac{\partial F_t(L_{ijmt})}{\partial L_{ijmt}} \right|_{L_{ijmt}} = A_t \alpha_{jt} \beta_{ijt} \left[ \sum_{j'} \alpha_{j't} \bar{L}_{j'mt}^{\rho} \right]^{\frac{1-\rho}{\rho}} \bar{L}_{j'mt}^{\rho-1}
\]

Equilibrium. A competitive equilibrium in this economy consists of a sequence of wages \(w_{ijmt}\), occupational choices (extensive margin labor supply) \(\mu_{ijmt}\), labor supply choices \(h_{ijmt}\) (intensive margin) and demanded labor \(L_{ijmt}\) such that:

1. given the wages and the realization of the preference shocks each worker solves the maximization problems described in equations (1) and (2);
2. given wages firms optimally demand labor in each occupation according to equation (5);
3. all labor markets clear, that is

\[
    L_{ijmt} = \mu_{ijmt} h_{ijmt}
\]

for all \(i, j, m\) and \(t\).

3 Data and Estimation

Our empirical implementation requires information about the cross-sectional distribution of labor supply and earnings across different demographic and occupation groups. Below we overview our data and estimation approach, after a very brief discussion of how utility and production parameters are identified. A more detailed discussion of identification can be found in Appendix A.1.

Identification of utility parameters. From equation (3), we obtain an expression linking relative employment in each occupation to the pecuniary surplus in that job. The latter is

\textsuperscript{2}This production function is consistent with classical models of heterogeneous human capital. For example, Willis (1986) assumes that while workers are perfect substitutes within occupations, they are imperfect substitutes across occupations “either because they perform different tasks within firms in a given industry or because they enter into the production of different final products which are imperfectly substitutable in consumption”.

5
rescaled by the parameter $\sigma_\theta$ that captures the dispersion of idiosyncratic preferences for different occupations.

\[
\log \left( \frac{\mu_{ijmt}}{\mu_{i0mt}} \right) = \frac{U_{ijmt} - U_{i0mt}}{\sigma_\theta}
\]  

Equation (11) suggests that, under parametric assumptions about the pecuniary utility function, one may use cross sectional variation in employment and wages for different job-worker matches to recover estimates of: (i) non-pecuniary returns $b_{ijt}$ for all occupations $j$, workers $i$ and time periods $t$; and (ii) the scaling parameter $\sigma_\theta$, which measures the cross-sectional dispersion of idiosyncratic preferences for specific occupations.

Identification of technology parameters. Equation (9) delivers an expression linking observable wages to technology parameters. Taking the ratio of wage returns within demographic groups across occupations, and using market clearing conditions, delivers the expression

\[
\frac{w_{ijmt}}{w_{ij'tmt}} = \frac{\alpha_{jt} \beta_{ijt}}{\alpha_{j't} \beta_{ij't}} \left( \frac{\bar{L}_{j'mt}}{\bar{L}_{jmt}} \right)^{1-\rho}
\]

where $\bar{L}_{jmt} = \sum_{i'} \beta_{i'jt} L_{i'jmt}$. As shown in Appendix A.1, worker-specific share parameters $\beta_{ijt}$ are identified by within occupation wage ratios, while the $\alpha_{jt}$ and $\rho$ can be recovered through estimation of the log linear approximation of equation (12).

3.1 Data

We use data from the 1980, 1990, and 2000 decennial Censuses and we pool together three years of the American Community Survey to get samples of comparable size for 2010 (2009-2011) and 2016 (2015-2017) (King et al., 2010).

We consider individuals aged between 25 and 54 and exclude those still in education, as well as workers in farming, forestry, and fishing occupations. We define the supply side heterogeneity, $i$, as a combination of gender, age (three groups: 25-34, 35-44, 45-55), and education (college graduates and above, and less than college). This implies that we have 12 different demographic groups. On the demand side, we consider a set of 13 occupations (to which we add the non-employment status), indexed by $j$. The list of occupations is shown in Table 1, along with the aggregation to four broad task clusters as in Acemoglu and Autor (2011). We also consider four geographical markets, indexed by $m$, corresponding to the standard Census regions (Northwest, Midwest, South, and West).

For each $(i,j,m,t)$ cell we compute total employment, mean wage rate, mean hours worked and, non-labor income. For total employment, we use the population weights. A worker is counted as employed if they report working at least 15 hours per week. Non-labor income is
Table 1: Occupational groupings used for the estimation of the model.

<table>
<thead>
<tr>
<th>No.</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Executive, Administrative, and Managerial</td>
</tr>
<tr>
<td>2</td>
<td>Management Related</td>
</tr>
<tr>
<td>3</td>
<td>Professional Specialty</td>
</tr>
<tr>
<td>4</td>
<td>Technicians and Related Support</td>
</tr>
<tr>
<td>5</td>
<td>Sales</td>
</tr>
<tr>
<td>6</td>
<td>Administrative Support</td>
</tr>
<tr>
<td>7</td>
<td>Protective Service</td>
</tr>
<tr>
<td>8</td>
<td>Other Service</td>
</tr>
<tr>
<td>9</td>
<td>Mechanics and Repairers</td>
</tr>
<tr>
<td>10</td>
<td>Construction Trades</td>
</tr>
<tr>
<td>11</td>
<td>Precision Production</td>
</tr>
<tr>
<td>12</td>
<td>Machine Operators, Assemblers, and Inspectors</td>
</tr>
<tr>
<td>13</td>
<td>Transportation and Material Moving</td>
</tr>
</tbody>
</table>

Managerial, Professional Specialty and Technical (Non-Routine Cognitive)

Sales and Administrative Support (Routine Cognitive)

Service (Non-Routine Manual)

Precision Production, Craft, Repair, Operators, Fabricators, and Laborers (Routine Manual)
obtained as the sum of incomes from businesses and farms.

3.2 Estimation

Our estimation approach follows a two-step procedure. We first recover non-pecuniary returns for different worker-job matches. Then, we estimate production shares and elasticity of substitution between different labor inputs.

Estimation of utility parameters. We use information about relative employment, wages and hours supplied in different worker-job matches to estimate the shape of the utility from consumption $u_c(\cdot)$ and the disutility from hours worked $u_h^i(\cdot)$, the dispersion of the idiosyncratic preference shock $\sigma_\theta$ and the non-pecuniary returns $b_{ijt}$. Through a GMM procedure, we estimate parameters to match as closely as possible the occupational shares observed in the data and the average hours of work supplied by individuals in each observation cell, given pecuniary returns to employment and non-labor income.

In practice, let $w$ and $y$ be the vector of all the labor and non-labor earnings as obtained from the data. From the first order condition associated to the maximization problem in equation (1) we compute the model-implied supply of hours and, using equation (11), we compute all the model-implied occupation shares (relative to non-employment) $\bar{\mu}_{ijmt}$ Let $m_{model} (\beta, w, y)$ be a column vector containing all the predicted shares and hours supplied, where $\beta$ is the vector of parameters to be estimated. Also, let $m_{data}$ be the empirical counterpart of $m_{model}$. Our estimator is given by

$$\hat{\Omega} = \arg \min_{\Omega} [m_{model} (\Omega, w, y) - m_{data}]^T W [m_{model} (\Omega, w, y) - m_{data}]$$

(13)

where $W$ is a positive definite weighting matrix.

For the implementation we assume the following functional forms for the utility of consumption and the disutility from work

$$u_c(c) = \frac{c^{1-\sigma} - 1}{1 - \sigma} \hspace{1cm} u_h^i (h) = \psi \frac{h^{1-\gamma_i}}{1 - \gamma_i}$$

(14)

where $\sigma > 0$ and $\gamma_i < 1$ are to be estimated.

Given the characterization of the heterogeneity on both sides of the labor market and the presence of four markets and five time periods, we have a total of 6,240 moments used to estimate 916 parameters. The estimated values for the utility of consumption and the

\footnote{In practice $W$ is a diagonal matrix with ones in correspondence of moments relating to the employment shares and $\sqrt{0.001}$ in correspondence of hours moments. This adjustment is necessary to avoid the latter to dominate the optimization process given the difference in the magnitudes of the two set of moments.}

\footnote{Notice that for $\gamma_i < 1$, $u_h^i (0) = 0$.}
dispersion of the idiosyncratic preference shock are reported in Table 2. Details about the exact moments used in the estimation and the identification of the parameters are in the Appendix. The remaining 913 estimates of the non-pecuniary benefits $b_{ijt}$ and of the shape of the disutility from working are not reported for obvious space reasons. They are all highly significant and we discuss them in further detail below.

**Estimation of production parameters.** The estimation of the production function closely follows the identification procedure described in detail in appendix A. In the appendix, we show that the $\beta$ parameters are identified up to a normalization by the within-occupation ratio of wages for two different occupation groups. This follows from the fact that within an occupation, workers from different demographic groups are perfect substitutes. Thus, after having normalized $\beta_{1jt} = 1$ for all $j = 1, ..., J$ and all $t$, we can estimate all the remaining $\beta$’s as the average wage ratio across markets:

$$\hat{\beta}_{ijt} = \frac{1}{M} \sum_{m=1}^{M} \frac{w_{ijmt}}{w_{1jmt}}$$  \hfill (15)

Estimated values of the remaining parameters are obtained through the estimation of the empirical counterpart of equation (32) in appendix A, reported here for convenience

$$\log \left( \frac{w_{ijmt}}{w_{1jmt}} \right) = \log \left( \frac{\alpha_{jt}}{\alpha_{1t}} \right) + \log \left( \frac{\beta_{ijt}}{\beta_{i1t}} \right) + (\rho - 1) \log \left( \frac{\sum_{j'} \beta_{ij't} L_{ij't} m_{j't}}{\sum_{j'} \beta_{i1't} L_{i1't} m_{i1't}} \right)$$  \hfill (16)

Ideally, in the empirical specification, we would include a full set of time-specific occupation dummies to estimate the $\alpha$ parameters fully flexibly. In practice, this substantially increases the number of parameters to be estimated and delivers very imprecise estimates. For this reason, we restrict $\frac{\alpha_{jt}}{\alpha_{1t}}$ to follow an occupation-specific quadratic trend. The values for each $\alpha_{jt}$ are recovered using the restriction that $\sum_{j} \alpha_{ij} = 1$ for all $t$.

The empirical counterpart of the second term is simply

$$\hat{Z}_{ijt} = \log \left( \frac{\hat{\beta}_{ijt}}{\hat{\beta}_{i1t}} \right)$$  \hfill (17)

and we expect the estimated coefficient on this term to equal one.
Finally, the empirical counterpart of the third term is

\[ \hat{X}_{jmt} = \log \left( \frac{\sum_{i'} \hat{\beta}_{i'jt} h_{i'jt} h_{i'jmt}}{\sum_{i'} \hat{\beta}_{i'j1t} h_{i'j1t} h_{i'j1mt}} \right) \] (18)

This quantity measures the relative supply of labor efficiency units to occupation \( j \). The estimated value of \( \rho \) is then recovered as \( \hat{\rho} = \hat{\phi} + 1 \).

To summarize our regression equation is

\[ W_{ijmt} = \gamma_{0j} + \gamma_{1jt} t + \gamma_{2jt} t^2 + \psi \hat{Z}_{ijt} + \phi \hat{X}_{jmt} + \epsilon_{ijmt} \] (19)

where \( W_{ijmt} = \log \left( \frac{w_{ijmt}}{w_{i1mt}} \right) \).

The latter specification exhibits a well-known endogeneity problem arising from the simultaneous determination of prices and quantities. To correct for this endogeneity we use two instruments for \( \hat{X}_{jmt} \), namely the lagged market-specific employment-weighted average wage for a given occupation, \( W_{jmt} \), and the change in the employment-weighted average of occupation-specific non-pecuniary returns to working, \( \Delta \hat{b}_{jt} \). The validity of the latter instrument is suggested by our theoretical restrictions. Estimates of heterogeneous non-pecuniary surplus across occupations (\( \hat{b}_{ij} \)) can be considered as instruments that induce non-technical labor supply shifts. By definition, non-pecuniary surplus exogenously changes labor supply by different demographic groups without directly affecting labor demand driven by technology parameters, making the estimation of (33) feasible.

Table 3 reports the estimated coefficients on \( \hat{X}_{jmt} \) and \( \hat{Z}_{ijt} \) for different specifications with bootstrapped standard errors for all specifications. Column 1 reports the OLS estimates. Comparing the OLS estimates to the IV estimates in the other columns, we immediately see the effects of the positive bias coming from the endogeneity of \( \hat{\phi} \). Moreover, the implied value of \( \rho \) is above one which is not acceptable\(^5\). Columns 2 and 3 report the estimation results obtained instrumenting \( \hat{X}_{ijmt} \) with the change in non-pecuniary returns to working in occupation \( j \) and the lagged average wage respectively. The two identify a range of values for \( \rho \) that correspond to an elasticity of substitution between occupations between 2.42 and 2.95. Our preferred specification is reported in column 4 where we use both our instruments to control for endogeneity. The implied value for \( \rho \) is 0.59 which corresponds to an elasticity of substitution of 2.46. We also report the p-value of the overidentification test (Sargan, 1958) that shows that we cannot reject the null hypothesis of the validity of the instruments at the conventional significance levels. Finally, it is worth noting that, as expected, the estimated coefficient on \( Z_{ijt} \) is never significantly different from 1. For space reasons, we do not individually report the 715 estimates of \( \beta_{ijt} \) and the 65 estimates of \( \alpha_{jt} \). These are all strongly significant and available upon request.

\(^5\)In the CES production function \( \rho \in (-\infty, 1] \).
### Table 3: Estimation results.

<table>
<thead>
<tr>
<th></th>
<th>OLS</th>
<th>IV</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>$\hat{\phi}$</td>
<td>0.110***</td>
<td>-0.413***</td>
</tr>
<tr>
<td></td>
<td>(0.011)</td>
<td>(0.137)</td>
</tr>
<tr>
<td>$\hat{\psi}$</td>
<td>0.992***</td>
<td>0.992***</td>
</tr>
<tr>
<td></td>
<td>(0.009)</td>
<td>(0.014)</td>
</tr>
</tbody>
</table>

Instrumental Variables

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta b_{jt}$</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>$W_{jmt-1}$</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>OverId p-val</td>
<td></td>
<td></td>
<td></td>
<td>0.944</td>
</tr>
<tr>
<td>Observations</td>
<td>3,120</td>
<td>2,496</td>
<td>2,496</td>
<td>2,496</td>
</tr>
</tbody>
</table>

Standard errors in parentheses

*** p < 0.01, ** p < 0.05, * p < 0.1

4 The Distribution of Employment Surplus

In its most basic form, our approach allows us to characterize the distribution of total worker surplus at different points in time. This distribution summarizes the total value of a job (relative to non-employment) for each subset of workers. The job surplus reflects both pecuniary and non-pecuniary returns.

The top panel of Figure 1 plots the distribution of the total (pecuniary and non-pecuniary) systematic worker surplus, across demographics and occupations, in three different years (1980, 2000, and 2016). It is apparent that the distribution of the systematic component of worker surplus has become more concentrated over time, with the probability mass of its left tail shifting noticeably to the right. This suggests that certain types of $i,j$ matches that were less frequent in the 1980s have become extremely common by 2016.

The latter observation can be better qualified by looking at more disaggregated distributions of the worker-job surplus. The bottom panel of Figure 1 separately reports the distributions of total surplus for gender-education groups. This more disaggregated view shows that the rightward shift of the left tail is mostly driven by growing surpluses among women. In fact, the distribution of total surpluses among women has become over time more similar to that of men. In contrast, the distribution of job surpluses for non-college men seems to have shifted to the left, suggesting that many of the matches involving less educated male workers are characterized by lower wages and/or worse non-pecuniary returns.

The mechanics of the changing distribution of worker surplus over time can be further examined by separately looking at the distribution of its pecuniary and non-pecuniary components. Figure 2 plots the density of total surplus, and of its three components, for different
Figure 1: Distribution of surplus: aggregate (top panel) and disaggregated (bottom four panels).
Figure 2: Distribution of surplus: this figure reports, for different years, the density functions of (i) total systematic surplus across worker-job matches; (ii) its pecuniary surplus component; (iii) its component concerning the disutility from hours worked; (iv) its non-pecuniary surplus component.
years. This makes it clear that (i) the density of pecuniary rewards lies to the right of total surplus density, while non-pecuniary rewards tend to be lower; (ii) pecuniary surplus is much more concentrated than non-pecuniary surplus; (iii) over time the distribution of the pecuniary component of surplus has become more disperse while (iv) the other two components have become more concentrated.

These facts convey information about match-specific shifts in the relative value of money, time, and other job-specific attributes. We re-examine these questions in some detail after estimating production technology parameters.

4.1 The Components of Worker Surplus

One key question is what accounts for the total surplus of specific worker-job matches. Specifically, one might ask how large are monetary rewards net of the disutility from working (“net-pecuniary”), relative to systematic non-pecuniary benefits, as well as whether different types of returns are positively or negatively related.

Figure 3 reports the weighted average of the absolute values of net-pecuniary and non-pecuniary rewards (in utils) within bins defined by gender, broad occupation, and year. The patterns over time are quite different for male and female workers: while men experience an increase in the absolute returns for both net-pecuniary (positive) and non-pecuniary (negative) components of the surplus, with these changes being stronger in cognitive occupations, women have a more nuanced set of changes. Specifically, the net-pecuniary surplus does not change as strongly, partly because of the underlying change in the value of non-employment. However, the non-pecuniary returns improve significantly in all non-routine occupations and worsen in routine manual occupations. Non-pecuniary returns stay almost unchanged for women in routine cognitive occupations.

A different way to gauge the composition of labor market rewards is presented in Figures 4 and 5, which reports scatter plots of net-pecuniary (average across markets) vs non-pecuniary surplus for each $i,j$ pair (demographic $i$ and occupation $j$ combination) for cognitive and non-cognitive occupations respectively.

With 12 demographic groups (defined by education, gender, and age) and 13 occupation categories, we have 156 observations per year. This allows us to draw some inference about the covariation of different components of total surplus: fitted lines in the scatter plots capture the linear relationship between pecuniary and non-pecuniary returns in different years. Negative comovement between money and non-money rewards would be consistent with the standard view of compensating differentials. Figure 4 shows that the latter relation is clearly detectable and stable over time among cognitive occupations for all the demographic groups, except for male college graduates. For the latter, we observe the compensating differential relation arising over time. In 1980 there appears to be little covariation at all between net-pecuniary
Figure 3: Net-pecuniary (pecuniary net of disutility from working) and non-pecuniary surplus by gender, occupation and year with bootstrapped confidence intervals. All values are expressed in utility equivalents (higher value corresponds to higher utility).
and non-pecuniary returns. A negative relation emerges, along with the compression of the distribution of non-pecuniary returns in 2000 and 2016.

Similarly, Figure 5 shows that there is a negative correlation between the two components of the surplus for all the demographic subgroups. This relation is overall steeper for men than it is for women.

4.2 The Structure of Worker-Job Matches over Time

The size of each job's surplus is closely related to the prevalence of an occupation in equilibrium. While the surplus is likely to change in response to higher or lower worker supply in a specific occupation, changes in the nature of work (e.g., work schedules, total hours worked, work environment, location, and amenities), as well as technological progress, may induce substantial shifts in the structure of employment and the distribution of worker-job matches.

The graphs in Figure 6 plot the relative frequency of employment (the size of each bubble) for different occupations: the plot is in the space of net-pecuniary (y-axis) and non-pecuniary (x-axis) payoffs. That is, each job corresponds to a different combination of worker surplus, with the size of the bubble indicating how many workers in a given year populate a particular occupation. For clarity, color-code our 13 occupations to highlight the four broader categories (non-routine cognitive, routine cognitive, non-routine manual, and routine manual).

Several observations can be made from these plots: first, the pecuniary surplus has grown for both genders, but proportionally more for men. This is consistent with the findings in the bar charts of Figure 3. This gender discrepancy is partly driven by changes in the pecuniary value of non-employment. Second, women have a significantly higher representation in non-routine and cognitive occupations. Third, the distribution of non-pecuniary surplus appears to have worsened for men while improving for women in most occupations.

These graphs relate, in a concise format, the changing distribution of surplus with the significant shifts in employment that have occurred over the past four decades.

4.3 Taking Stock

The evidence presented so far indicates that a significant amount of heterogeneity exists in the match value of jobs to workers. This heterogeneity reflects both pay and non-pecuniary surplus. The split between these components is rather uneven across different demographic-occupation matches. Moreover, the relative magnitude of specific returns also changed over time, occasionally in unexpected ways. The following observations summarize our basic findings so far:

1. The distribution of surplus has changed for all workers since the 1980s. This aggregate shift, with average surplus becoming higher over time, masks vast heterogeneity.
Figure 4: Each panel plots the distribution of net-pecuniary and non-pecuniary returns over the set of cognitive occupations within an education-gender group. Data points are color-coded to distinguish between different years.
Figure 5: Each panel plots the distribution of net-pecuniary and non-pecuniary returns over the set of non-cognitive occupations within an education-gender group. Data points are color-coded to distinguish between different years.
Figure 6: Relative frequency of employment (bubble size) for different occupation groups, by gender and year. The y-axis measures net-pecuniary surplus and the x-axis measures non-pecuniary surplus for each occupation category.
2. The pecuniary surplus has grown more for men than it did for women; this shift does not suggest that pay went up more strongly for men; rather, the pecuniary premium relative to non-employment increased for men more than it did for women.

3. At the same time, the non-pecuniary surplus worsened (it became more negative) for men, while it went in the opposite direction for most women. As a consequence, gains associated with employment for men have become more dependent on monetary rewards, since the latter provide compensation for a worsening non-pecuniary surplus. For women, instead, increased participation aligns more closely with improvements in non-pecuniary surplus.

4. Compensating differentials (negative covariation of pecuniary and non-pecuniary surplus) are detected for almost all workers in non-college jobs. This relation has been fairly stable in time for most demographic groups in most of the occupations. The only exception is observed for college men in cognitive occupations for which we detect an increase of the intensity of this relationship over time.

5. Non-pecuniary returns are more dispersed in the population than pecuniary ones; this is true for all genders, educations, and time periods.

6. The structure of employment (the frequency of different worker-job matches) has also changed significantly over different time periods. In particular, women exhibit increasing participation in managerial and non-routine positions and are able to extract a higher surplus from these jobs.

Some of these observations are the outcome of the equilibrium interaction of demand for, and supply of, different worker-occupation inputs. We examine the mechanics of these interactions using a production technology that aggregates labor inputs from heterogeneous occupation and demographic groups.

5 Production Share Estimates

Changes in the pecuniary component of workers’ surplus depend on the relative productivity of their labor input. In turn, the productivity of each worker-job pair is the by-product of both supply of and demand for their labor input. From wage equation (9), we know that the marginal productivity of a type $i$ worker in occupation $j$ at time $t$ is proportional to $\alpha_{jt}\beta_{ijt}$. Figure 7 shows the evolution of an employment-weighted average of the labor loadings $\alpha_{jt}\beta_{ijt}$ for four major occupation categories (the left panel reports levels, while the right one shows growth relative to the 1980 baseline). All occupation types have experienced growth in their production shares with the exception of Routine Manual occupations, for which we observe a
steady decline reaching -18% in 2016 relative to 1980. The occupation types exhibiting growth show a similar pattern during the 1980s (+15%); however, their paths started diverging in the 1990s with non-routine cognitive jobs the only group sustaining a steady growth rate until the turn of the century after which we observe an increase in the rate of growth. This resulted in a staggering jump in the production shares of non-routine cognitive occupations between 1980 and 2016, adding up to about 62%. Non-routine manual occupations and routine cognitive occupations, while exhibiting lower increments, also experienced significant long-term growth (respectively of +36% and +30% over the sample period). The fanning out of the production shares is a key determinant of pecuniary surplus and employment across different demographic groups and, as we show below, plays an important role for the striking labor market changes of the past decades.

Production shares by type of worker. To zoom into the evolution of production shares for heterogeneous labor inputs, Figure 8 shows the counterparts of the left panel in Figure 7 for different demographic groups. The fall in the aggregate technology share of Routine Manual occupations is almost entirely due to the falling productivity of non-college men; the loading of women in these occupations has remained roughly constant while the college-educated rarely work in these occupations.

Despite being lower in levels, in relative terms production shares for men and women have increased in a similar fashion for non-routine cognitive occupations (+65%) while the growth has been higher for women in routine cognitive occupations (+61% vs. +40%) but lower in non-routine manual occupations (+42% vs. +49%). Interestingly, the productivity of college men has declined in routine manual occupations (-10%), while college women display a slight increase in productivity for the same (+10%) occupation type. Among non-college workers,
Figure 8: Weighted average production shares \((\alpha_{jt}\beta_{ijt})\) of major occupation groups for different demographics.
the growth in productivity of women has been higher in all occupation types.

Figure 8 highlights the productivity share premium of cognitive over manual occupations. This gap is much wider for college graduates. Perhaps not surprisingly, a college degree does not improve the productivity loading when working in manual occupations.

6 Equilibrium and Counterfactuals

To establish the quantitative importance of different forces for the large structural changes observed in the labor market, one needs to account for the way employment and returns are determined in equilibrium. Given parameter estimates, we compute equilibrium prices and quantities for different time periods. These reflect changes in both technology and non-pecuniary returns from jobs. Then, by holding constant the value of selected parameters, we leverage this structure to perform counterfactual exercises and quantitative comparisons.

Prices and quantities: model vs data. One can compute all prices and quantities in our model. In Figure 9 we compare data on the average wages and employment within each \((i,j)\) cell (demographic-occupation pair) to their counterparts obtained by solving the equilibrium of the model for each market and year.
Model-generated prices and quantities closely match data observations, accounting for, respectively, 98.8%, 93.9%, and 70.1% of total variation in employment, wage rates, and hours worked. The tight correspondence between model output and data is reassuring since the equilibrium restrictions from Section 2, while implicitly assumed, are not explicitly targeted in the estimation.

6.1 Counterfactual Analysis

We perform two sets of numerical exercises. In the first set, we explore how labor supply and wages would have changed if the value of non-pecuniary returns to working had stayed at its 1980 levels. In the second, we examine the impact of non-technology shifters; specifically, we compute the counterfactual employment and wages had production function parameters stayed the same as in 1980.

In each set of experiments, we perform partial equilibrium exercises. For employment, we compute counterfactual scenarios keeping wages at 1980 levels. By comparing the latter to the ones prevailing when prices are free to adjust, we quantify the impact of equilibrium adjustments on both the level and composition of labor demand and employment. In a related set of experiments for wages, we consider the partial equilibrium effects of changing technology parameters while holding labor supply at its 1980 levels.
Employment. Figure 10 reports the labor force participation (share) of four demographic groups defined by gender and education. The blue lines show the historical evolution of labor force participation. Overall participation has declined for men during the 1980-2016 period; the drop was small for college-educated men (-2 percentage points) and more substantial for the less educated (-7 percentage points). Among women, instead, both high (+13) and low (+12) education individuals increased their labor force participation.

Our counterfactual experiments reveal interesting aspects of these changes. The red lines in each of the four panels show the counterfactual scenarios in which we keep non-pecuniary returns at their 1980 level, while the yellow lines refer to experiments in which we keep technology constant. It is clear that most of the change in the labor force participation of men is explained by changes in non-pecuniary returns. In fact, had these returns stayed the same, the participation of both high and low educated men would be higher by, respectively, 5 and 7 percentage points. Technological change played a different role depending on the education level. For college-educated men the yellow line lies below the blue line suggesting that technological change partially offset the negative impact of changes in non-pecuniary returns on labor force participation. The opposite is true for non-college men.

Both the change in non-pecuniary returns and technology have supported the growing labor force participation of women. For low-education women non-pecuniary returns take the lion’s share, explaining most of the observed change in participation (the red line is always below the yellow). For college-educated women, changes in non-pecuniary returns are the main driver of the increase in participation until the early 2000s; after that year, technological change becomes the most important factor.

Finally, the dashed lines report the partial equilibrium effects of changes in non-pecuniary returns and are obtained by holding wages at their 1980 level. In all four quadrants, the dashed lines almost perfectly overlap with the yellow ones, suggesting that price adjustments due to general equilibrium effects play a negligible role once we account for technological change. This suggests that most of the price adjustments over time are due to changes in production shares, rather than changes in input quantities.

Occupation Groups. In Figure 11, we perform the same counterfactual to assess the role of technology and non-pecuniary returns for the evolution of employment across the four major occupation groups defined in Table 1. The blue line in the top-left panel shows the well-documented increase in the fraction of the population employed in non-routine cognitive (NRC) occupations. From 1980 to 2016 this fraction has climbed about 8 percentage points. Contrasting this pattern with the yellow line, which shows the counterfactual scenario

---

6For easier comparability, these quantities are obtained by simulating data from the full model. The high predictive power of the model implies that the simulated histories are almost identical to actual data.

7Technological changes subsume possible changes in wage discrimination as in Hsieh et al. (2019).
Figure 11: Evolution of the shares of the population employed in four major occupation groups in the baseline model (which replicates the data) and in the counterfactual scenarios.

where production parameters are held at their 1980 values, we can gauge the role played by technological progress: without the latter, we would observe essentially no increase in NRC employment.

The dashed line reports the partial equilibrium counterfactual where we keep wages at their 1980 levels; this should be compared to the yellow line, as in both cases the effects of technological change are muted. The difference between dashed and yellow lines is that the latter allows for price adjustments in response to labor supply changes. It is apparent that the effects of price responses in negligible for the observed growth in the share of population employed in NRC occupations. The red line, describing the counterfactual scenario in which non-pecuniary benefits are kept at their 1980 levels, shows a different picture. Since the turn of the century, non-pecuniary returns have become a drag for NRC employment. Without the change in non-pecuniary returns, the fraction of the population employed in NRC occupations would have been almost 4.5 percentage points higher in 2016.

The top-right panel performs the same exercises for routine cognitive occupations. Overall, we observe a slight decline in the fraction of the population employed in such occupations (-1 percentage point), but the pattern is not monotonic as we initially observe an increase of about 2 percentage points between 1980 and 1990. The counterfactual experiments show that both technological changes and changes to non-pecuniary returns have contributed to the historical trajectory. Nevertheless, by 2016 the effects of changes in technology, entirely disappear for this group.

Non-routine manual occupations are shown in the bottom-left panel. The share of non-
routine manual workers has increased steadily in our sample moving from 6.7% to 11.5%. Both changes to technology and non-pecuniary returns have played a similar role in this growth. Unlike for non-routine and routine cognitive occupations, general equilibrium effects have played a role in the observed changes, reducing the growth in the share of the population working in non-routine manual occupations.

Finally, the bottom-right panel shows the results for routine manual occupations. Both technology and non-pecuniary returns have contributed to the observed 6 percentage point fall in the fraction of workers in routine manual occupations. Technology had a much stronger impact over the whole sample period. Interestingly, the difference between the dashed and the yellow line suggests that general equilibrium effects mitigate the negative impact of technological change on the share of the population working in routine manual occupations: as workers flow out of these jobs, marginal returns tend to go up and slow down the outflows.

To summarize, we observe that technological change has been the main driver of the employment boom in NRC jobs and of the employment bust in RM jobs. In NRM occupations a big part of the changes can be attributed to non-pecuniary returns. Finally, in RC jobs there is little change overall, but technological change still plays the biggest role.

Wages. Figure 12 shows both actual and counterfactual changes in the average wage rate of different demographic groups, following the same color coding used in Figure 11. Between 1980 and 2016 wage income experienced a massive increase for college graduates; this is almost entirely driven by technological change. The equilibrium effects originating from changes in
non-pecuniary returns are relatively small and have opposite implications for men and women. Changes in non-pecuniary returns have essentially no implications for college-educated women while they caused an increase in the hourly wage rate of college-educated men of about 1$.

For non-college men, wages have been mostly stagnant in the first half of the sample and declining afterwards. Neither technology nor non-pecuniary returns have played a predominant role in determining this trajectory but both have contributed to the drop observed after 2000. Interestingly, the equilibrium effects, as captured by the differences in the dashed purple line and the yellow line, have helped to reduce the drop in the wages of this demographic group.

The bottom right panel shows that low-educated women have experienced a substantial increase in their average labor income in the first half of the sample and a slight decline afterwards; this is almost entirely explained by technological changes. As for non-college men, general equilibrium effects have contributed positively to the wage rates paid to non-college women.

7 Discussion and Conclusions

Striking labor market shifts, often occurring within the span of a decade, have been documented since the 1980s. From the point of view of workers, changes in labor market structure are embodied in the pecuniary and non-pecuniary surplus derived from different job matches. In this paper, we characterize the structure of employment and wages in the US labor market as the equilibrium outcome of the interaction between technological progress and changes in the non-pecuniary value of jobs for different workers. The latter changes capture a variety of elements that have reshaped the nature of work, contributing to structural changes in the distribution of individual labor market outcomes.

In the first part of the paper, we develop a simple procedure to estimate the match-specific surplus associated with alternative types of employment and show that similar jobs are valued very differently by different sets of workers.

Heterogeneity in the values attached to similar occupations reflects measurable heterogeneity in the way individuals regard monetary and non-monetary returns from employment. Interestingly, workers don’t always trade-off pecuniary and non-pecuniary surplus in their jobs. When we explore the hypothesis of compensating differentials (negative covariation of pecuniary and non-pecuniary surplus), we find that these trade-offs can be detected among almost all workers; the magnitude of this relationship is very heterogeneous across demographic groups but essentially stable for all of them except for college graduates in cognitive occupations. For the latter, a negative relation between the two quantities seems to have emerged over time.

Our analysis suggests that shifts in non-pecuniary surplus are key to account for changes
in the total surplus that workers derive from employment. Men experienced a severe deterioration of their non-pecuniary surplus on average; this is in contrast to women who enjoyed a significant improvement. The pecuniary surplus increased for both groups.

In the second part of the paper, we posit a production technology describing the aggregation of different labor factors and estimate its parameters, separately, over different decades. To by-pass the endogeneity of prices and quantities in the estimation of technology parameters, we employ moment restrictions based on (pre-determined) values of the non-pecuniary worker surplus estimated in the first part of the paper.

Having recovered production technology parameters for different decades, we explore the quantitative contribution of specific demand and supply forces to the observed shifts in employment and wages. By contrasting the impact of technological change to that of heterogeneous job valuations, these counterfactual exercises shed light on the mechanics of structural change in the labor market and provide a way to rationalize some of the observations described in the first part of the paper.

Counterfactual exercises confirm that non-pecuniary factors played a central role in the employment patterns of men. Had these stayed at their 1980 levels, the participation of both high and low educated men would be much higher in 2016. Technological change played a smaller role, partially offsetting the negative impact of changes in non-pecuniary returns on the labor force participation of college-educated men while further reducing the participation of non-college men in the second half of our sample.

The story for women is quite different as changes in both non-pecuniary returns and technology have bolstered their labor force participation. For low-education women, non-pecuniary returns and technological change play a similar role in explaining the increase in participation. For college-educated women, changes in non-pecuniary returns are an important driver of the increase in participation until the early 2000s, while technological change becomes the most important factor after that year.

Finally, we document that while employment patterns are the by-product of both technology and preferences, the evolution of wages is almost entirely explained by technological progress. This is true across all gender and education groups.
References


A Identification and Estimation

This section discusses the identification and estimation of model parameters and provides an overview of the empirical analysis.

A.1 Identification of Utility and Technology Parameters

To show the identification of the structural parameters, first consider the time-consumption problem described in equation (1). With the assumed functional forms, the problem becomes

\[
U_{ijmt} = \max_{h_{ijmt}} \frac{c_{ijmt}^{1-\sigma}}{1-\sigma} (c_{ijmt}) - \psi \frac{h_{ijmt}^{1-\gamma_i}}{1-\gamma_i} (h_{ijmt}) + b_{ijmt}
\]

\[\text{s.t. } c_{ijmt} = w_{ijmt} h_{ijmt} + y_{imt}\] (20)

the associated first order condition in logarithmic form is

\[
\log (h_{ijmt}) = -\frac{1}{\sigma - \gamma_i} \log (\psi) + \frac{1-\sigma}{\sigma - \gamma_i} \log (w_{ijmt})
\] (21)

The empirical counterpart of this is

\[
\log (h_{ijmt}) = \alpha_i + \beta_i \log (w_{ijmt}) + \epsilon_{ijmt}
\] (22)

with

\[
\alpha_i = -\frac{1}{\sigma - \gamma_i} \log (\psi) \quad \beta_i = \frac{1-\sigma}{\sigma - \gamma_i}
\] (23)

From the estimation\(^8\) of the latter equation we can obtain \(\gamma_i\) and \(\psi\) as a function of \(\sigma\):

\[
\gamma_i = \sigma - \frac{1-\sigma}{\beta_i} \quad \psi = \exp \left( -\frac{1-\sigma}{\beta_i} \alpha_i \right)
\] (24)

Substituting \(\gamma_i\) and \(\psi\) in the FOC we can write the optimal hours as a function of the wage rate and \(\sigma\): \(h_{ijmt} = g_i (w_{ijmt}; \sigma)\).

From equation (3) in the main text we have that

\[
\log \left( \frac{\mu_{ijmt}}{\mu_{\emptyset mt}} \right) = \frac{U_{ijmt} - U_{\emptyset mt}}{\sigma \theta}
\] (25)

where

\[
U_{ijmt} = u_c (w_{ijmt} g_i (w_{ijmt}; \sigma)) - u_l (g_i (w_{ijmt}; \sigma)) + b_{ijt}
\] (26)

\(^8\)The theory also provides a testable restriction, namely \(\frac{\alpha_i}{\beta_i} = \kappa = \frac{\log (\psi)}{\sigma - 1}\) for all \(i\).
and

$$U_{i0mt} = u_c(y_{imt})$$  \hspace{1cm} (27)$$

Taking the derivative of (25) with respect to $w_{ijmt}$ gives

$$\frac{\partial \log \left( \frac{u_{ijmt}}{u_{i0mt}} \right)}{\partial w_{ijmt}} = \frac{u_c(w_{ijmt}g_i(w_{ijmt};\sigma)) [g_i(w_{ijmt};\sigma) + w_{ijmt}g_i'(w_{ijmt};\sigma)] - u_t(g_i(w_{ijmt};\sigma))g_i'(w_{ijmt};\sigma)}{\sigma_\theta}$$  \hspace{1cm} (28)$$

which, taking the ratio across markets, gives

$$\frac{\partial \log \left( \frac{u_{ijmt}}{u_{i0mt}} \right)}{\partial w_{ijmt}} = \frac{u_c(w_{ijmt}g_i(w_{ijmt};\sigma)) [g_i(w_{ijmt};\sigma) + w_{ijmt}g_i'(w_{ijmt};\sigma)] - u_t(g_i(w_{ijmt};\sigma))g_i'(w_{ijmt};\sigma)}{u_c(w_{ijmt}g_i'(w_{ijmt};\sigma) + w_{ijmt}g_i'(w_{ijmt};\sigma))}$$ \hspace{1cm} (29)$$

that is a nonlinear equation in one unknown. Solving this gives $\sigma$. Once $\sigma$ is known, eq. (28) gives the identification of $\sigma_\theta$. Finally, $b_{ijt}$ is identified by equation (25).

**Identification of production function parameters.**  On the firm side, taking the ratio between the wages for two demographic groups within an occupation (eq. (9)), we have that

$$\frac{w_{ijmt}}{w_{ijmt'}} = \frac{\beta_{ijt}}{\beta'_{ijt}}$$  \hspace{1cm} (30)$$

which shows that the $\beta$’s are directly identifiable from wage data as long as we normalize the value of the $\beta$’s for one demographic group (e.g. setting $\beta_{1jt} = 1$ for all $j$ and $t$). Taking a similar ratio within demographic groups across occupations and using market clearing gives

$$\frac{w_{ijmt}}{w_{ij'mt'}} = \frac{\alpha_{jt}\beta_{ijt}}{\alpha_{j't}\beta'_{ij't}} \left( \frac{L_{j'mt}}{L_{jmt}} \right)^{1-\rho} = \frac{\alpha_{jt}\beta_{ijt}}{\alpha_{j't}\beta'_{ij't}} \left( \frac{\sum_{j'}\beta_{ij't}L_{ij'mt}}{\sum_{j'}\beta'_{ij't}L_{ij'mt}} \right)^{1-\rho}$$  \hspace{1cm} (31)$$

Once we know the $\beta$’s, we can identify the $\alpha$’s (up to a normalization) and $\rho$’s as follows. Taking the log of eq. (31) for $j' = 1$ gives

$$\log \left( \frac{w_{ijmt}}{w_{i1mt}} \right) = \log \left( \frac{\alpha_{jt}}{\alpha_{1't}} \right) + \log \left( \frac{\beta_{ijt}}{\beta_{1't}} \right) + (\rho - 1) \log \left( \frac{\sum_{j'}\beta_{ij't}L_{ij'mt}}{\sum_{j'}\beta'_{ij't}L_{ij'mt}} \right)$$  \hspace{1cm} (32)$$
Since, at this point, the $\beta$’s are known, one can compute $X_{jmt} = \log \left( \frac{\sum_i \beta_{ijt} L_{i1mt}}{\sum_i \beta_{ij1t} L_{ij1mt}} \right)$, $Z_{ijt} = \frac{\beta_{ijt}}{\beta_{1jt}}$ and $W_{ijmt} = \log \left( \frac{w_{ijmt}}{w_{i1mt}} \right)$ and regress the latter on $X_{jmt}$ and a set of occupation dummies $\gamma$, separately for each year:

$$W_{ijmt} = \gamma_{jt} + \psi Z_{ijt} + \phi X_{jmt} + \epsilon_{ijmt}$$  (33)

Then the $\alpha$’s are identified by $\frac{\alpha_{jt}}{\alpha_{1t}} = e^{\dot{e}_{jt}}$ imposing $\sum_j \alpha_{jt} = 1$ for each $t$, and $\rho$ by $\rho = \left( 1 + \dot{\phi} \right)$.

Once all these parameters are identified, the TFP parameters $A$’s are identified as residuals using the fact that in our model, thanks to the constant returns to scale assumption, total production is $Y_{mt} = \sum_i \sum_j w_{ijmt} L_{ijmt}$.

A.2 Moments and estimation of preferences

The estimation approach sequentially recovers labor supply and demand parameters. In this appendix, we discuss the details of the first step of our estimation, namely the estimation of the utility parameters. We estimate utility parameters and systematic surplus values (both pecuniary and non-pecuniary) for different workers using a GMM method.

The set of utility parameters to estimate is composed of:

- two shape parameters for the utility of consumption, $\gamma$ and $\sigma$;
- the scale parameter of the type-I extreme value distribution of idiosyncratic preference shock, $\sigma_{\theta}$;
- $I \times J \times T$ non pecuniary surpluses, $b_{ijt}$.

To estimate these parameters we target $(I \times J \times M \times T)$ employment ratios capturing the proportion of workers in each demographic group $i$ choosing occupation $j$ relative to the non-employed in all markets and all time periods, $\frac{\mu_{ijmt}}{\mu_{i0mt}}$.

In our specific application, $I = 12$, $J = 15$, $M = 4$ and $T = 5$. Thus, we estimate a total of 903 parameters using 3,600 moments.