Estimating On-the-job Search as an Intensive Margin*

Guillaume Wilemme

Sciences Po Paris and AMSE

October 2015

Preliminary

Abstract

Job-to-job transition is a major source of wage growth over a lifetime. This paper individually estimates how job search efforts respond to expected gains from mobility, accounting from observed characteristics. Both gains and efforts are unobserved, calling for a structural approach. I build a random search model with endogenous search effort and individual heterogeneity. Workers have forward-looking expectations on job opportunities and household characteristics. The optimal search effort function is jointly estimated with the distribution of job offers from panel data. Results show that individuals search slightly less when their expected gain from searching is lower.

JEL Code: C41, H31, J62

Keywords: job mobility, endogenous search, on-the job search, structural econometrics

*I thank Bruno Decreuse, Fabien Postel-Vinay, Jean-Marc Robin, Linas Tarasonis. The implementation of the NBER TAXSIM program with the PSID relies on part of the code from Eric Zwick, available on the NBER website. The collection of data used in this study was partly supported by the National Institutes of Health under grant number R01 HD069609 and the National Science Foundation under award number 1157698.

1 Introduction

Job mobility turns out to be an important channel for wage increases over a working career.[1] The literature has emphasized its numerous benefits such as absorbing negative shocks or improving the matching between jobs and workers. Although individual heterogeneity along this dimension has been noticed, no conclusion can be drawn about the causal determinants of job mobility. Yet, the causal relationships matter for explaining individual behaviors and the technology behind job mobility. Such an understanding is necessary to evaluate the impact of public policies and income taxation in particular.

A job change is the outcome of a complex process involving search frictions, job opportunities and individual behaviors. The efficiency of search activities, meaning how search efforts account in this process, is a key component. If search behaviors have a strong impact, individuals have interest in searching more intensively as the gain from a job change increases. Progressive income taxation is then likely to be a disincentive for workers to search as it reduces the net gain. The heterogeneity of in search behaviors has been neglected by most of the empirical literature, at least for employed workers.[2] This approach can be reasonable if the search technology has low efficiency or it can generate non-negligible biases if not. The limited knowledge about the search technology leaves us in a predicament.

Do workers search on-the-job in response to their net gains from searching? The answer I find is yes. This paper is the first one to provide an estimate of the search technology efficiency. I consider on-the-job search as an endogenous decision made by workers, controlling for individual heterogeneity. The intellectual challenge requires to encompass several difficulties: i) search efforts are not observable, only successful search is known; ii) search efforts are the outcome of an optimization problem, implying a strong non-linearity between the variables impacting the search cost, the exposure to job offers and the gain from changing jobs; iii) individuals face a taxation function which is, by definition, of infinite dimension.

I tackle this issue through a structural model that I estimate on tenure and wage data from the Panel Study of Income Dynamics. By explicitly modeling the optimization problem, I can infer the expected net-of-tax gain from searching for each agent. The structure of the model provides a relationship between job change, individual characteristics and the expected gain, which is used in the estimation part of the analysis. I propose a naive two-step procedure to estimate the model. The first step consists in estimating a Mincer-type equation to obtain individual productivities and job contracts. The second

[1] The seminal work of Topel and Ward (1992) found that one third of wage growth in the first ten years is due to job mobility.

[2] Two notable exceptions is Christensen et alii (2005) and Menzio et alii (2015).

step is the maximization of the likelihood conditionally on observing productivities and job contracts in order to recover the job offer distribution and the parameters of the search technology. A more efficient estimation method using simulated moments is still work in progress.

My paper contributes to the literature on the determinants of job turnover. Gentry and Hubbard (2004) show evidence of a significant negative impact of income taxation on job mobility. They use a reduced-form model (precisely a probit model) to estimate the impact of the tax system. Though their set of control variables is rich, their approach is not immune of the non-linearity problem I mentioned above. My work constitutes a complementary approach to document the effect of taxes on job changes. The results I find go into the same direction, though they are less attenuate. In addition, the structure of the model enables me to identify the causal channels and to simulate counterfactual situations at the cost of explicit hypotheses. The strength of the structural approach leans on the logical process to interpret the predictions of the model: if the model is not satisfying when confronted to reality, then one can question the precise hypotheses. On this methodological aspect, this article feeds the recent debate about the place of theory in applied microeconomics.[3]

My research is a first step to open the black box of on-the-job search behaviors by considering individual heterogeneity. Burdett (1978) is the first proposing a theoretical framework with job transitions of employed workers. Burdett and Mortensen (1998) go beyond and formulates a general equilibrium model with firms posting wages and both employed and unemployed workers searching. Christensen et alii (2005) use their framework and add endogenous search efforts. Heterogeneity is not in the scope of their paper. Among others, Bontemps et alii (2000) and Van den Berg, Ridder (1998) estimate a general equilibrium model of the labor market using wage data to recover the distribution of job opportunities. My approach and assumptions to these papers'. Contrary to these two papers, I do not consider the distribution of job offers as endogenous but I model explicitly endogenous search efforts. The seminal work of Topel and Ward (1992) emphasized the role of job mobility as a wage growth factor in early careers. This question has been recently documented through a search and matching model in Bagger et alii (2014). On-the-job search is a technology for increasing wages, workers can be poached and they can renegotiate their current contract as in Postel-Vinay, Robin (2002) and Cahuc et alii (2006). In my paper, the heterogeneity of firms is less explicitly formulated : workers climb on the job ladder in the same spirit as Barlevy (2008).

This research bridges the gap between the search and matching literature and a long tradition in economics that have focused on the labor supply to taxation and the partic-

[3]See the book by Wolpin (2013) and the responses/contributions of several economists.

ipation margin. Workers take into account the tax schedule when choosing whether or not to participate in the labor market (the extensive margin) and the amount of time to devote to work (the intensive margin). Much less is known regarding the job search margin: whether or not a worker looks for a better job while he is currently employed. A raise in taxes is then likely to reduce on-the-job search according to the benchmark search model: the gain from job shopping through a better paid job is more taxed whereas the cost of searching is less affected. This paper estimates job search behaviors and the extent to which individuals adjust them regarding job opportunities, tax (dis)incentives and possible individual costs.

2 Empirical Evidence

The rich dataset is provided by the Panel Study of Income Data (PSID).[4] It follows heads of household for several consecutive years. Thus, I observe wage dynamics, job mobility and job characteristics on one hand, and household characteristics (property incomes, marital status, children, spouse incomes) on the other hand. The on-line program TAXSIM from the National Bureau of Economic Research enables me to compute federal and state taxes (and marginal rates) for each observation from the individual characteristics. The exact tax function is then recovered for each worker.

2.1 Data construction

I use the sample of the PSID from 1979 to 1992. To avoid labor participation effects, I keep only males between 18 and 60, employee at a nonzero wage in two consecutive years, working between 35 and 65 hours a week, with no extra job. I use the weights defined in the PSID. The software TAXSIM is used to compute elements of the taxation of each individual according to their taxpayer characteristics[5]. TAXSIM cannot provide information on the individual income taxation prior to 1979.

The PSID data for job mobility suffers from inconsistency as noticed by Brown and Light (1992). We would expect the difference between declared tenure in year t and in year $t+1$ to be equal to the calendar time elapsed between the two interviews if the job has not changed. However, this is not coherent with the declared job mobility. I then follow the consolidation procedure proposed by Brown, Light (1992) to define job changes and tenure: if an individual says her job tenure in year $t+1$ is below the job tenure in year t plus the calendar time passed between the two interviews, then I consider that she has switched job, otherwise I consolidate the tenure in $t+1$ so that it is consistent

[4]Panel Study of Income Dynamics, public use dataset. Produced and distributed by the Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor, MI (2014).

[5]http://www.nber.org/taxsim/

with the tenure in t. I drop also individuals saying they changed jobs whereas it is not consistent with the tenure. A change of job is a change of employer in the PSID, except for years 1979 and 1980, for which it is a change of position. Promotion and intra-firm mobility are then excluded from the analysis.

The income tax is paid annually. Individuals may, however, have several jobs in a given year. I assume that individuals consider only long-term jobs in order to compute the net gain from the job change,. The taxation function is computed by TAXSIM using a counterfactual labor income, which is the labor income if the employee keeps the same job and the same wage for the entire year. The taxation includes federal taxes, state taxes and FICA taxes.

2.2 Descriptive Statistics

Table 1 provides some summary statistics relative to the main characteristics. The final sample contains 21,481 observations relative to 4,961 heads of households. The United States are geographically partitioned in 4 clusters in the PSID: Northeast, North Central, South and West. Importantly, a job change is observed for 6.3% of the observations.

A first step before looking at the structural estimation is to explore the more simple models. I estimate first a logit model, following Gentry and Hubbard (2004) for the choice of the regressor variables.[6] To benefit from the panel structure of the data, I also estimate a proportional-hazard Cox model, which is now a standard tool for duration data analysis. The main interest is to avoid the potential bias from a correlation between a regressor variable and job tenure. Average marginal effects cannot be computed through a Cox model so I have chosen a logit specification (instead of a probit for instance) because the coefficients are comparable on the same scale.[7] The estimation results are given tables 2 and 3.

The two models provide almost the same estimates. The main focus is on the contribution of the marginal tax rate for explaining the instant probability of switching job. I do not find a significant effect of the marginal rate at 10%, contrary to Gentry and Hubbard (2004). The sign, however, is negative and goes in the same direction: taxes may have a disincentive effect on the probability to switch job. The annual gross labor income has a significant non-linear impact. Some estimates corroborate a story of mobility cost: an individual is less mobile if he is homeowner, married and less educated. The number of children and the spouse labor income has a positive impact on the probability to change job, which is not intuitive at first sight.

[6] Gentry and Hubbard (2004) have one more variable regarding the taxation function, an index of the convexity for each individual, and they add the square of job tenure.

[7] In both models, the coefficients define the (constant) marginal effects of the regressors on the logit function applied to the probability of a job change.

Table 1 – Description of the sample

Variable	**Mean**	**SD**
Marginal tax rate (%)	39.68	8.79
Labor income (annual, in $10,000)	2.31	1.30
Spouse labor income (if spouse is working, annual, in $10,000)	1.03	0.91
Property income (annual, in $10,000)	-0.016	0.733
Age (in 10 years)	3.52	1.10
Homeowner	54%	
Member of a job union	24%	
Member of a minority group	14%	
Single	28%	
Education:		
- Less than high school	21%	
- High school graduate	23%	
- Some college or training	33%	
- Bachelor degree	16%	
- More than a bachelor degree	6%	
Number of children		
- No children	44%	
- 1 child	22%	
- 2 children	19%	
- 3 children or more	15%	
Region		
- Northeast	23%	
- North Central	28%	
- South	30%	
- West	19%	
Observations	21,481	
Individuals	4,961	
Job change (share of observations)	6.3%	

The sample contains male individual, between 18 and 60, employee at a nonzero wage in two consecutive years, working between 35 and 65 hours a week, with no extra job. The job change equals 1 if the worker gets a new job between two interviews. I use the weights of the survey.

Table 2 – Estimation results (part1)

Variable	**Logit** (Std. Err.)	**Cox** (Std. Err.)
Job tenure (months)	-0.014**	.
	(0.001)	
Marginal tax rate (%)	-0.009	-0.006
	(0.006)	(0.005)
Labor income (annual, in $10,000)	-0.667**	-0.668**
	(0.098)	(0.087)
Labor income square	0.061**	0.061**
	(0.012)	(0.010)
Spouse labor income (if spouse is working, annual, in $10,000)	0.200**	0.170**
	(0.072)	(0.064)
Spouse labor income square	-0.018	-0.015
	(0.011)	(0.011)
Property income (annual, in $10,000)	-0.008	-0.007
	(0.054)	(0.045)
Homeowner	-0.410**	-0.379**
	(0.092)	(0.082)
Member of a job union	-0.662**	-0.650**
	(0.124)	(0.118)
Age (in 10 years)	0.230	0.327
	(0.322)	(0.290)
Age square	-0.054	-0.069†
	(0.042)	(0.038)

Significance levels : $\dagger$: 10% $*$: 5% $**$: 1%. The explained variable is whether an individual has a job change. Each regression has controls with dummy variables for crossed time-region effects, occupation in 2 digits, industry in 2 digits. I used the weights of the survey.

Table 3 – Estimation results (part 2)

Variable	Logit	Cox
	(Std. Err.)	(Std. Err.)
Member of a minority group	-0.405**	-0.361**
	(0.115)	(0.104)
Single	0.330**	0.263**
	(0.115)	(0.098)
Education: (ref. less than high school)		
- High school graduate	0.036	0.007
	(0.117)	(0.106)
- Some college or training	0.414**	0.329**
	(0.117)	(0.105)
- Bachelor degree	0.640**	0.541**
	(0.154)	(0.140)
- More than a bachelor degree	0.842**	0.717**
	(0.206)	(0.188)
Number of children: (ref. no children)		
- 1 child	0.103	0.076
	(0.103)	(0.092)
- 2 children	0.072	0.045
	(0.108)	(0.099)
- 3 children or more	0.286*	0.218†
	(0.129)	(0.114)
Controls	Yes	Yes

Significance levels : † : 10% ∗ : 5% ∗∗ : 1%. The explained variable is whether an individual has a job change. Each regression has controls with dummy variables for crossed time-region effects, occupation in 2 digits, industry in 2 digits. I used the weights of the survey.

3 Structural Model

I introduce here the theoretical model. The first part provides the core idea of the model, meaning the search and matching technology and the timing of search in a static framework. The second part adds another set of hypotheses, incorporating forward-looking agents and endogenous search.

3.1 A two-step job search model

Consider a segment of the labor market in which workers are equally productive, within one period. Workers can be unemployed (state $S = U$) or employed at a wage w (state $S = E(w)$). There is a continuous distribution of job offers $\mathcal{F}(w)$ with probability density function $f(w)$. Job mobility is a two-step procedure. A worker in state S draws a job offer with probability p_S. This offer is then a draw from the job offer distribution. In addition, employed workers loose their job with probability δ, independent of the current wage. Following the literature, p_S can be interpreted as the search efficiency or intensity. It is a mix of exposure to job opportunities and active search efforts. The search technology as defined here is non-directed in a sense that a worker can draw any offer in the distribution after the first step.[8] The timing is the following: individuals search and receive an offer with probability p_S, then job separation occurs with probability δ, lastly workers with an offer w from $\mathcal{F}$ decide to accept it or not.

Utility supposed to be strictly increasing in the wage and state-invariant. There is a reservation wage, $\underline{w}$, so that individuals are indifferent between state U and state $E(\underline{w})$. Table 4 provides the transition probabilities between states. An employed in state $E(w)$ becomes unemployed if her current job breaks, with probability δ, and she does not receive an acceptable offer, with probability $1 - p_{E(w)}(1 - \mathcal{F}(\underline{w}))$. She switches to a worse-paid job, state $E(w')$ with $w' < w$, if she looses her job and receives the offer w'. She gets a better-paid job, state $E(w')$ with $w < w'$, if she receive the offer w' independently of whether her job stops at the end of the period or not. She may also keep her current job if her job does not break and she does not receive any wage-improving offer.[9] An unemployed gets a job with wage w' with probability $p_U f(w')$. She remains unemployed is she does not receive any acceptable offer.

3.2 Endogenous search

In this section, I provide microfoundations for the search intensities in order to disentangle the role of exposure and of search efforts. Consider the problem of a worker. Notations,

[8] Search is often called random in the literature.

[9] Drawing the same wage on the market occurs with a null probability because of continuous distribution

Table 4 – Transition Probabilities and Densities

From\To	$E(w')$, $\underline{w} \leq w'$			U
	$w' < w$	$w < w'$	$w' = w$	
$E(w)$	$\delta p_{E(w)} f(w')$	$p_{E(w)} f(w')$	$(1-\delta)[1-p_{E(w)}(1-\mathcal{F}(w))]$	$\delta[1-p_{E(w)}(1-\mathcal{F}(\underline{w}))]$
U	$p_U f(w')$			$1-p_U(1-\mathcal{F}(\underline{w}))$

Note: One can check that the probabilities integrate to 1 row by row.

state-space $\Omega_t = \{x, y, z, t\}$, w_t, $\tilde{w}_t$, ϵ_t

The Bellman equations:

$$
\begin{aligned}
V(\Omega_t, w_t) &= \begin{cases} V_0(\Omega_t) & \text{if} \\ \max\{V_0(\Omega_t), V_1(\Omega_t, \tilde{w}_t)\} & \text{if} \\ V_1(\Omega_t, w_t) & \text{if} \\ \max\{V_1(\Omega_t, w_t), V_1(\Omega_t, \tilde{w}_t)\} & \text{if} \end{cases} \\
V_k(\Omega_t, w_t) &= \max_{\lambda} \{R_k(\Omega_t, w_t, \lambda) + \beta \mathbb{E}\left[V(\Omega_{t+1}, w_{t+1}) | \Omega_t, \lambda, w_t, k\right]\} \\
V_k(\Omega_T, w_T) &= R_k(\Omega_T, w_T, 0)
\end{aligned}
$$

(simplistic case when $V_0(\Omega_t) < V_1(\Omega_t, w_t)$)

$$\mathbb{E}\left[V(\Omega_{t+1}, w_{t+1}) | \Omega_t, \lambda, w_t, k=1\right] = \delta(1-p_1)\mathbb{E}\left[V_0(\Omega_{t+1}) | \Omega_t\right] + ...$$

Denote $S_t = E$ or U the employment status, w_t the current wage, x_t the set of individual characteristics impacting the taxation function, y_t the set of characteristics impacting individual productivity, z_t the set of characteristics impacting the individual search technology. Workers maximize the expected discounted utility:

$$\max_{\{\lambda_\tau\}_{t \leq \tau \leq T}} \mathbb{E}_t \left\{ \sum_{\tau=t}^{T} \beta^{\tau-t} \mathcal{U}(S_\tau, w_\tau, \lambda_\tau, x_\tau, z_\tau) \right\}.$$

The utility writes

$$\mathcal{U}(E, w, \lambda, x, z) = \mathcal{N}_t(w, x) - c^E(\lambda, z)$$

and

$$\mathcal{U}(U, w, \lambda, x, z) = b - c^U(\lambda, z)$$

the state space (S_t, Ω_t, t) with $\Omega_t = (x_t, y_t, z_t)$ .... the stochastic dynamics of the state space The Bellman equations:

$$
\begin{aligned}
&V_t^E(\Omega_t) = \max_{\lambda} \{\} \\
&\mathcal{U}(E, w_t, \lambda, x_t, z_t) + \delta \left[(1-p^E(\lambda)) V_{t+1}^U(h_{t+1}) + p^E(\lambda) \int \max\left[V_{t+1}^U(h_{t+1}), V_{t+1}^E(h_{t+1}, 0, \theta)\right] f(w|y_t) d\theta \right] \\
&+ (1-\delta) \left[(1-p^E(\lambda)) V_{t+1}^E(h_{t+1}, \tau_{t+1}, \tilde{\theta}_t) + p^E(\lambda) \int \max\left[V_{t+1}^E(h_{t+1}, \tau_{t+1}, \tilde{\theta}_t), V_{t+1}^E(h_{t+1}, 0, \theta)\right] f(w|y_t) d\theta \right]
\end{aligned}
$$

Need to define cost function and a gain function. Let Φ be the gain from receiving an offer (unconditional acceptable or not).

Consider an economy populated by N individuals. All of them are employed and working full-time. Each agent receives a before-tax wage $\mathcal{W}(h_t, \tau_t, \tilde{\theta}_t)$, which depends on a set of human capital characteristics h_t, the current tenure τ_t and a job effect $\tilde{\theta}_t$ at period t. The dependence in h_t captures changes in general abilities to produce, whatever the job a worker has. Human capital variables are typically the levels of education and experience. Individuals can improve their job-specific capacities over time, hence tenure is accounted in the wage function. The job effect measures the quality of the job. For a same intrinsic productivity, a worker can be paid differently according to the firm's type. This variable can be simply interpreted as the firm's payment contract. The Bellman equations:

$$
\begin{aligned}
&V_t^E(h_t, \tau_t, \tilde{\theta}_t) = \max_{\lambda} \{\} \\
&\mathcal{N}\mathcal{W}(h_t, \tau_t, \tilde{\theta}_t) - c^E(\lambda) + \delta \left[(1 - p^E(\lambda)) V_{t+1}^U(h_{t+1}) + p^E(\lambda) \int \max \left[V_{t+1}^U(h_{t+1}), V_{t+1}^E(h_{t+1}, 0, \theta) \right] f(\theta) d\theta \right] \\
&+ (1 - \delta) \left[(1 - p^E(\lambda)) V_{t+1}^E(h_{t+1}, \tau_{t+1}, \tilde{\theta}_t) + p^E(\lambda) \int \max \left[V_{t+1}^E(h_{t+1}, \tau_{t+1}, \tilde{\theta}_t), V_{t+1}^E(h_{t+1}, 0, \theta) \right] f(\theta) d\theta \right]
\end{aligned}
$$

$$
\begin{aligned}
&V_t^U(h_t) = \max_{\lambda} \{\} \\
&b - c^U(\lambda) + (1 - p^U(\lambda)) V_{t+1}^U(h_{t+1}) + p^U(\lambda) \int \max \left[V_{t+1}^U(h_{t+1}), V_{t+1}^E(h_{t+1}, 0, \theta) \right] f(\theta) d\theta
\end{aligned}
$$

$$
\begin{aligned}
&V_T^E(h_T, \tau_T, \tilde{\theta}_T) = \mathcal{N}\mathcal{W}(h_T, \tau_T, \tilde{\theta}_T) \\
&V_T^U = b
\end{aligned}
$$

Given the functional forms of c and p, the objective function of the problem is concave and the first-order condition is equivalent to

$$c^{j\prime}(\lambda_t^j) = p^{j\prime}(\lambda_t^j) \Phi_t^j$$

with

$$\Phi_t^E = \delta \left[\int \max \left[V_{t+1}^E(h_{t+1}, 0, \theta) - V_{t+1}^U(h_{t+1}), 0 \right] f(\theta) d\theta \right] \tag{1}$$

$$+ (1 - \delta) \left[\int \max \left[V_{t+1}^E(h_{t+1}, 0, \theta) - V_{t+1}^E(h_{t+1}, \tau_{t+1}, \tilde{\theta}_t), 0 \right] f(\theta) d\theta \right] \tag{2}$$

$$\Phi_t^U = \int \max \left[V_{t+1}^E(h_{t+1}, 0, \theta) - V_{t+1}^U(h_{t+1}), 0 \right] f(\theta) d\theta \tag{3}$$

Whereas the dynamics of human capital h_t is exogenous, individuals can impact their job-specific characteristics, τ_t and $\tilde{\theta}_t$, through searching on-the-job. Each period, a worker chooses a one-dimension search effort λ. The probability to receive a random offer in a small interval $(\theta, \theta+\Delta\theta)$ is $p(\lambda)f(\theta)\Delta\theta$, increasing in λ. The search technology is assumed random, implying this probability to be multiplicatively separable. Equivalently, the search proceeds in two steps. The worker receives a job offer with probability $p(\lambda)$ in $(0, 1)$. In this case, she then receives an offer drawn from a distribution with probability density function (p.d.f.) f and cumulative density function (c.d.f.) $\mathcal{F}$. The density of job offers is predetermined when the worker chooses its effort.[10] Search, however, has a monetary costly $c(\lambda)$. The entire search technology cannot be identified with the data used in this paper. I use the following specification:

$$p(\lambda) = 1 - e^{-(\mu_0+\mu\lambda)},$$
$$c(\lambda) = \frac{\lambda^2}{2},$$

where μ_0 and μ are parameters that may vary across individuals. The very constrained form of the cost function is due to identifiability restrictions.[11] These deterministic functions are general but specific enough to fit the following interpretation. λ is analogous to a number of applications sent. As the number of applications increases, it is getting harder for a worker to apply to new jobs, meaning the cost function is convex. The worker receives $\lambda_0 + \mu\lambda$ job offers from her search strategy o0n average per period. Assuming a Poison distribution, the probability to receive at least one offer during the period is the definition of $p(\lambda)$. If the individual receives several offers, she takes one of then from the uniform distribution. Believing that a search technology is occurring this way in practice is unrealistic. Nevertheless, this hypothesis provides microfoundations and exhibits the parameters of interest. The coefficient μ captures the search efficiency. When it equals 0, the individual cannot influence the arrival rate of job offers.

Wages are taxed. The taxation function faced by each individual depends on some characteristics of the household, as well as other incomes. These variables are predetermined when the worker chooses her search effort. Denote $\mathcal{N}$ the individual after-tax function at the period. The timing is the following. At the beginning of period t, the worker chooses a search effort λ_t and incurs the search cost $c(\lambda_t)$. Then, a job offer drawn from $\mathcal{F}$ is received with probability $p(\lambda_t)$. The worker can end up in two different positions: she keeps her job and receives the gross wage $\mathcal{W}(h_t, \tau_t, \tilde{\theta}_t)$, or she switches to a new job θ yielding $\mathcal{W}(h_t, 0, \theta)$. Individuals maximize the current-period expected utility

[10] One can possibly endogenize the distribution. See Bontemps, Robin, van den Berg (2000).

[11] If $c(\lambda) = \frac{\kappa}{2}\lambda^2$, one cannot identify separately κ from μ using the technique developed in this paper. I choose to normalize $\kappa = 1$.

without accounting for the next periods. A job offer θ may not be preferable for an agent, she may rather reject it. The program of the agent is

$$\max_{\lambda}\left\{-c(\lambda)+\left[(1-p(\lambda))\mathcal{NW}(h_t,\tau_t,\tilde{\theta}_t)+p(\lambda)\int\max\left[\mathcal{NW}(h_t,\tau_t,\tilde{\theta}_t),\mathcal{NW}(h_t,0,\theta)\right]f(\theta)d\theta\right]\right\}.$$

Given the functional forms of c and p, the objective function of the problem is concave and the first-order condition is equivalent to

$$c'(\lambda_t)=p'(\lambda_t)\Phi_t$$

with

$$\Phi_t=\int\max\left[\mathcal{NW}(h_t,0,\theta)-\mathcal{NW}(h_t,\tau_t,\tilde{\theta}_t),0\right]f(\theta)d\theta. \tag{4}$$

The optimal search effort is such that the marginal cost equals the marginal gain from searching. Φ_t defines the net returns from search activities. The optimal search efforts and the derived probability to receive an offer write

$$\lambda_t=\frac{1}{\mu}\Omega\left(\mu^2e^{-\mu_0}\Phi_t\right), \tag{5}$$

$$p_t=p(\lambda_t)=1-e^{-\mu_0-\Omega\left(\mu^2e^{-\mu_0}\Phi_t\right)}. \tag{6}$$

Function Ω is the Lambert W-function, defined as the solution of $\Omega(x)e^{\Omega(x)}=x$, and it is increasing on $(0,+\infty)$. The search effort and the probability to get an offer rise with the returns from searching. The worker does not respond to the returns from searching when μ is equal to 0.

The model has particular features. They are formulated in the following propositions.

Proposition 1 *The search effort λ_t decreases in the job quality $\tilde{\theta}_t$ if*

- *the after-tax wage function $\mathcal{NW}$ is an increasing function of job quality, $\frac{\partial\mathcal{NW}}{\partial\tilde{\theta}}>0$.*

A worker who has already a good job (equivalently a high $\tilde{\theta}_t$) has less chance to find a better job, so her expected returns from searching are lower. She then invests less in search.

Proposition 2 *The search effort λ_t increases in h_t*

- *the after-tax wage function $\mathcal{NW}$ is an increasing function of job quality, $\frac{\partial\mathcal{NW}}{\partial\tilde{\theta}}>0$,*
- *the after-tax wage function $\mathcal{NW}$ exhibits supermodularity between human capital and job quality, $\frac{\partial^2\mathcal{NW}}{\partial\tilde{\theta}\partial h}>0$.*

Supermodularity implies that a productive worker takes more benefits from a high-type job than a worker less endowed with human capital. The returns from searching are higher for this worker. The proofs are detailed in the appendix.

The model does not account for job changes with wage lost, though it occurs in the data. The model will be extended in this direction (to be done). In the estimation, wage losses will considered as noise. I assume that wages are observed with error by the econometrician (but not by the worker). The observed wage w_t is defined by

$$\ln(w_t) = \ln\left(\mathcal{W}(h_t, \tau_t, \tilde{\theta}_t)\right) + \epsilon_t, \tag{7}$$

with ϵ_t an error term.

4 The Econometric Model

The data contains individual wages, job mobility and a set of household and job characteristics. We do not observe directly the job effects, the probabilities to receive an offer and the job offer distribution. In this section, I state the parametric assumptions and the identification hypotheses required to estimate the structural model. The econometric model can be formulated as non-linear Kalman filter in a state-space representation. The variables will be indexed by two subscripts for the individual and the period. The observed wages $w_{i,t}$ and the job change dummies $m_{i,t}$ are the two measurement variables. The two state variables are the job quality $\tilde{\theta}_{i,t}$ and the returns from searching $\Phi_{i,t}$.

I choose a log-linear specification of the wage function so that the observed wage follows a classical Mincer function,

$$\ln(w_{i,t}) = \alpha' h_{i,t} + \gamma \tau_{i,t} + \tilde{\theta}_{i,t} + \epsilon_{i,t}, \tag{8}$$

where α is a vector of the same size as $h_{i,t}$ and γ is a scalar. The error term $\epsilon_{i,t}$ is assumed to follow a centered normal distribution $\mathcal{N}(0, \sigma_\epsilon)$.

I characterize the event of receiving a job offer. I estimate a model in which the search efficiency parameter μ is identical for everyone and the fixed search term μ_0 is a linear combination of observable variables $\beta' z_{i,t}$. Denote $\mathcal{E}(1)$ the exponential distribution of parameter 1. From equation (6), the probability to get a job offer for individual i at time t, $p_{i,t}$, is equal to the c.d.f. of $\mathcal{E}(1)$ evaluated at $\beta' z_{i,t} + \Omega\left(\mu^2 e^{-\beta' z_{i,t}} \Phi_{i,t}\right)$. Equivalently, a worker receives a job offer if and only if she draws $\nu_{i,t}$ from $\mathcal{E}(1)$ such that

$$\nu_{i,t} < \beta' z_{i,t} + \Omega\left(\mu^2 e^{-\beta' z_{i,t}} \Phi_{i,t}\right).$$

Once the worker gets a job offer $\theta_{i,t}$ from $\mathcal{F}$, she accepts it if and only if

$$\gamma \tau_{i,t} + \tilde{\theta}_{i,t-1} < \theta_{i,t}.$$

In this case, the worker is better off starting the new job even if she looses the returns to tenure in the previous job. Define $m_{i,t}$ as a dummy variable whether the worker changes job or not. A worker switches job when these two events occur, otherwise she keeps her current job:

$$m_{i,t} = \begin{cases} 1 \text{ if } \begin{cases} \nu_{i,t} < \beta' z_{i,t} + \Omega\left(\mu^2 e^{-\beta' z_{i,t}} \Phi_{i,t}\right) \\ \gamma \tau_{i,t} + \tilde{\theta}_{i,t-1} < \theta_{i,t} \end{cases} \\ 0 \text{ if not} \end{cases} . \tag{9}$$

The dynamics of the job effect is

$$\tilde{\theta}_{i,t} = \begin{cases} \theta_{i,t} \text{ if } \begin{cases} \nu_{i,t} < \beta' z_{i,t} + \Omega\left(\mu^2 e^{-\beta' z_{i,t}} \Phi_{i,t}\right) \\ \gamma \tau_{i,t} + \tilde{\theta}_{i,t-1} < \theta_{i,t} \end{cases} \\ \tilde{\theta}_{i,t-1} \text{ if not} \end{cases} . \tag{10}$$

From the wage specification and equation (4), I characterize the returns from searching as

$$\Phi_{i,t} = \int_{\theta > \gamma \tau_{i,t} + \tilde{\theta}_{i,t-1}} \left[\mathcal{N}_{i,t}(e^{\alpha' h_{i,t} + \theta}) - \mathcal{N}_{i,t}(e^{\alpha' h_{i,t} + \gamma \tau_{i,t} + \tilde{\theta}_{i,t-1}})\right] f(\theta) d\theta. \tag{11}$$

5 Estimation

The model will be estimated through maximum simulated likelihood. As a preliminary stage, I use a two-step method. This is an inefficient strategy but the main advantage lies on the speed of computation. It will also provides good priors for the one-step method.

5.1 Two-step procedure

Step 1 The first measurement equation (8) can be used to estimate parameters α, γ and the job components $\{\tilde{\theta}_{i,t}\}$. I denote separately human capital characteristics that vary other time (like experience) $h^v_{i,t}$ from those that keep constant (like education). I range individuals with the same fixed characteristics in different groups denoted by r. $r(i)$ is the group to which individual i belongs to and α_r the group fixed effect. $j(i,t)$ is the index of worker i job at time t. First, estimate the following equation:

$$\ln(w_{i,t}) = \alpha^{v\prime} h^v_{i,t} + \gamma \tau_{i,t} + \alpha_{r(i)} + \tilde{\theta}_{j(i,t)} + \epsilon_{i,t} \tag{12}$$

I denote $\tilde{\theta}_{j(i,t)}$ to emphasize that the job component remains constant over a job spell. The framework is a linear regression with individual fixed effects. The parameters α^v and γ can be estimated by using the *within* transformation,

$$\overline{\ln(w)}_{i,t} = \alpha^{v\prime} \overline{h}^v_{i,t} + \gamma \overline{\tau}_{i,t} + \overline{\epsilon}_{i,t}, \tag{13}$$

denoting

$$\overline{x}_{i,t} = x_{i,t} - \frac{\sum\limits_{j(i',t')=j(i,t)} x_{i',t'}}{\sum\limits_{j(i',t')=j(i,t)} 1}$$

for any variable $x_{i,t}$.

Minimizing ordinary least squares provides estimates $\hat{\alpha}$ and $\hat{\gamma}$. For each job k, define the average residual

$$\hat{R}_k = \frac{\sum\limits_{j(i,t)=k} (\ln(w_{i,t}) - \hat{\alpha}^{v\prime} h_{i,t} - \hat{\gamma}\tau_{i,t})}{\sum\limits_{j(i,t)=k} 1}.$$

The key identification assumption to estimate separately the α_r and the $\tilde{\theta}_j$ is the upper bound on the job offer distribution: $\tilde{\theta}_j \leq 0$. An estimate of α_r is derived:

$$\hat{\alpha}_r = \max\left\{\hat{R}_k | \text{for } (i,t) \text{ such that } j(i,t) = k \text{ and } r(i) = r\right\}.$$

The estimates $\hat{\tilde{\theta}}_j(i,t)$ are obtained as $\hat{\tilde{\theta}}_j(i,t) = \hat{R}_{j(i,t)} - \hat{\alpha}_{r(i)}$.

Step 2 We can write the likelihood of observing a job change or not $m_{i,t}$ and of drawing the new job effect in the first case. Define the likelihood of either not changing job or either changing job and drawing a new job offer, conditionally on observing the job effects $\tilde{\theta}_{i,t}$.

$$\mathcal{L}\left(\beta, \mu, \mathcal{F}; \{h_{i,t}, z_{i,t}, m_{i,t}, N_{i,t}, \tilde{\theta}_{i,t}\}\right) = \prod_{i,t} \left(p_{i,t}\mathcal{F}(\gamma\tau_{i,t} + \tilde{\theta}_{i,t-1}) + 1 - p_{i,t}\right)^{1-m_{i,t}} \left(p_{i,t} f(\tilde{\theta}_{i,t})\right)^{m_{i,t}} \tag{14}$$

$$\text{with} \left\{ \begin{array}{l} p_{i,t} = \Psi\left(\beta' z_{i,t} + \Omega\left(\mu^2 e^{-\beta' z_{i,t}} \Phi_{i,t}\right)\right) \\ \Phi_{i,t} = \int_{\theta > \gamma\tau_{i,t} + \tilde{\theta}_{i,t-1}} \left[\mathcal{N}_{i,t}(e^{\alpha' h_{i,t} + \theta}) - \mathcal{N}_{i,t}(e^{\alpha' h_{i,t} + \gamma\tau_{i,t} + \tilde{\theta}_{i,t-1}})\right] f(\theta) d\theta \end{array} \right.$$

I maximize this likelihood after substituting the parameters α, γ and $\tilde{\theta}_{i,t}$ by their estimates from the first step. The c.d.f. of the exponential distribution is Ψ. To deal with the nonparametric part, I use a sieve approximation of $\mathcal{F}$ with stepwise functions. On one hand, the computation of $\Phi_{i,t}$ is made numerically simpler, as the integral becomes a finite sum and as only a finite number of evaluation of the function $N_{i,t}$ (computed using TAXSIM) are required. On the other hand, the likelihood must be amended because the approximation of $\mathcal{F}$ is not continuous nor differentiable. I substitute $\mathcal{F}$ and f in the first line of (14) by a smooth kernel of the sieve approximation.

Define a partition of the segment $(0,1)$ of $M+1$ elements by a sequence $s = \{s_m\}$ with $0 < s_1 < s_2 < ... < s_M \leq 1$, $M > 1$. I estimate $\hat{\kappa} = \{\hat{\kappa}_m\}$ with $0 \leq \hat{\kappa}_m \leq 1$ and $\sum_{m=1}^{M} \hat{\kappa}_m = 1$ such that $\sum_{m=1}^{M} \hat{\kappa}_m \mathbf{1}(e^\theta > s_m)$ is a stepwise approximation of $\tilde{\mathcal{F}}$. This approximation is used to compute the $\Phi_{i,t}$. I define a density estimator with Epanechnikov kernel function:

$$\tilde{f}_\kappa(\theta) = \sum_{m=1}^{M} \kappa_m \mathcal{K}(e^\theta - s_m) \qquad \tilde{\mathcal{F}}_\kappa(\theta) = \sum_{m=1}^{M} \kappa_m \int_0^\theta \mathcal{K}(e^u - s_m) du$$

$$\text{with } \mathcal{K}(v) = \frac{3}{4b}\left(1 - \left(\frac{v}{b}\right)^2\right) \mathbf{1}(|v| < b)$$

I use the textbook formula for the optimal bandwidth of a normal distribution to set b.[12] The program to maximize writes:

$$\max_{\beta,\mu,\kappa} \prod_{i,t} \left(p_{i,t}\tilde{\mathcal{F}}_\kappa(\hat{\gamma}\tau_{i,t} + \hat{\tilde{\theta}}_{i,t-1}) + 1 - p_{i,t}\right)^{1-m_{i,t}} \left(p_{i,t}\tilde{f}_\kappa(\hat{\tilde{\theta}}_{i,t})\right)^{m_{i,t}}$$

$$\text{with} \begin{cases} p_{i,t} = \Psi\left(\beta' z_{i,t} + \Omega\left(\mu^2 e^{-\beta' z_{i,t}} \Phi_{i,t}\right)\right) \\ \Phi_{i,t} = \sum\limits_{m | s_m > e^{\hat{\gamma}\tau_{i,t} + \hat{\tilde{\theta}}_{i,t-1}}} \left[\mathcal{N}_{i,t}(s_m e^{\hat{\alpha}' h_{i,t}}) - \mathcal{N}_{i,t}(e^{\hat{\alpha}' h_{i,t} + \hat{\gamma}\tau_{i,t} + \hat{\tilde{\theta}}_{i,t-1}})\right] \kappa_m \end{cases}$$

The solution of this maximization converges to the solution of the maximization of (14) when M tends to ∞. There is a natural way of setting the parameter M in practice. We increase M until the relative distance between the former and the new estimates of β and μ is smaller than a fixed tolerance parameter.

5.2 Identification

The identification of the first step relies on the ad-hoc normalization of the δ_r. Here, I show the identification in the second step. If the true parameters are $(\beta^*, \mu^*, \mathcal{F}^*)$, I prove that there is no tuple $(\beta, \mu, \mathcal{F})$ different from $(\beta^*, \mu^*, \mathcal{F}^*)$ such that for any θ,

$$p_{i,t}\mathcal{F}(\theta) + 1 - p_{i,t} = p^*_{i,t}\mathcal{F}^*(\theta) + 1 - p^*_{i,t}$$

$$p_{i,t} f(\theta) = p^*_{i,t} f^*(\theta)$$

$$\text{where} \begin{cases} p_{i,t} = \Psi\left(\beta' z_{i,t} + \Omega\left(\mu^2 e^{-\beta' z_{i,t}} \Phi_{i,t}\right)\right) \\ \Phi_{i,t} = \int_{\theta > \gamma\tau_{i,t} + \tilde{\theta}_{i,t-1}} \left[\mathcal{N}_{i,t}(e^{\alpha' h_{i,t} + \theta}) - \mathcal{N}_{i,t}(e^{\alpha' h_{i,t} + \gamma\tau_{i,t} + \tilde{\theta}_{i,t-1}})\right] f(\theta) d\theta \end{cases}$$

If one integrates the second equation, it follows that $\lambda^*_{i,t}\mathcal{F}^*(\theta) = \lambda_{i,t}\mathcal{F}(\theta)$, hence $\lambda^*_{i,t} = \lambda_{i,t}$ and $f^*(\theta) = f(\theta)$ with the first equation. For non-degenerated distribution $\mathcal{F}^*$, some variation in $z_{i,t}$ and $\Phi_{i,t}$ enables us to identify β^* and μ^* as long as the two variables are not collinear and Ψ is invertible. Identification relies on observing enough change of jobs with different contracts of θ so that we can proxy the probability of receiving a non-acceptable offer.

[12]See Wasserman, *All of Statistics*.

5.3 Results

I follow the estimation strategy described above. Only the estimates of a model with job change defined by panel A is currently available, without taxation ($N_{i,t}(.)$ is the identity function). In this case, $\Phi_{i,t}$ is the expected gain in the pretax labor income. The standard errors are computed by bootstrap methods (100 replications), which is time-consuming given the numerical complexity of step 2. The dimension of the sieve M is fixed to 10, the partition is uniform ($s_{m+1} - s_m = 1/M$). The estimated of the Mincer equation are given table 5. The results from the first step are not surprising. Experience has decreasing returns and the more educated workers earns the more. Job tenure has also a positive impact.

Table 5 – Structural Estimates

Variable		
$\ln(w_{i,t}) = \alpha' h_{i,t} + \gamma \tau_{i,t} + \theta_{i,t} + \epsilon_{i,t}$		
- $h_{i,t}$	$\hat{\alpha}$	
experience (in 10 years)		0.085 (0.047)
experience^2 (in 10 years)		-0.031 (0.005)
Less than high school		10.4 (0.12)
Some college or training		10.5 (0.14)
Bachelor degree		10.8 (0.10)
More than a bachelor degree		11.1 (0.11)
High school graduate		11.5 (0.16)
- $\tau_{i,t}$	$\hat{\gamma}$	
Job tenure (in months)		0.0049 (0.0012)
$p_{i,t} = \Psi\left(\beta' z_{i,t} + \Omega\left(\mu^2 e^{-\beta' z_{i,t}} \Phi_{i,t}\right)\right)$		
- efficiency parameter	$\hat{\mu}$	556 (191)
- $z_{i,t}$	$\hat{\beta}$	
intercept		9.27 (12)
experience (in 10 years)		6.4 (21)
not single		4.2 (46)
number of children		4.0 (711)
spouse labor income (in \$10,000)		4.1 (5.9)

Standard errors are given in parentheses. They are computed through a bootstrap method with 100 replications. $\Phi_{i,t}$ is expressed in \$10.000.

Figure 1 provides the estimated job effectsand the estimated distribution of job offers, after an exponential transformation.[13] As expected, the distribution of observed (accepted) jobs first-order stochastically domainates the distribution of job offers.

From table 5, it is obvious that the second step is poorly estimated as no coefficient is significant for $\hat{\beta}$. The efficiency parametr, however, is significantly positive. The

[13] This transformation is conveninent both for the estimation and the interpretation.

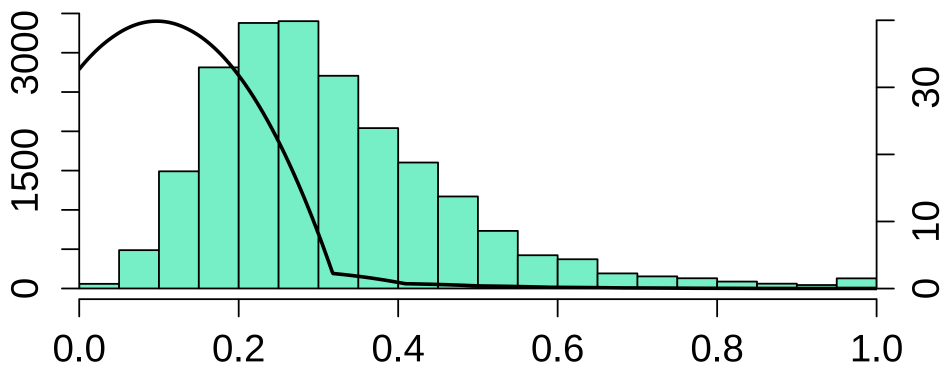

Figure 1 – Distribution of estimated accepted job contracts and estimated density of the job offer distribution
The histogram corresponds to the distribution in the sample of estimated jobs after an exponential transformation $e^{\hat{\hat{\theta}}_j}$ (left y-axis). The curve is a kernel estimate of the job offer distribution after exponential transformation too (right y-axis).

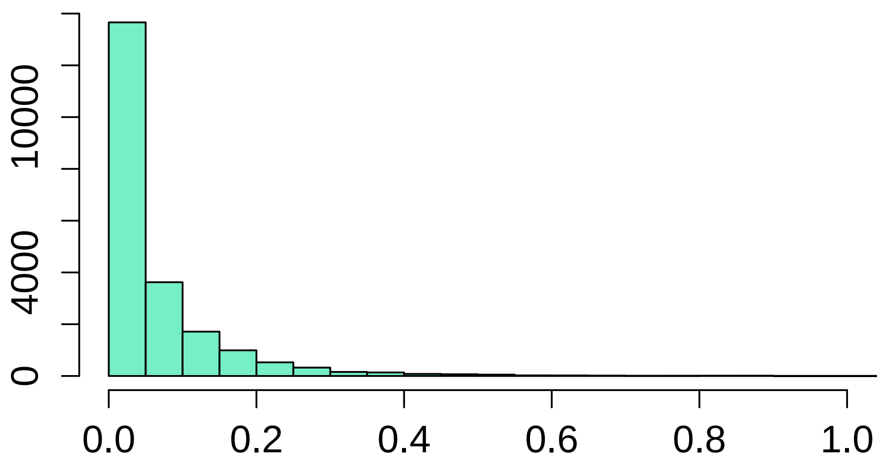

Figure 2 – Distribution of estimated returns to search
The returns to search are the estimates of $\Phi_{i,t}$ among the observations. The x-axis is expressed in \$10.000, it is right-side truncated.

distribution of the estimated returns to search is right-skewed in figure 2. Most of the workers have net returns below \$2.000.

References