Wages, Job Mobility, and Duration Dependence.
A Proposed Estimation Toolbox

Bocar A. Ba

Software Engineering Workshop

May 7 , 2015

Outline

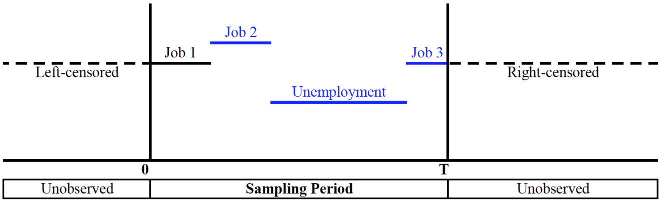
1. Introduction
2. Empirical Strategy
3. Code

Introduction

Motivations

Over the sampling period, notice that workers can be in one of the following states

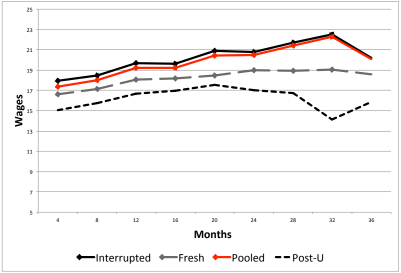
1. Job 1: Interrupted tenure
2. Job 2: Fresh tenure
3. Job 3: Post-unemployment tenure



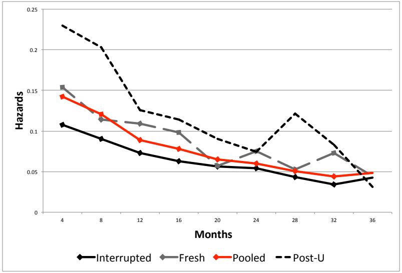
Introduction

Motivations: Wages and Hazards by Spell Type

- Bias toward negative duration dependence
    - "More mobility prone persons are the first to leave the population leaving the less mobile behind and hence creating the illusion of stronger negative duration dependence than actually exists" (Heckman and Singer (1984))
- Overstatement of interrupted/fresh tenure spells



(b) Wages



(c) Job Separation Hazards

Introduction

Potential Explanations

- Selection into employment
- Selection into tenure
- Pure duration dependence
- Unobserved heterogeneity

Introduction

- Estimate wages conditional on the following transitions
    1. Job-to-Job (Tenure)
    2. Job-to-Unemployment (Employment participation and Tenure)
    3. Unemployment-to-Job (Employment participation)
- but, those transitions are hazard rates
- So need to control for spell structures
    - Stock samples or interrupted spells are composed of individuals with
        - long spells
        - short spells
    - Flow samples or fresh spells are only composed of individuals with short spells

Introduction

- Estimate wages conditional on the following transitions
  1. Job-to-Job (Tenure)
  2. Job-to-Unemployment (Employment participation and Tenure)
  3. Unemployment-to-Job (Employment participation)
- but, those transitions are hazard rates
- So need to control for spell structures
  - Stock samples or interrupted spells are composed of individuals with
    - long spells
    - short spells
  - Flow samples or fresh spells are only composed of individuals with short spells

Introduction

- Estimate wages conditional on the following transitions
  1. Job-to-Job (Tenure)
  2. Job-to-Unemployment (Employment participation and Tenure)
  3. Unemployment-to-Job (Employment participation)
- but, those transitions are hazard rates
- So need to control for spell structures
  - Stock samples or interrupted spells are composed of individuals with
    - long spells
    - short spells
  - Flow samples or fresh spells are only composed of individuals with short spells

Empirical Strategy

Setup:

- Assumptions
  1. Self-selection across jobs
  2. Self-selection into employment
- How? Five spell type:
  1. Interrupted tenure: continuously employed with the same employer since the beginning of the sampling period
  2. Fresh tenure: continuously employed, but changed employer at least once since the beginning of the sampling period
  3. Post-unemployment tenure: previously unemployed but currently employed
  4. Interrupted unemployment: continuously unemployed since the beginning of the sampling period
  5. Fresh unemployment spells: currently unemployed but has been employed at least once since the beginning of the sampling period

Empirical Strategy

Setup:

- Assumptions
  1. Self-selection across jobs
  2. Self-selection into employment

- How? Five spell type:
  1. Interrupted tenure: continuously employed with the same employer since the beginning of the sampling period
  2. Fresh tenure: continuously employed, but changed employer at least once since the beginning of the sampling period
  3. Post-unemployment tenure: previously unemployed but currently employed
  4. Interrupted unemployment: continuously unemployed since the beginning of the sampling period
  5. Fresh unemployment spells: currently unemployed but has been employed at least once since the beginning of the sampling period

Empirical Strategy

Wage Equation

- At each period t, individual i in spell j earns

$$\log(w_{j,it}) = [z_{j,it} \ {h^w}_j(t)]\gamma_j + \xi_{j,i} + \varepsilon_{j,it}$$

- with
  - $z_{j,it}$ = time varying and constant Mincerian covariates
  - ${h^w}_j(t)$ = time effect, function of the worker's potential experience and tenure on the job
  - $\xi_{j,i}$ = unobserved heterogeneity
  - The residuals $\varepsilon_{j,it} \sim N(0, \sigma_j^2)$.
  - $Z_{j,it} = [z_{j,it} \ {h^w}_j(t)]$
  - $\Gamma_{w,j} = [\gamma_j \ \xi_j]$
- Assuming

$$E(\varepsilon_{j,it} | Z_{j,it}, \Gamma_{w,j}) = 0$$

$$E(\xi_{j,i} | Z_{j,it}) \neq 0$$

Empirical Strategy

Wage Equation

- At each period t, individual i in spell j earns

$$\log(w_{j,it}) = [z_{j,it}\ h^{w}{}_{j}(t)]\gamma_j + \xi_{j,i} + \varepsilon_{j,it}$$

- with
  - $z_{j,it}$ = time varying and constant Mincerian covariates
  - $h^{w}{}_{j}(t)$ = time effect, function of the worker's potential experience and tenure on the job
  - $\xi_{j,i}$ = unobserved heterogeneity
  - The residuals $\varepsilon_{j,it} \sim N(0, \sigma_j^2)$.
  - $Z_{j,it} = [z_{j,it}\ h^{w}{}_{j}(t)]$
  - $\Gamma_{w,j} = [\gamma_j\ \xi_j]$
- Assuming

$$E(\varepsilon_{j,it} | Z_{j,it}, \Gamma_{w,j}) = 0$$

$$E(\xi_{j,i} | Z_{j,it}) \neq 0$$

Empirical Strategy

Wage Equation

- At each period t, individual i in spell j earns

$$\log(w_{j,it}) = [z_{j,it}\ {h^w}_j(t)]\gamma_j + \xi_{j,i} + \varepsilon_{j,it}$$

- with
  - $z_{j,it}$ = time varying and constant Mincerian covariates
  - ${h^w}_j(t)$ = time effect, function of the worker's potential experience and tenure on the job
  - $\xi_{j,i}$ = unobserved heterogeneity
  - The residuals $\varepsilon_{j,it} \sim N(0, \sigma_j^2)$.
  - $Z_{j,it} = [z_{j,it}\ {h^w}_j(t)]$
  - $\Gamma_{w,j} = [\gamma_j\ \xi_j]$
- Assuming

$$E(\varepsilon_{j,it} | Z_{j,it}, \Gamma_{w,j}) = 0$$

$$E(\xi_{j,i} | Z_{j,it}) \neq 0$$

Empirical Strategy

Hazard Rates

- let $D_{j,it} = 1$ if agent i exits spell j at the end of period t, otherwise $D_{j,it} = 0$.
- Hazard rate of spell j in period $t$is

$$\lambda(t|X_{j,it}, \Gamma_{D,j}) = \Pr(D_{j,it} = 1|X_{j,it}, \Gamma_{D,j})$$

$$\lambda_j(t|X_{j,it}, \Gamma_{D,j}) = \frac{1}{1 + \exp(-([x_{j,it}\ h_j^D(t)]\beta_j + \theta_{j,i}))}$$

- with
  - $x_{j,it}$ = time varying and constant covariates
  - $h^D{}_j(t)$ = time effect or duration dependence
  - $\theta_{j,i}$ = unobserved heterogeneity
  - $\Gamma_{D,j} = [\beta_j\ \theta_j]$

Empirical Strategy

Hazard Rates

- let $D_{j,it} = 1$ if agent i exits spell j at the end of period t, otherwise $D_{j,it} = 0$.
- Hazard rate of spell j in period $t$is

$$\lambda(t|X_{j,it}, \Gamma_{D,j}) = \Pr(D_{j,it} = 1|X_{j,it}, \Gamma_{D,j})$$

$$\lambda_j(t|X_{j,it}, \Gamma_{D,j}) = \frac{1}{1 + \exp(-([x_{j,it}\ h_j^D(t)]\beta_j + \theta_{j,i}))}$$

- with
  - $x_{j,it}$ = time varying and constant covariates
  - $h^D{}_j(t)$ = time effect or duration dependence
  - $\theta_{j,i}$ = unobserved heterogeneity
  - $\Gamma_{D,j} = [\beta_j\ \theta_j]$

Empirical Strategy

Hazard Rates

- let $D_{j,it} = 1$ if agent i exits spell j at the end of period t, otherwise $D_{j,it} = 0$.
- Hazard rate of spell j in period $t$is

$$\lambda(t|X_{j,it}, \Gamma_{D,j}) = \Pr(D_{j,it} = 1|X_{j,it}, \Gamma_{D,j})$$

$$\lambda_j(t|X_{j,it}, \Gamma_{D,j}) = \frac{1}{1 + \exp(-([x_{j,it}\ h_j^D(t)]\beta_j + \theta_{j,i}))}$$

- with
  - $x_{j,it}$ = time varying and constant covariates
  - $h^D{}_j(t)$ = time effect or duration dependence
  - $\theta_{j,i}$ = unobserved heterogeneity
  - $\Gamma_{D,j} = [\beta_j\ \theta_j]$

Empirical Strategy

Hazard Rates

- let $D_{j,it} = 1$ if agent i exits spell j at the end of period t, otherwise $D_{j,it} = 0$.
- Hazard rate of spell j in period $t$is

$$\lambda(t|X_{j,it}, \Gamma_{D,j}) = \Pr(D_{j,it} = 1|X_{j,it}, \Gamma_{D,j})$$

$$\lambda_j(t|X_{j,it}, \Gamma_{D,j}) = \frac{1}{1 + \exp(-([x_{j,it}\ h_j^D(t)]\beta_j + \theta_{j,i}))}$$

- with
  - $x_{j,it}$ = time varying and constant covariates
  - $h^D{}_j(t)$ = time effect or duration dependence
  - $\theta_{j,i}$ = unobserved heterogeneity
  - $\Gamma_{D,j} = [\beta_j\ \theta_j]$

Empirical Strategy

Hazard Rates

- Why the logistic distribution?
  - It tends to a mixed proportional hazard (MPH) function as the hazards get smaller
  - I.E. the hazard rates are proportionally related to covariates, unobserved heterogeneity, and duration dependence
  - The normal distribution could be an alternative if one is willing to sacrifice the MPH property
- The survival function, contribution to the likelihood if right-censored, is

$$S_j(t|X_{j,it},\Gamma_{D,j}) = \prod_{k=1}^{t}\left(1-\lambda_j(k|X_{j,it},\Gamma_{D,j})\right)$$

- The conditional density, contribution to the likelihood if complete spell, is

$$f_j(t|X_{j,it},\Gamma_{D,j}) = \lambda_j(t|X_{j,it},\Gamma_{D,j})S_j(t-1|X_{j,it},\Gamma_{D,j})$$

Empirical Strategy

Hazard Rates

- Why the logistic distribution?
  - It tends to a mixed proportional hazard (MPH) function as the hazards get smaller
  - I.E. the hazard rates are proportionally related to covariates, unobserved heterogeneity, and duration dependence
  - The normal distribution could be an alternative if one is willing to sacrifice the MPH property
- The survival function, contribution to the likelihood if right-censored, is

$$S_j(t|X_{j,it},\Gamma_{D,j}) = \prod_{k=1}^{t}\left(1-\lambda_j(k|X_{j,it},\Gamma_{D,j})\right)$$

- The conditional density, contribution to the likelihood if complete spell, is

$$f_j(t|X_{j,it},\Gamma_{D,j}) = \lambda_j(t|X_{j,it},\Gamma_{D,j})S_j(t-1|X_{j,it},\Gamma_{D,j})$$

Empirical Strategy

Hazard Rates

- Why the logistic distribution?
  - It tends to a mixed proportional hazard (MPH) function as the hazards get smaller
  - I.E. the hazard rates are proportionally related to covariates, unobserved heterogeneity, and duration dependence
  - The normal distribution could be an alternative if one is willing to sacrifice the MPH property
- The survival function, contribution to the likelihood if right-censored, is

$$S_j(t|X_{j,it},\Gamma_{D,j}) = \prod_{k=1}^{t}\left(1-\lambda_j(k|X_{j,it},\Gamma_{D,j})\right)$$

- The conditional density, contribution to the likelihood if complete spell, is

$$f_j(t|X_{j,it},\Gamma_{D,j}) = \lambda_j(t|X_{j,it},\Gamma_{D,j})S_j(t-1|X_{j,it},\Gamma_{D,j})$$

Empirical Strategy

Unobserved Heterogeneity

- Define $(\xi, \theta) \sim G(\Theta)$ as the joint distribution of the unobserved components
- $G(\Theta)$ follows a discrete distribution with M mass points.
- This approach has the following properties
  - Robust to misspecifications of θ (Heckman and Singer (1984))
  - The selection is captured through $G(\Theta)$, i.e. affects
    - Staying at the same firm
    - Wage negotiation
    - Probability of finding a job if unemployed
  - Conditioning on the spell type and (ξ, θ), $\varepsilon_{j,it}$ becomes i.i.d

Empirical Strategy

Likelihoods

- The likelihood contribution for individual i in Model 2 is

$$\begin{aligned} L_{it} &= \sum_{j=1}^{3} d_{j,it} \int \frac{1}{\sigma_j} \phi\left(\frac{\varepsilon_{j,it}}{\sigma_j}\right) \times f_j(t)^{D_{j,it}} \times S_j(t)^{1-D_{j,it}} dG(\Theta) \\ &+ \sum_{k=4}^{5} d_{k,it} \int f_k(t)^{D_{k,it}} \times S_k(t)^{1-D_{k,it}} dG(\Theta) \end{aligned}$$

- where $d_{j,it} = 1$ if i is in spell j at date t

- I maximize the following log likelihood over the set of parameters Ω

$$l(\Omega) = \sum_{i=1}^{N} \sum_{i=1}^{t_i} \log L_{it}(\Omega)$$

$$l_{it}(\Omega) = \log L_{it}(\Omega)$$

Empirical Strategy

Likelihoods

- The likelihood contribution for individual i in Model 2 is

$$\begin{aligned} L_{it} \quad &= \sum_{j=1}^{3} d_{j,it} \int \frac{1}{\sigma_j} \phi\left(\frac{\varepsilon_{j,it}}{\sigma_j}\right) \times f_j(t)^{D_{j,it}} \times S_j(t)^{1-D_{j,it}} dG(\Theta) \\ &+ \sum_{k=4}^{5} d_{k,it} \int f_k(t)^{D_{k,it}} \times S_k(t)^{1-D_{k,it}} dG(\Theta) \end{aligned}$$

- where $d_{j,it} = 1$ if i is in spell j at date t

- I maximize the following log likelihood over the set of parameters Ω

$$l(\Omega) = \sum_{i=1}^{N} \sum_{i=1}^{t_i} \log L_{it}(\Omega)$$

$$l_{it}(\Omega) = \log L_{it}(\Omega)$$

Empirical Strategy

Standard Errors

- Optimization ensures
  - Score values are close to zero ($< 10^{-4}$)
  - Hessian is nonsingular
  - Global optimum
- The standard errors are calculated using the following matrix variance covariance

$$\hat{V} = \hat{A}^{-1}\hat{B}\hat{A}^{-1}$$

$$\hat{A} = -\frac{1}{N}\sum_{i=1}^{N}\sum_{t=1}^{t_i}(\nabla^2_{\Omega} l_{it}(\hat{\Omega}))$$

$$\hat{B} = \frac{1}{N}\sum_{i=1}^{N}\left(\sum_{t=1}^{t_i}\nabla_{\Omega} l_{it}(\hat{\Omega})\right)\left(\sum_{t=1}^{t_i}\nabla_{\Omega} l_{it}(\hat{\Omega})\right)'$$

Empirical Strategy

Standard Errors

- Optimization ensures
  - Score values are close to zero ($< 10^{-4}$)
  - Hessian is nonsingular
  - Global optimum
- The standard errors are calculated using the following matrix variance covariance

$$\hat{V} = \hat{A}^{-1}\hat{B}\hat{A}^{-1}$$

$$\hat{A} = -\frac{1}{N}\sum_{i=1}^{N}\sum_{t=1}^{t_i}(\nabla^2_{\Omega} l_{it}(\hat{\Omega}))$$

$$\hat{B} = \frac{1}{N}\sum_{i=1}^{N}\left(\sum_{t=1}^{t_i}\nabla_{\Omega} l_{it}(\hat{\Omega})\right)\left(\sum_{t=1}^{t_i}\nabla_{\Omega} l_{it}(\hat{\Omega})\right)'$$

Empirical Strategy

Standard Errors

- Optimization ensures
  - Score values are close to zero ($< 10^{-4}$)
  - Hessian is nonsingular
  - Global optimum
- The standard errors are calculated using the following matrix variance covariance

$$\hat{V} = \hat{A}^{-1}\hat{B}\hat{A}^{-1}$$

$$\hat{A} = -\frac{1}{N}\sum_{i=1}^{N}\sum_{t=1}^{t_i}(\nabla^2_{\Omega} l_{it}(\hat{\Omega}))$$

$$\hat{B} = \frac{1}{N}\sum_{i=1}^{N}\left(\sum_{t=1}^{t_i}\nabla_{\Omega} l_{it}(\hat{\Omega})\right)\left(\sum_{t=1}^{t_i}\nabla_{\Omega} l_{it}(\hat{\Omega})\right)'$$