



PARIS | LONDON | BERLIN | MADRID | TORINO BUSINESS SCHOOL

# Les données de panel. Applications en management

19 septembre 2013  
Sandrine Macé



## Plan



1. Les données de panel, de quoi parlons-nous ?
2. Les avantages ? Les inconvénients
3. Spécificités de la modélisation des données de panel

19 septembre 2013

2



# 1. LES DONNÉES DE PANEL

19 septembre 2013

3



## Données individuelles en coupe transversale

| Individu<br>$i$ | Sommes<br>dépensées | Quantités<br>achetées | Prix   | Nb jours depuis<br>dernier achat |
|-----------------|---------------------|-----------------------|--------|----------------------------------|
| <b>1</b>        | $SD_1$              | $QA_1$                | $Pr_1$ | $J_1$                            |
| <b>2</b>        | $SD_2$              | $QA_2$                | $Pr_2$ | $J_2$                            |
| <b>3</b>        | $SD_3$              | $QA_3$                | $Pr_3$ | $J_3$                            |
| <b>4</b>        | $SD_4$              | $QA_4$                | $Pr_4$ | $J_4$                            |
| ...             | ...                 | ...                   | ...    | ...                              |
| <b>n</b>        | $SD_n$              | $QA_n$                | $Pr_n$ | $J_n$                            |



« Photo » instantanée à un moment donné  
Permet d'étudier l'hétérogénéité entre individus

19 septembre 2013

4



## Données temporelles

- Prenons maintenant un seul individu, le n°1 par exemple, et regardons l'évolution de sa consommation dans le temps

| Individu<br><i>i</i> | Temps<br><i>t</i> | Sommes<br>dépensées | Quantités<br>achetées | Prix       | Nb jours<br>depuis dernier<br>achat |
|----------------------|-------------------|---------------------|-----------------------|------------|-------------------------------------|
| 1                    | 1                 | $SD_{1,1}$          | $QA_{1,1}$            | $Pr_{1,1}$ | $J_{1,1}$                           |
| 1                    | 2                 | $SD_{1,2}$          | $QA_{1,2}$            | $Pr_{1,2}$ | $J_{1,2}$                           |
| 1                    | 3                 | $SD_{1,3}$          | $QA_{1,3}$            | $Pr_{1,3}$ | $J_{1,3}$                           |
| 1                    | 4                 | $SD_{1,4}$          | $QA_{1,4}$            | $Pr_{1,4}$ | $J_{1,4}$                           |
| ...                  | ...               | ...                 | ...                   | ...        | ...                                 |
| 1                    | <i>t</i>          | $SD_{1,t}$          | $QA_{1,t}$            | $Pr_{1,t}$ | $J_{1,t}$                           |

➔ Permet d'étudier le comportement dynamique, les évolutions des relations dans le temps



## Mélangeons les 2 types de données

Données en coupe  
transversale  
(*cross sectional*)

+

Données  
temporelles  
(*time series*)

= Données de panel  
(*cross sectional time series data  
or longitudinal*)



## Mélangeons les 2 types de données

| Individu i | Temps t | Sommes dépensées | Quantités achetées | Prix       | Nb jours depuis dernier achat |
|------------|---------|------------------|--------------------|------------|-------------------------------|
| 1          | 1       | $SD_{1,1}$       | $QA_{1,1}$         | $Pr_{1,1}$ | $J_{1,1}$                     |
| 1          | 2       | $SD_{1,2}$       | $QA_{1,2}$         | $Pr_{1,2}$ | $J_{1,2}$                     |
| ...        | ...     | ...              | ...                | ...        | ...                           |
| 1          | t       | $SD_{1,t}$       | $QA_{1,t}$         | $Pr_{1,t}$ | $J_{1,t}$                     |
| 2          | 1       | $SD_{2,1}$       | $QA_{2,1}$         | $Pr_{2,1}$ | $J_{2,1}$                     |
| 2          | 2       | $SD_{2,2}$       | $QA_{2,2}$         | $Pr_{2,2}$ | $J_{2,2}$                     |
| ...        | ...     | ...              | ...                | ...        | ...                           |
| 2          | t       | $SD_{2,t}$       | $QA_{2,t}$         | $Pr_{2,t}$ | $J_{2,t}$                     |
| 3          | 1       | $SD_{3,1}$       | $QA_{3,1}$         | $Pr_{3,1}$ | $J_{3,1}$                     |
| 3          | 2       | $SD_{3,2}$       | $QA_{3,2}$         | $Pr_{3,2}$ | $J_{3,2}$                     |
| ...        | ...     | ...              | ...                | ...        | ...                           |
| 3          | t       | $SD_{3,t}$       | $QA_{3,t}$         | $Pr_{3,t}$ | $J_{3,t}$                     |
| ...        | ...     | ...              | ...                | ...        | ...                           |
| n          | 1       | $SD_{n,1}$       | $QA_{n,1}$         | $Pr_{n,1}$ | $J_{n,1}$                     |
| n          | 2       | $SD_{n,2}$       | $QA_{n,2}$         | $Pr_{n,2}$ | $J_{n,2}$                     |
| ...        | ...     | ...              | ...                | ...        | ...                           |
| n          | t       | $SD_{n,t}$       | $QA_{n,t}$         | $Pr_{n,t}$ | $J_{n,t}$                     |

7

19 septembre 2013



## De quoi parlons-nous ?

- Les données de panel sont des observations relatives aux mêmes individus (entreprises, ménages, salariés, ...) à des périodes successives.
- Elles possèdent les deux dimensions :
  - Temporelle : longitudinale
  - Individuelle : transversale
- Données comportementales, collectées systématiquement voire automatiquement

8

19 septembre 2013



## Quelques exemples

- Economie
  - Ménages, entreprises, pays, ...
  - Pauvreté, chômage, endettement des entreprises, défaillance, financement des entreprises, ...
- Marketing
  - Panels de ménages, Panels de magasin, ecommerce, ...
  - Décisions d'achat(quoi, combien, quand), élasticité-prix, ...
- RH
  - Entreprises, salariés, ...
  - Turn over, efficacité d'une formation, etc.
- Comptabilité & Finance
  - Entreprises
  - Rendement d'un titre, beta, bénéfices par action, coûts du capital, informations comptables, ...



## EXEMPLE N°1



## Does fast food marketing cause consumption to rise? (1/4)



- Et si la communication faite par la restauration rapide était responsable de l'obésité ?
- Les promotions par réduction de prix ont-elles un impact sur la consommation totale en fast food ?
  - Provoquent-elles une augmentation des quantités consommées ?
  - Poussent-elles à consommer plus souvent ?
- NON, répond la restauration rapide !  
La concurrence est forte et les effets des promotions s'annulent (brand switching), n'augmentant pas la demande primaire.
- Alors, augmentation de la demande primaire ou impact sur la demande secondaire ?

11



## Does fast food marketing cause consumption to rise? (2/4)



- Résultats d'une recherche menée au Canada
- Données de panel NPD :
  - 5657 visites de restaurants par 139 ménages pendant 6 ans
  - Variables socio démographiques
  - Nature des achats (produit, prix, ...)
  - Caractéristiques du restaurant
- Modèle stochastique combinant la probabilité d'aller dans un fast food (versus autre restaurant), la probabilité de choisir un fast food donné & quantités achetées

12



## Does fast food marketing cause consumption to rise? (3/4)

**Table 1. Summary of Household and Fast Food Data**

| Variable <sup>a</sup> | N                 | Mean  | Std. Dev. | Min.    | Max.    |          |
|-----------------------|-------------------|-------|-----------|---------|---------|----------|
| Individu              | Age               | 5,657 | 48.937    | 12.319  | 21.000  | 86.000   |
|                       | Household size    | 5,657 | 2.773     | 1.306   | 1.000   | 6.000    |
|                       | Marital status    | 5,657 | 0.744     | 0.437   | 0.000   | 1.000    |
|                       | Education         | 5,657 | 3.248     | 1.451   | 0.000   | 6.000    |
|                       | Occupation        | 5,657 | 0.489     | 0.500   | 0.000   | 1.000    |
| Stimulus              | Income            | 5,657 | 53.419    | 23.885  | 7.500   | 80.000   |
|                       | Combo             | 5,657 | 0.146     | 0.353   | 0.000   | 1.000    |
|                       | Buy one, get one  | 5,657 | 0.022     | 0.148   | 0.000   | 1.000    |
| Dépense               | Special           | 5,657 | 0.345     | 0.475   | 0.000   | 1.000    |
|                       | Spending per trip | 5,657 | \$1.015   | \$1.004 | \$0.052 | \$34.500 |
|                       | Price             | 5,657 | \$1.457   | \$2.165 | \$0.043 | \$57.500 |
| Produit consommé      | Grams             | 5,657 | 1.008     | 0.841   | 0.022   | 7.857    |
|                       | Calories          | 5,657 | 1.342     | 1.151   | 0.000   | 10.440   |
|                       | Protein grams     | 5,657 | 0.053     | 0.036   | 0.000   | 0.273    |
|                       | Fat grams         | 5,657 | 0.065     | 0.041   | 0.000   | 0.271    |
|                       | Carbo grams       | 5,657 | 0.165     | 0.088   | 0.000   | 0.728    |
|                       | Water grams       | 5,657 | 0.716     | 0.165   | 0.040   | 1.087    |

Note: Marital status is defined as 0 = single, 1 = married; Education is from 0 = no high school, to 6 = post-graduate degree; Occupation is 0 = blue collar, 1 = white collar; Combo, BOGO and Special are defined as 0 = no promotion and 1 = promotion. <sup>a</sup>Income, grams, and calories are reported in '000 of units. Price is in \$/100 gram. All values per meal.



## Yes, fast food marketing does cause consumption to rise...! (4/4)

| Restaurant <sup>a</sup> | Unit Sales |           |
|-------------------------|------------|-----------|
|                         | Primary    | Secondary |
| McDonalds               | 0.845      | 0.155     |
| A&W                     | 0.823      | 0.177     |
| Subway                  | 0.797      | 0.203     |
| Burger King             | 0.869      | 0.131     |
| Wendy's                 | 0.842      | 0.158     |
| KFC                     | 0.732      | 0.268     |
| Dairy Queen             | 0.639      | 0.361     |
| Pizza Hut               | 0.589      | 0.410     |
| Harvey's                | 0.798      | 0.201     |
| Arby's                  | 0.807      | 0.193     |
| Other Fast Food         | 0.764      | 0.236     |
| Mr. Submarine           | 0.754      | 0.246     |
| Pizza Pizza             | 0.317      | 0.683     |
| La Belle Province       | 0.849      | 0.151     |
| Little Caesar's         | 0.419      | 0.581     |
| Taco Bell               | 0.014      | 0.986     |
| Dominos                 | 0.407      | 0.593     |
| Cultures                | 0.658      | 0.342     |
| Taco Time               | 0.823      | 0.176     |
| Pizza Delight           | 0.703      | 0.296     |
| Average                 | 0.673      | 0.327     |

the average primary response (purchase incidence & quantities) is 67.3% while the average secondary, (brand-switching) response is 32.7%.





## EXEMPLE N°2

19 septembre 2013

15



### Les cabinets d'audit, source de valeur ?

- There is very little empirical evidence on the value of an external audit per se.
- What is the value of voluntary external audits of financial statements with respect to the cost of debt?
  - Do private companies with an external audit pay a significantly lower interest rate on their debt than do private companies without an audit?
- Données : KIS-DATA (agence de notation de crédit en Corée), 1987-2002, 2 916 entreprises auditées volontairement versus 69 961 non auditées



## Les variables étudiées



### Variable definitions

| Variable                           |   | Definition   |
|------------------------------------|---|--|
| 1. Dependent variable              |   |  |
| <i>Interest Rate Spread</i>        | = | Interest rate – the average prime rate for the year. The interest rate in year $t$ is measured as interest expenses in year $t$ divided by the average of short-term and long-term debt at the beginning and end of year $t$ . |
| 2. Test variables                  |   |  |
| <i>VAudit</i>                      | = | 1 if a company has voluntary audits in year $t$ and 0 otherwise.   |
| <i>Big4</i>                        | = | 1 if a company is audited by one of the Big 4 (previous 5, 6, and 8) auditors and 0 otherwise.   |
| 3. Firm-specific control variables |   |  |
| <i>Size</i>                        | = | Log (total assets at the end of the fiscal year).  |
| <i>Profitability</i>               | = | Ratio of profit to sales. Profit represents income before extraordinary items.   |
| <i>Leverage</i>                    | = | Ratio of total liabilities to total assets.  |
| <i>Tangible</i>                    | = | Ratio of tangible assets to total assets. Tangible assets represent the difference between non-current assets and intangible assets.   |
| <i>Sales Growth</i>                | = | Percentage change in sales dollars in year $t$ .   |
| <i>Negative Equity Dummy</i>       | = | 1 if a company has negative equity and 0 otherwise.  |

Kim, Simunic, Stein & Yi, 2011

19 septembre 2013



## Modèle



$$\text{Interest Rate Spread} = \beta_0 + \beta_1 \text{VAudit} + \beta_2 \text{VAudit} * \text{Big4} + \sum \gamma_k \text{Control}_k + (\text{Year Dummies}) + (\text{Industr Dummies}) + \text{error}$$

- Si  $\beta_1, \beta_2$  sont négatifs, nous pouvons conclure à un impact de l'audit volontaire

18

19 septembre 2013



TABLE 4  
Regression results using the combined sample of unaudited companies and companies with voluntary audits

| Variable                                  | (1)    |           | (2)           |           | (3)    |           | (4)           |           |
|---|--------|-----------|---------------|-----------|--------|-----------|---------------|-----------|
|   | Pooled |           | Fixed effects |           | Pooled |           | Fixed effects |           |
|   | Coef.  | t-stat    | Coef.         | t-stat    | Coef.  | t-stat    | Coef.         | t-stat    |
| <i>VAudit</i> ( $\beta_1$ )               | -1.240 | -11.87*** | -0.562        | -3.69***  | -1.151 | -9.52***  | -0.501        | -2.83**   |
| <i>VAudit</i> * <i>Big4</i> ( $\beta_2$ ) |        |           |               |           | -0.261 | -1.41     | -0.195        | -0.68     |
| <i>Size</i>                               | -0.013 | -0.41     | -0.120        | -1.71*    | -0.013 | -0.40     | -0.118        | -1.68*    |
| <i>Profitability</i>                      | -0.277 | -0.90     | -1.565        | -4.29***  | -0.278 | -0.90     | -1.566        | -4.30***  |
| <i>Leverage</i>                           | 1.025  | 8.07***   | -1.776        | -12.30*** | 1.024  | 8.07***   | -1.777        | -12.31*** |
| <i>Tangible</i>                           | -4.205 | -31.18*** | -3.658        | -16.85*** | -4.205 | -31.18*** | -3.657        | -16.85*** |
| <i>Sales Growth</i>                       | -0.135 | -5.19***  | 0.269         | 10.57***  | -0.135 | -5.18***  | 0.269         | 10.57***  |
| <i>Neg Eq. Dummy</i>                      | 1.654  | 12.01***  | 1.297         | 7.19***   | 1.657  | 12.04***  | 1.298         | 7.20***   |
| <i>Intercept</i>                          | 4.080  | 7.81***   | 6.786         | 6.80***   | 4.076  | 7.81***   | 6.763         | 6.77***   |
| <i>Year dummies</i>                       |        | Yes       |               | Yes       |        | Yes       |               | Yes       |
| <i>Industry dummies</i>                   |        | Yes       |               | No        |        | Yes       |               | No        |
| Adj. $R^2$                                |        | 0.075     |               | 0.065     |        | 0.075     |               | 0.065     |
| <i>N</i> (years)                          |        | 72,577    |               | 72,577    |        | 72,577    |               | 72,577    |

Notes

All variables are as defined in the appendix. Columns 1 and 3 report the results of pooled ordinary least squares regressions. Columns 2 and 4 report the results of firm fixed effects regressions. Reported t-values are on an adjusted basis using standard errors corrected for firm-level clustering. \*\*\*, \*\*, and \* denote the 1%, 5%, and 10% levels of significance, respectively.



## Modèle sur les changements

- What is the impact of a change in borrower status, from being unaudited to being audited, on the interest rate spread?

$$\Delta \text{Interest Rate Spread} = \beta_0 + \beta_1 \Delta \text{Engagement}(N, V) + \Delta \text{Engagement}(N, M) + \sum \gamma_k \Delta \text{Control}_k + (\text{Industry \& Year Dummies}) + \text{error} \quad (2),$$

where  $\Delta \text{Engagement}(N, V)$  is a dummy variable that equals one when the engagement status changes from no audit in year  $t - 1$  to a voluntary audit in year  $t$  and zero otherwise;  $\Delta \text{Engagement}(N, M)$  is a dummy variable that equals one when the engagement status changes from no audit in year  $t - 1$  to a mandatory audit in year  $t$  and zero otherwise;  $\Delta$

- Results: change in engagement status from being unaudited to being voluntarily or mandatorily audited leads to significant savings in the interest cost of borrowing.

- $\beta_1$  is more than two times larger than  $\beta_2$



## EXEMPLE N°3

19 septembre 2013

21



## La mobilité externe

- Combien de temps reste-t-on salarié dans une entreprise ?
- Y a-t-il des différences entre hommes et femmes ?
- Quelles sont les variables explicatives ?
  
- Données provenant d'une base de données danoises, Integrated Database for Labour Market Research (IDA), qui comprend de l'information sur les employés de tous les établissements du secteur privé au Danemark, 1980-2000
  
- 638 515 observations (5% de l'échantillon total)
  
- Discrete time duration models with group-level heterogeneity:
  - modèle logit sur une variable égale à 1 si l'individu sort de l'entreprise
  - Prise en compte dans les variables explicatives d'une variable mesurant l'ancienneté dans l'entreprise



Table 1  
Descriptive statistics, 1980–2000

|                                      | Women              | Men                | All                |
|--------------------------------------|--------------------|--------------------|--------------------|
| Gender                               | –                  | –                  | 0.386              |
| Age less than 30 years               | 0.491              | 0.450              | 0.466              |
| Age 30–50 years                      | 0.377              | 0.400              | 0.391              |
| Age above 50 years                   | 0.132              | 0.150              | 0.143              |
| Unskilled                            | 0.427              | 0.375              | 0.395              |
| Skilled                              | 0.541              | 0.583              | 0.567              |
| High skilled                         | 0.032              | 0.042              | 0.038              |
| Children                             | 0.345              | 0.345              | 0.345              |
| Married/cohabiting                   | 0.458              | 0.467              | 0.464              |
| Manufacturing                        | 0.234              | 0.327              | 0.291              |
| Primary sector                       | 0.019              | 0.043              | 0.033              |
| Electricity, gas and water supply    | 0.003              | 0.009              | 0.007              |
| Construction                         | 0.024              | 0.132              | 0.090              |
| Retail and trade                     | 0.310              | 0.238              | 0.266              |
| Transportation                       | 0.043              | 0.073              | 0.062              |
| Financial                            | 0.202              | 0.127              | 0.156              |
| Service                              | 0.166              | 0.052              | 0.096              |
| Years of tenure                      | 2.314<br>(3.133)   | 2.471<br>(3.241)   | 2.411<br>(3.201)   |
| Workplace size*                      | 198<br>(63)        | 208<br>(62)        | 204<br>(62)        |
| Workplace size*<br>(lagged one year) | 186<br>(61)        | 195<br>(61)        | 192<br>(61)        |
| Payroll per worker*<br>(1980-prices) | 78,361<br>(39,325) | 90,107<br>(37,154) | 85,576<br>(38,434) |
| # Observations                       | 246,316            | 392,199            | 638,515            |

Note: Based on a 5% sample. Standard deviations are in parentheses.  
\*These numbers are employee-weighted.

Fredericksen, Honoré & Hu, 2007

15 SEPTEMBER 2010

Job separation models, men

|                                   | Conventional<br>logit model | Conventional<br>logit model | Conventional<br>logit model | Fixed-effects<br>model,<br>$\tau = \max$ | Fixed-effects<br>model,<br>$\tau = \max$ |
|-----------------------------------|-----------------------------|-----------------------------|-----------------------------|--|--|
| Constant                          | 0.110<br>(0.020)            | 0.477<br>(0.021)            | 0.341<br>(0.023)            | –  | –  |
| Age less than 30 years            | –                           | –                           | –                           | –  | –  |
| Age 30–50 years                   | –0.361<br>(0.010)           | –0.281<br>(0.010)           | –0.286<br>(0.010)           | –0.442<br>(0.017)                        | –0.438<br>(0.017)                        |
| Age more than 50 years            | –0.355<br>(0.014)           | –0.277<br>(0.014)           | –0.278<br>(0.014)           | –0.354<br>(0.022)                        | –0.348<br>(0.022)                        |
| Unskilled                         | –                           | –                           | –                           | –  | –  |
| Skilled                           | –0.018<br>(0.008)           | 0.041<br>(0.008)            | 0.040<br>(0.008)            | 0.022<br>(0.013)                         | 0.025<br>(0.013)                         |
| High skilled                      | –0.150<br>(0.020)           | –0.001<br>(0.020)           | –0.020<br>(0.021)           | –0.081<br>(0.035)                        | –0.076<br>(0.035)                        |
| Children                          | –0.133<br>(0.010)           | –0.120<br>(0.010)           | –0.116<br>(0.010)           | –0.115<br>(0.016)                        | –0.114<br>(0.016)                        |
| Married/cohabiting                | –0.135<br>(0.011)           | –0.105<br>(0.011)           | –0.099<br>(0.011)           | –0.110<br>(0.016)                        | –0.110<br>(0.016)                        |
| Lagged workplace size*            | –                           | 0.078<br>(0.015)            | 0.138<br>(0.016)            | –  | 0.795<br>(0.094)                         |
| Lagged workplace size**           | –                           | –0.029<br>(0.003)           | –0.039<br>(0.003)           | –  | –0.086<br>(0.017)                        |
| Payroll per worker**              | –                           | –0.796<br>(0.014)           | –0.806<br>(0.015)           | –  | –0.362<br>(0.059)                        |
| Std. dev. of payroll per worker** | –                           | 0.224<br>(0.020)            | 0.247<br>(0.020)            | –  | 0.149<br>(0.065)                         |
| Manufacturing                     | –                           | –                           | –                           | –  | –  |
| Primary sector                    | –                           | –                           | 0.402<br>(0.018)            | –  | –  |
| Electricity, gas and water supply | –                           | –                           | –0.268<br>(0.049)           | –  | –  |
| Construction                      | –                           | –                           | 0.141<br>(0.012)            | –  | –  |
| Retail and trade                  | –                           | –                           | 0.105<br>(0.010)            | –  | –  |
| Transportation                    | –                           | –                           | 0.299<br>(0.015)            | –  | –  |
| Financial                         | –                           | –                           | 0.257<br>(0.013)            | –  | –  |
| Service                           | –                           | –                           | 0.084<br>(0.018)            | –  | –  |
| Year dummies                      | YES                         | YES                         | YES                         | YES                                      | YES                                      |
| Tenure dummies                    | YES                         | YES                         | YES                         | YES                                      | YES                                      |
| Log likelihood/objective function | –227,456                    | –225,269                    | –224,742                    | –111,454                                 | –111,159                                 |

Note: Based on 392,199 observations.

Fredericksen, Honoré & Hu, 2007

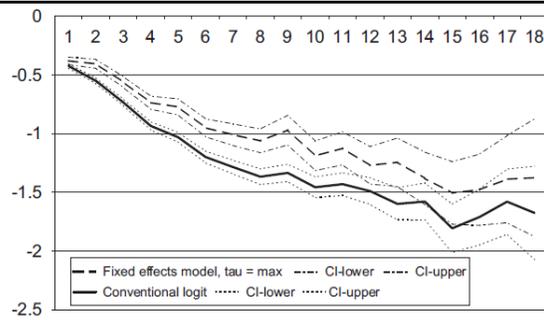
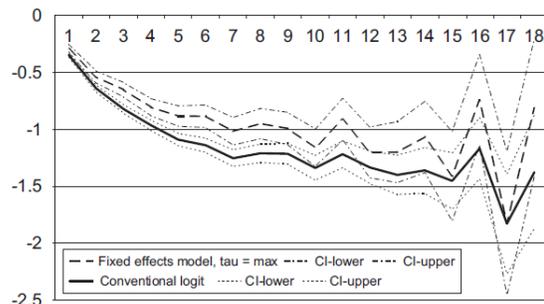
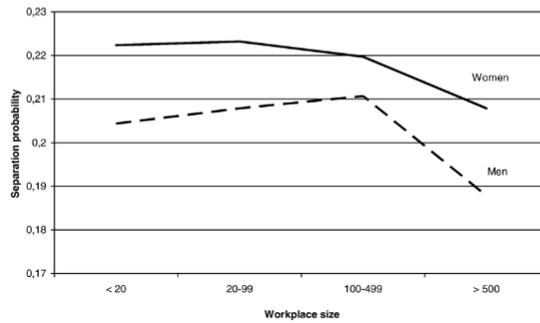
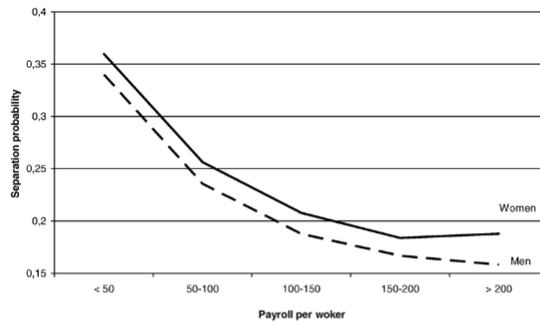


Fig. 1. Coefficients on tenure dummies (men).



Fredericksen, Honoré & Hu, 2007



Fredericksen 2004

19 septembre 2013



## EXEMPLE N°4

19 septembre 2013

27



### Les enchères inversées dans le B2B (1/3)

- Recherches essentiellement centrées sur les résultats des enchères (prix final, gain financier)
- Quel est le comportement des enchérisseurs ? Peut-on expliquer pourquoi il est plus ou moins offensif ?
- Les relations entre acheteurs et fournisseurs en dehors des enchères agissent-elles sur les offres d'enchères ?
  - Les fournisseurs qui souhaitent établir une relation à long terme avec l'acheteur présentent-ils un comportement plus offensif pendant les enchères ?
  - Plus d'offres soumises ?
  - Plus rapidement ?
  - Prix concédé plus important ?

28



## Les enchères inversées dans le B2B (2/3)



- Information collectée :
  - suivi complet de 12 événements d'enchères inversées, comportant 60 participants
  - questionnaire adressé aux enchérisseurs : satisfaction sur la relation avec l'entreprise acheteuse, vision globale ou individuelle, investissements spécifiques en personnels, technologies, capacités de production, etc. dédiés à l'entreprise acheteuse
- 132 observations (44 enchérisseurs différents)
- Analyse du comportement de l'enchérisseur au cours des offres (lots)

29

19 septembre 2013



## Les enchères inversées dans le B2B (3/3)



- Caractère offensif mesuré 3 façons différentes
  - nombre d'offres soumises
  - rapidité dans les offres (nb total d'offres / durée de l'enchère)
  - prix concédé (différence entre 1<sup>er</sup> prix et prix final)
- Caractère offensif = f(Etat des relations avant l'enchère; Fournisseur officiel; nombre total d'enchérisseurs, nombre de soumissions, durée de l'enchère)
- Résultat : Comportement moins offensif lorsque l'enchérisseur est un fournisseur officiel, beaucoup d'acteurs pendant l'enchère mais peu d'offres soumises, souhait d'investir dans l'entreprise acheteuse

30

19 septembre 2013



# UNE 1<sup>ÈRE</sup> CONCLUSION

19 septembre 2013

31



## Pourquoi les données de panels sont-elles bien adaptées au management ?

- La disponibilité des données de panel s'accroît.
- Le management est dynamique.
- Possibilité de mesurer l'impact de plusieurs actions (stimuli), seules ou simultanées et les synergies (interactions).
- Possibilité de mesurer l'effet provenant de l'âge ou bien de la cohorte.

19 septembre 2013

32



## 2. AVANTAGES & INCONVÉNIENTS

19 septembre 2013

33



### Avantages

- Données nombreuses ayant pour conséquence une précision des estimations
  - Propriétés asymptotiques, biais et variance de l'estimation tendent vers 0;
- Modélisation visant à expliquer l'hétérogénéité individuelle et temporelle observable
- Permet de tenir compte de l'influence de variables omises, sources d'hétérogénéité non observable :
  - Omission de caractéristiques individuelles
  - Ou omission de variables explicatives communes à tous les individus mais variant dans le temps
- Possibilité d'analyse multiniveau (hiérarchique)
  - Salariés / entreprise d'appartenance

34

19 septembre 2013



## Hétérogénéité : observée et non-observée



| Exemple             | Hétérogénéité observable   | Hétérogénéité non observable          |
|---------------------|--|---------------------------------------|
| <b>Entreprise</b>   | Indicateurs de performance : Rentabilité CA,                       | Qualité des équipes de management     |
| <b>Salarié</b>      | Ancienneté, salaire, fonction, nombre de personnes encadrées, etc. | Motivation, dynamisme                 |
| <b>Consommateur</b> | Nature des achats : quantité, catégorie de produits, prix, etc...  | Satisfaction                          |
| <b>Travailleur</b>  | Temps passé au chômage : formation, sexe, ...                      | Efforts dans la recherche d'un emploi |

Caractéristiques constantes ou variant avec le temps

Caractéristiques individuelles constantes dans le temps

35

19 septembre 2013



## Inconvénients



- Constitution de l'échantillon (couverture)
- Données manquantes partielles ou complètes (biais de sélection)
- Présence de points atypiques à détecter

36

19 septembre 2013



### 3. SPÉCIFICITÉS DE LA MODÉLISATION DES DONNÉES DE PANEL

19 septembre 2013

37



#### Deux types de modèles

##### 1. Modèles stochastiques

- « *A stochastic model is a model in which the probability components are built in at the outset rather than being added ex post facto to accommodate discrepancies between predicted and actual results* » Massy, Montgomery & Morrison, 1970

##### 2. Modèles économétriques

- Dans les 2 cas, l'enjeu est d'intégrer une hétérogénéité non observée.

19 septembre 2013

38



## Modèles stochastiques

- On prend en compte la nature probabiliste du comportement étudié : une loi de distribution est appliquée sur le phénomène étudié.
- L'hétérogénéité non observée est prise en compte en ajoutant une loi de distribution sur un des paramètres (Mixture)



## Par exemple, en marketing

- Purchase incidence:
  - how many purchases will occur in a specified interval of time?
  - Poisson & Negative Binomial Distribution (hétérogénéité gamma)
- Purchase timing:
  - interpurchase time
  - Erlang 2, Gamma, Hazard model
- Brand choice:
  - Binomial, multinomial
  - Beta-binomial, Dirichlet multinomial
- Integrated models of incidence, timing and choice



## Notations

- $i$  : individu
- $t$  : temps
  
- $Y$  : variable à expliquer
- $y_{it}$  : valeur de  $Y$  pour l'individu  $i$  au temps  $t$
- $X_k$  : variable explicative
- $x_{kit}$  : valeur de  $X_k$  pour l'individu  $i$  au temps  $t$
  
- $b_0$  : constante
- $b_k$  : paramètre à estimer pour la variable  $X_k$
- $w_{it}$  : terme d'erreur



## Modèles économétriques Régresser $y_{it}$ par $x_{kit}$

- Modèle que l'on souhaiterait estimer :

$$y_{it} = b_{0it} + \sum_k b_{kit} \cdot x_{it}^k + w_{it}$$

- Deux modèles très fréquemment utilisés pour intégrer l'hétérogénéité non observée :
  - *Modèle à effets fixes* : Les effets individuels non observés sont corrélés aux autres variables explicatives  $X_k$
  - *Modèle à erreurs composées* (effets aléatoires) : l'hétérogénéité non observée varie de façon aléatoire selon les individus



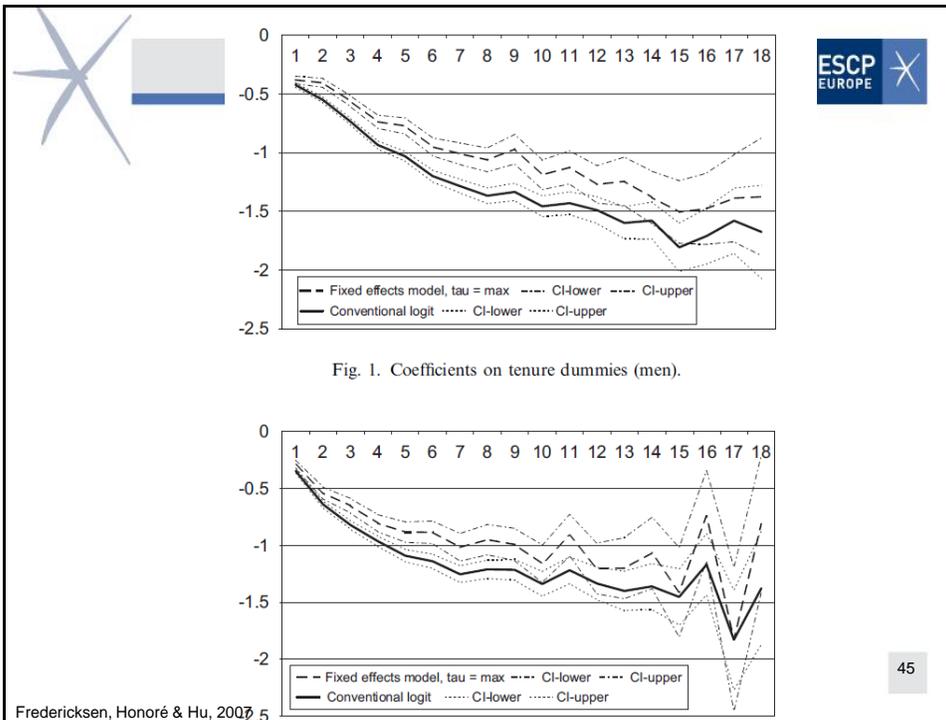
## Attention, danger

- Ne pas prendre en compte l'hétérogénéité non observée risque de conduire à des estimations biaisées.



Job separation models, men

|                                     | Conventional<br>logit model | Conventional<br>logit model | Conventional<br>logit model | Fixed-effects<br>model,<br>$\tau = \max$ | Fixed-effects<br>model,<br>$\tau = \max$ |
|-------------------------------------|-----------------------------|-----------------------------|-----------------------------|--|--|
| Constant                            | 0.110<br>(0.020)            | 0.477<br>(0.021)            | 0.341<br>(0.023)            |  |  |
| Age less than 30 years              | -                           | -                           | -                           | -  | -  |
| Age 30-50 years                     | -0.361<br>(0.010)           | -0.281<br>(0.010)           | -0.286<br>(0.010)           | -0.442<br>(0.017)                        | -0.438<br>(0.017)                        |
| Age more than 50 years              | -0.355<br>(0.014)           | -0.277<br>(0.014)           | -0.278<br>(0.014)           | -0.354<br>(0.022)                        | -0.348<br>(0.022)                        |
| Unskilled                           | -                           | -                           | -                           | -  | -  |
| Skilled                             | -0.018<br>(0.008)           | 0.041<br>(0.008)            | 0.040<br>(0.008)            | 0.022<br>(0.013)                         | 0.025<br>(0.013)                         |
| High skilled                        | -0.150<br>(0.020)           | -0.001<br>(0.020)           | -0.020<br>(0.021)           | -0.081<br>(0.035)                        | -0.076<br>(0.035)                        |
| Children                            | -0.133<br>(0.010)           | -0.120<br>(0.010)           | -0.116<br>(0.010)           | -0.115<br>(0.016)                        | -0.114<br>(0.016)                        |
| Married/cohabiting                  | -0.135<br>(0.011)           | -0.105<br>(0.011)           | -0.099<br>(0.011)           | -0.110<br>(0.016)                        | -0.110<br>(0.016)                        |
| Lagged workplace size*              |                             | 0.078<br>(0.015)            | 0.138<br>(0.016)            |  | 0.795<br>(0.094)                         |
| Lagged workplace size <sup>2a</sup> |                             | -0.029<br>(0.003)           | -0.039<br>(0.003)           |  | -0.086<br>(0.017)                        |
| Payroll per worker**                |                             | -0.796<br>(0.014)           | -0.806<br>(0.015)           |  | -0.362<br>(0.059)                        |
| Std. dev. of payroll per worker**   |                             | 0.224<br>(0.020)            | 0.247<br>(0.020)            |  | 0.149<br>(0.065)                         |
| Manufacturing                       |                             |                             | -                           |  |  |
| Primary sector                      |                             |                             | 0.402<br>(0.018)            |  |  |
| Electricity, gas and water supply   |                             |                             | -0.268<br>(0.049)           |  |  |
| Construction                        |                             |                             | 0.141<br>(0.012)            |  |  |
| Retail and trade                    |                             |                             | 0.105<br>(0.010)            |  |  |
| Transportation                      |                             |                             | 0.299<br>(0.015)            |  |  |
| Financial                           |                             |                             | 0.257<br>(0.013)            |  |  |
| Service                             |                             |                             | 0.084<br>(0.018)            |  |  |
| Year dummies                        | YES                         | YES                         | YES                         | YES                                      | YES                                      |
| Tenure dummies                      | YES                         | YES                         | YES                         | YES                                      | YES                                      |
| Log likelihood/objective function   | -227,456                    | -225,269                    | -224,742                    | -111,454                                 | -111,159                                 |



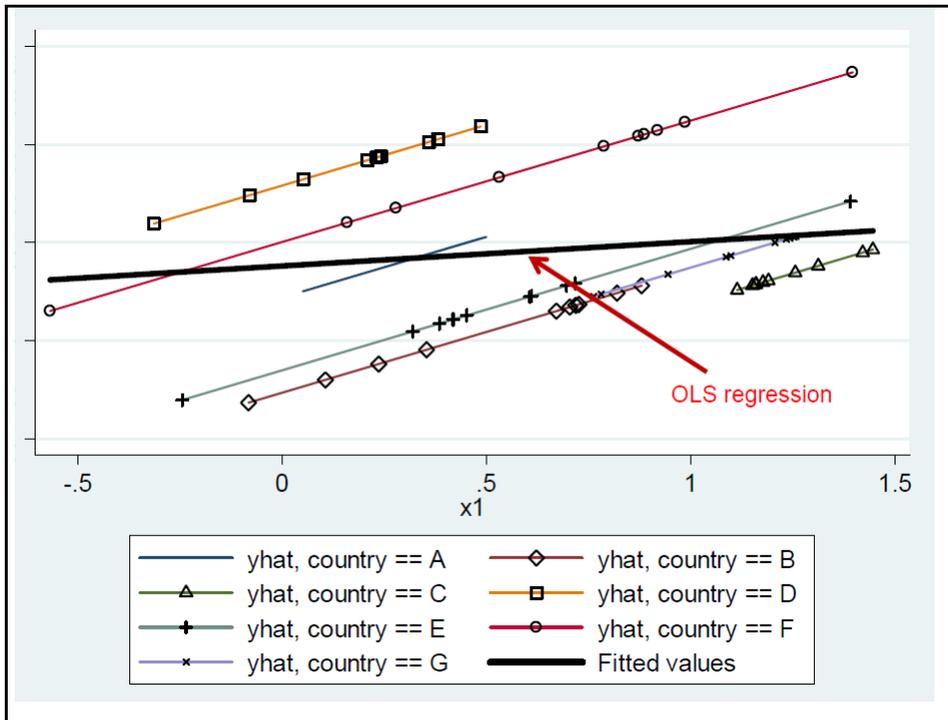
### 3.1. Modèle à effets fixes

- Les effets individuels non observés sont corrélés aux autres variables explicatives  $X_k$
- Ajoutons un effet spécifique fixe (ne varie pas dans le temps) par individu ( $a_i$ ) et un effet spécifique fixe pour chaque moment observé ( $d_t$ ). Alors :

$$b_{0it} = b_0 + a_i + d_t$$

$$y_{it} = b_0 + a_i + d_t + \sum_k b_k \cdot x_{kit} + w_{it}$$

- L'hétérogénéité non observée est captée par une constante spécifique à chaque individu





### 3.2. Modèle à effets aléatoires



- Modèle à erreurs composées
- Les effets spécifiques individuels ( $a_i$ ) et temporels ( $d_t$ ) ne sont pas constants mais suivent une distribution aléatoire.

$$y_{it} = b_0 + \sum_k b_k \cdot x_{kit} + \varepsilon_{it}$$

$$\varepsilon_{it} = u_i + v_t + w_{it}$$

$\uparrow$   
 Hétérogénéité  
aléatoire spécifique  
à l'individu  $i$

$\uparrow$   
 Effet temporel  
aléatoire

$$u_i \sim iid(0; \sigma_u^2)$$

$$v_t \sim iid(0; \sigma_v^2)$$

$$w_{it} \sim iid(0; \sigma_w^2)$$

$$\sigma_\varepsilon^2 = \sigma_u^2 + \sigma_v^2 + \sigma_w^2$$

48

19 septembre 2013



- L'estimateur des MCO n'est plus le meilleur estimateur linéaire sans biais.
- Estimateur des moindres carrés généralisés (MCG)
- Estimateur des moindres carrés quasi généralisés (MCQG)
  
- Modèles non linéaires : hypothèse sur la loi de distribution des  $u_i$
  
- Pas de corrélation entre les effets individuels non observés et les autres régresseurs  $X_k$ 
  - Hypothèse restrictive
  - Parfois problème de non convergence



### Fixe ou aléatoire, quel modèle utiliser ?

- Test de Hausman (1978) : orthogonalité entre les effets aléatoires et les régresseurs



## En conclusion

- Les données de panel, un format de données très approprié aux problématiques du management
- Les possibilités de modélisations sont nombreuses :
  - Modèles linéaires
  - Modèles de choix discrets
  - Modèles de comptage
  - Modèle de durée
  - Modèles dynamiques
  - Les panels spatiaux
  - ...



## Références

- Batalgi, B. H., 2008, *Econometric analysis of panel data*, 4. ed. New York: John Wiley and Sons.
- J.-J. Dreesbeke & G. Saporta, 2010, *Analyse statistique des données longitudinales*, Technip
- Greene, W. H., 2007, *Econometric analysis*. 6 ed. Upper Saddle River: Prentice Hall.
- Maddala, G. S., 1993, *The econometrics for panel data*. Brookfield: Elgar.
- A. Pirotte, 2011, *Econométrie des données de panel*, Economica
- Sevestre A., 2002, *Econométrie des données de panel*, Dunod
- Wooldridge J. M. (2001), *Econometric analysis of cross section and panel data*, 2. ed. Cambridge: MIT Press, 2010.