

Séance 1 – Introduction

Pierre Pora

Qui suis-je

- ▶ Administrateur de l'Insee
- ▶ Actuellement chargé de mission à l'Unité des études démographiques et sociales à la Direction générale de l'Insee
- ▶ Par ailleurs chercheur affilié au Crest-Ensaë et à l'Ined
- ▶ Thématiques de recherche : économie du travail, économie de la famille, en particulier parentalité et inégalités de genre
- ▶ Dans la statistique publique : production de données d'enquête et de données administratives, micro-simulation, évaluation *ex-ante* et *ex-post*...
- ▶ pierre.pora [at] ensae.fr

Introduction à l'économétrie

- ▶ Quartz : At the start of your career, what do you wish you had known ? What, if anything, do you wish you had not believed ?
- ▶ Claudia Goldin : Wish I had known : more econometrics. Wish I had not believed : that I needed to know more econometrics.

Introduction à l'économétrie

- ▶ Utilisation de l'outil statistique pour quantifier les faits économiques dans leur dimension empirique
- ▶ Façon de mettre à l'épreuve les théories économiques
- ▶ La panoplie d'outils potentiels est très vaste !

Introduction à l'économétrie

- ▶ L'usage de l'outil statistique soulève (au moins !) trois questions
 - ▶ Comment collecter des données quantitatives
 - ▶ Comment faire des comparaisons statistiques
 - ▶ Comment faire le lien entre une comparaison statistique et une théorie économique

Introduction à l'économétrie

- ▶ L'usage de l'outil statistique soulève (au moins !) trois questions
 - ▶ Comment collecter des données quantitatives
 - ▶ Domaine de la statistique d'enquête
 - ▶ Comment faire des comparaisons statistiques
 - ▶ Ce sur quoi on va se concentrer ce semestre
 - ▶ (et sur une façon particulière de procéder)
 - ▶ Comment faire le lien entre une comparaison statistique et une théorie économique
 - ▶ Questions d'inférence causale, économétrie structurelle

Introduction à l'économétrie

- ▶ Agnostique sur les raisons théoriques qui poussent à s'intéresser à telle ou telle comparaison
- ▶ Cet enseignement ne parle pas de cause et d'effets !
 - ▶ Ce sont des concepts qui font toujours appel à une (proto-)théorie économique
 - ▶ Ce sont des questions importantes d'économétrie
 - ▶ Mais elles seront abordées ailleurs et plus tard
 - ▶ En pratique c'est souvent ce sur quoi les économistes se concentrent
 - ▶ Mais pédagogiquement préférable de traiter cela séparément

Introduction à l'économétrie

- ▶ *In fine* cet enseignement se concentre sur
 - ▶ l'usage de la régression linéaire
 - ▶ indépendamment de son interprétation causale
 - ▶ en l'interprétant comme un outil permettant de générer des indicateurs synthétiques de comparaisons de moyennes
 - ▶ et en pratiquant son implémentation sur R

L'économétrie comme langage(s)

- ▶ Faire de l'économie c'est écrire à des économistes, dans le jargon des économistes
 - ▶ idéalement à propos de quelque chose qui intéresse d'autres économistes
- ▶ L'économétrie c'est simplement le(s) jargon(s) dans le(s)quel(s) les économistes parlent des faits empiriques
- ▶ 4 jargons au total :
 - ▶ Celui dans lequel on s'adresse aux ordinateurs
 - ▶ Un langage graphique qui consiste à exprimer ses résultats sous forme de figures et de tableaux
 - ▶ Un langage mathématique qui formalise le sens statistique de ce que l'on fait
 - ▶ Et une couche de jargon académique / professionnel pour recouvrir tout ça

L'économétrie comme langage(s)

- ▶ Cet enseignement est principalement un cours de langues !
- ▶ J'essaie de trouver des exemples et des illustrations pas *trop* inintéressants
- ▶ Mais ce qu'il faut retenir est relativement indépendant de ces exemples
- ▶ J'essaie simplement de vous apprendre un petit morceau de la langue dans laquelle vous serez amené-e-s à communiquer avec vos pairs

En pratique

- ▶ 12 semaines
- ▶ Une séance de TD (présence obligatoire) + une séance de cours chaque semaine
 - ▶ Sauf la dernière séance de cours = examen sur ordinateur
- ▶ Usage du logiciel R, indissociable de l'enseignement
 - ▶ On ne peut pas faire d'économétrie sans un ordinateur et sans un logiciel statistique !
 - ▶ Apprentissage par la pratique et par l'exemple

En pratique

- ▶ Articulation TD / cours
 - ▶ Au départ relativement déconnectées
 - ▶ TD très orientés sur l'acquisition du bagage minimal en R
 - ▶ Cours plus sur des questions de statistique et d'économétrie
 - ▶ Objectif : raccrocher les wagons au bout de quelques semaines
 - ▶ Utiliser R pour mettre en oeuvre les techniques étudiées en cours
 - ▶ Ca n'empêchera pas non plus les séances de cours de discuter parfois de l'implémentation en R

En pratique

- ▶ Pas mal de temps pour un contenu pas si vaste que ça
 - ▶ On peut en profiter pour entrer dans certains détails
 - ▶ Aussi pour échanger sur vos questions
 - ▶ N'hésitez pas à intervenir !

En pratique

- ▶ Validation : examen sur ordinateur la dernière séance
 - ▶ L'objet n'est pas de vous piéger
 - ▶ *A priori* :
 - ▶ Un peu de manipulations de R
 - ▶ Lecture de tableaux / graphiques tirés d'articles empiriques
 - ▶ Quelques questions théoriques
- ▶ Portera sur l'intégralité du programme
 - ▶ Pour l'instant : pas très intéressant de trop se demander ce qu'il y aura dedans

En pratique

- ▶ Supports en ligne sur mon site (<https://pierrepora.github.io/>)
 - ▶ Slides des séances de cours
 - ▶ Pas de notes rédigées à ce stade
 - ▶ Enoncés des TD + éléments de correction
 - ▶ Dans les dernières semaines : un exemple de choses que je pourrais mettre à l'examen

Quel usage de l'IA générative ?

- ▶ Bon adjuvant pour coder mais pas toujours exact ni efficace
 - ▶ Recourir à un modèle d'IA ne permet pas de se passer de compétences de bases dans le maniement de R
 - ▶ Être capable de tester ce que propose le modèle requiert une compréhension non-nulle du langage
- ▶ Sur la dimension plus théorique de la statistique et de l'économétrie : les modèles disponibles font le même genre d'erreurs que ce que l'on trouve dans un certain nombre de ressources en ligne
- ▶ Performances qui peuvent être amenées à évoluer dans le temps
- ▶ Mon conseil : se limiter dans l'usage de ces outils dans la phase d'apprentissage
 - ▶ Il faut faire un peu de muscle avant de pouvoir les utiliser efficacement

Un exemple pour motiver le reste de l'enseignement

- ▶ Goldin, Claudia. 2014. "A Grand Gender Convergence : Its Last Chapter." *American Economic Review* 104 (4) : 1091–1119.
- ▶ Papier consacré aux évolutions historiques des inégalités de genre sur le marché du travail aux Etats-Unis et aux mécanismes qui expliquent leur persistance
- ▶ Travail essentiellement empirique : illustre une des façons possibles de mobiliser un argument économétriques
- ▶ A la fin du semestre tout ce qui y est exposé dans la partie empirique devrait être clair pour vous !

Un exemple pour motiver le reste de l'enseignement

- ▶ Les ingrédients pour un travail convaincant d'économétrie appliquée :
 - ▶ Définir un indicateur clair
 - ▶ Faire attention aux difficultés posées par les comparaisons retenues
 - ▶ Utiliser des données de bonne qualité

Un exemple pour motiver le reste de l'enseignement

The ratio of (mean) annual earnings between male and female workers (full-time, full-year, 25 to 69 years) was 0.72 in 2010 and that of the medians was 0.77. The ratio of the medians for the same group was 0.74 in 2000 but 0.56 in 1980.¹⁵ These aggregate ratios have been somewhat sticky for the last ten years or so after greatly increasing in the preceding decades, especially in the 1980s. The same is true looking only at college graduates, for whom the ratios are lower—0.65 in 2010 for the means and 0.72 for the medians, about the same as it was in 2000. Interestingly, across the past decade the gender pay gap has narrowed *within* almost all age groups even though the aggregate has not budged as much. How can that be?

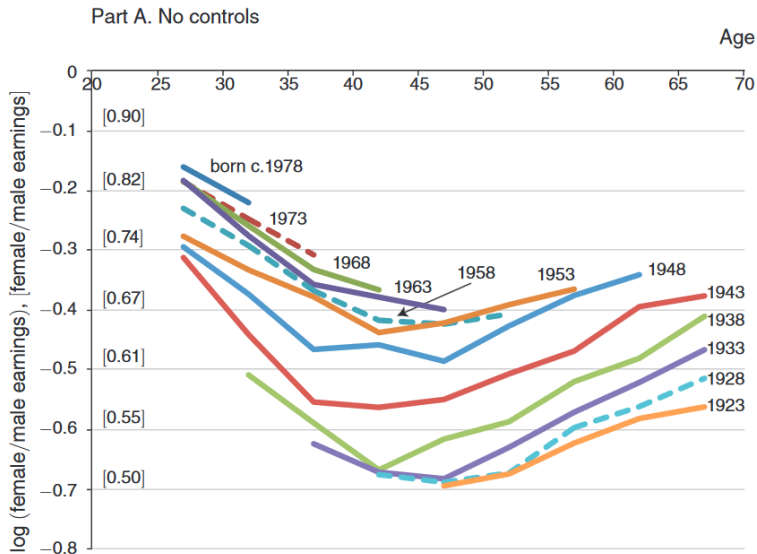
The answer concerns what happens to the gender gap over the life-cycle. The ratio of female to male earnings greatly decreases for some time as cohorts age. It is lower for individuals in their forties compared with the same individuals in their twenties. And because the baby boom is still working its way through the population, the aggregate ratio can be fairly stable even though the underlying components are actually increasing.

¹⁵ Full-time, full-year means 35 hours or more per week and 40 weeks or more per year, throughout this piece. The sample excludes earnings outliers (see notes to Figure 1) and members of the armed forces. The 2010 ratios are an average for 2009 to 2011 from the American Community Survey micro-data. Those for 1980 and 2000 are computed from the U.S. Census micro-data.

Un exemple pour motiver le reste de l'enseignement

- ▶ Comprenez-vous l'argument développé par Goldin ici ?
- ▶ Quelle serait une bonne façon de le mettre en évidence ?

Un exemple pour motiver le reste de l'enseignement



Un exemple pour motiver le reste de l'enseignement

Residual differences by occupation in earnings by gender, I will demonstrate, are largely due to the value placed on the hours and job continuity of workers, including the self-employed.²⁴ Individuals in some occupations work 70 hours a week and receive far more than twice the earnings of those who work 35 hours a week. But in other occupations they do not. Some occupations exhibit linearity with respect to time worked whereas others exhibit nonlinearity (convexity).²⁵ When earnings are linear with respect to time worked the gender gap is low; when there is nonlinearity the gender gap is higher.

Un exemple pour motiver le reste de l'enseignement

In many workplaces employees meet with clients and accumulate knowledge about them. If an employee is unavailable and communicating the information to another employee is costly, the value of the individual to the firm will decline. Equivalently, employees often gain from interacting with each other in meetings or through random exchanges. If an employee is not around that individual will be excluded from the information conveyed during these interactions and has lower value unless the information can be fully transferred in a low cost manner.

The point is quite simple. Whenever an employee does not have a perfect substitute nonlinearities can arise.²⁶ When there are perfect substitutes for particular workers and zero transactions costs, there is never a premium in earnings with respect to the number or the timing of hours. If there were perfect substitutes earnings would be linear with respect to hours. But if there are transactions costs that render workers imperfect substitutes for each other, there will be penalties from low hours depending on the value to the firm. A sparse framework will demonstrate these points and develop them further.

Un exemple pour motiver le reste de l'enseignement

- ▶ Comment mettre cette théorie à l'épreuve des faits ?
- ▶ Quels seraient les ingrédients nécessaires pour cela ?

Un exemple pour motiver le reste de l'enseignement

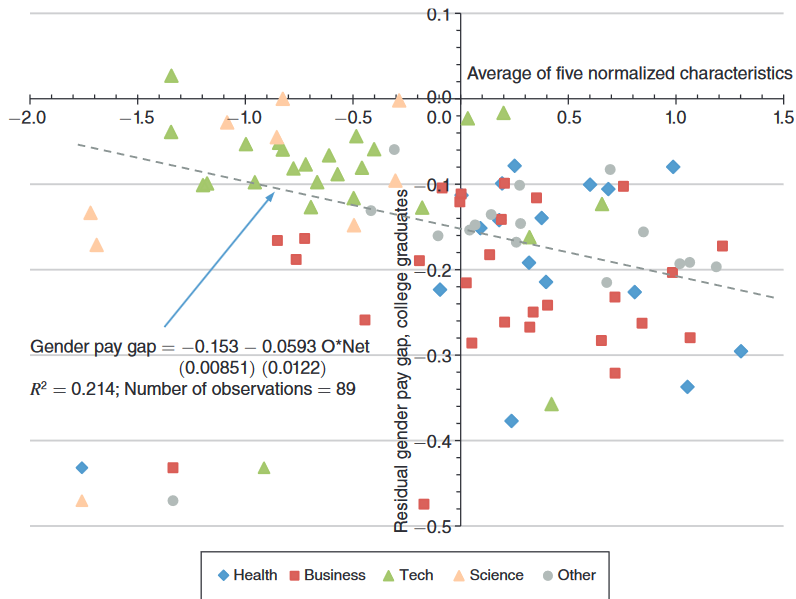
TABLE 2—O*NET CHARACTERISTICS: MEANS (NORMALIZED) BY OCCUPATIONAL GROUP

O*Net characteristics	Technology and science	Business	Health	Law
1. Time pressure	−0.488	0.255	0.107	1.51
2. Contact with others	−0.844	0.171	0.671	0.483
3. Establishing and maintaining interpersonal relationships	−0.611	0.548	0.276	0.781
4. Structured vs. unstructured work	−0.517	0.313	0.394	1.22
5. Freedom to make decisions	−0.463	−0.00533	0.974	0.764
Number of occupations	31	28	16	1

Notes: The occupations are those in Figure 2, part A. When there is more than one O*Net occupation for an ACS occupation, the characteristic is weighted by the number of workers in each of the O*Net occupations. Each of the O*Net characteristics has been normalized to have a mean of 0 and a standard deviation of 1. The work setting characteristics and questions most relevant to the issues raised here are:

1. Time pressure: *How often does this job require the worker to meet strict deadlines?*
Lower pressure means worker does not have to be around at particular times.
2. Contact with others: *How much does this job require the worker to be in contact with others (face-to-face, by telephone, or otherwise) in order to perform it?*
Less contact means greater flexibility.
3. Establishing and maintaining interpersonal relationships: *Developing constructive and cooperative working relationships with others, and maintaining them over time.*
The more working relationships, the more workers and clients the employee must be around.
4. Structured versus unstructured work: *To what extent is this job structured for the worker, rather than allowing the worker to determine tasks, priorities, and goals?* If the job is highly structured to the worker, there would be a lower chance that the worker would have close substitutes.
5. Freedom to make decisions: *How much decision making freedom, without supervision, does the job offer.*
Generally means that the worker determines what each client should receive, rather than being given a specific project, and thus workers are poorer substitutes for each other the greater are these freedoms.

Un exemple pour motiver le reste de l'enseignement



Un autre exemple

- ▶ Acemoglu, Daron, Simon Johnson, et James A. Robinson. 2001. "The Colonial Origins of Comparative Development : An Empirical Investigation." *American Economic Review* 91 (5) : 1369–1401.

Un autre exemple

1. There were different types of colonization policies which created different sets of institutions. At one extreme, European powers set up “extractive states,” exemplified by the Belgian colonization of the Congo. These institutions did not introduce much protection for private property, nor did they provide checks and balances against government expropriation. In fact, the main purpose of the extractive state was to transfer as much of the resources of the colony to the colonizer.

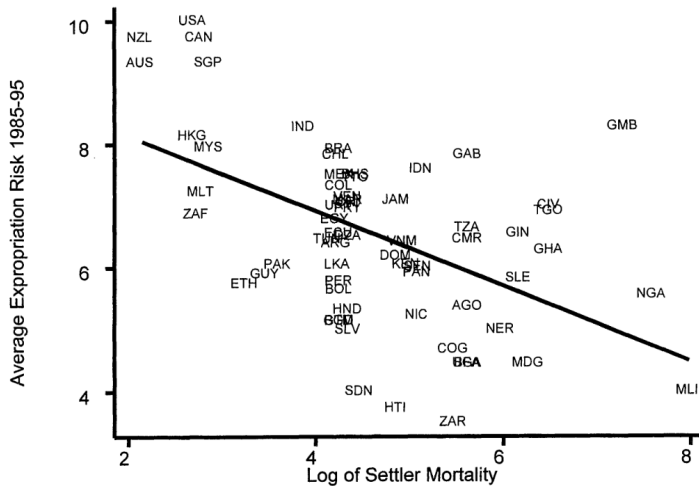
At the other extreme, many Europeans migrated and settled in a number of colonies, creating what the historian Alfred Crosby (1986) calls “Neo-Europes.” The settlers tried to replicate European institutions, with strong emphasis on private property and checks against government power. Primary examples of this include Australia, New Zealand, Canada, and the United States.

2. The colonization strategy was influenced by the feasibility of settlements. In places where the disease environment was not favorable to European settlement, the cards were stacked against the creation of Neo-Europes, and the formation of the extractive state was more likely.
3. The colonial state and institutions persisted even after independence.

Un autre exemple

- ▶ Comment tester empiriquement cette idée ?
- ▶ Quels sont les ingrédients nécessaires pour le faire ?

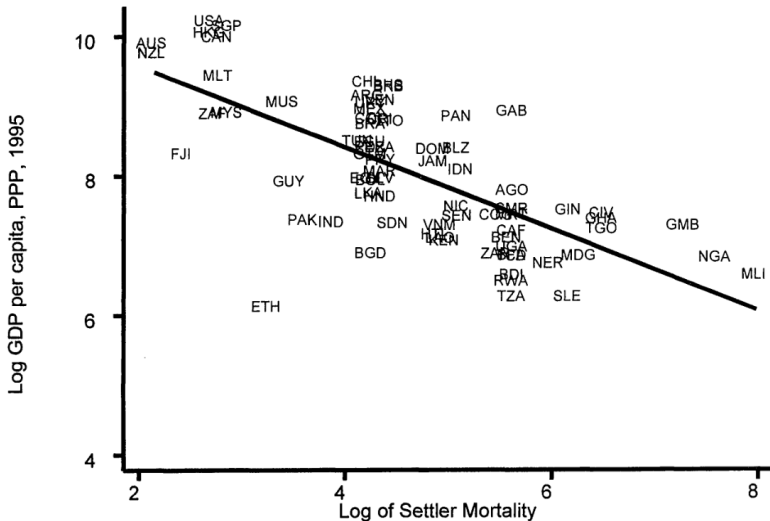
Un autre exemple



Un autre exemple

- ▶ Comment utiliser cette idée pour étudier l'effet des institutions coloniales sur le développement économique ?

Un autre exemple



La suite

- ▶ L'objet de cet enseignement : ces droites qui traversent des nuages de point
 - ▶ Comment les interpréter ?
 - ▶ Indépendamment de leur interprétation causale !
 - ▶ Comment les construire ?
 - ▶ Théorie de leur construction (sans aller dans un formalisme trop élaboré)
 - ▶ Comment faire en pratique à partir de données quantitatives empiriques en utilisant le logiciel R

La semaine prochaine

- ▶ Attention cours et TD inversés (cours le jeudi TD le vendredi)
- ▶ Cours : voir de façon un peu générale ce que signifie travailler avec des données quantitatives en économie
- ▶ TD : continuer sur les bases de R