L'accès aux données sur les entreprises du point de vue des sociologues

Emmanuel Lazega, Lise Mounier

▶ **To cite this version:**

Emmanuel Lazega, Lise Mounier. L'accès aux données sur les entreprises du point de vue des sociologues. 1999. hal-01071205

HAL Id: hal-01071205

https://sciencespo.hal.science/hal-01071205

Submitted on 6 Oct 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

3. L'accès aux données sur les entreprises du point de vue des sociologues

Emmanuel Lazega, Lise Mounier (Lasmas-IdL)

Pour les économistes, la France dispose de bases d'enquêtes sur les entreprises parmi les meilleures du monde, en termes de couverture (peu d'omissions, peu de répétitions), de contenu et de codification (informations enregistrées), bien que les relations entre statistiques et comptabilité d'entreprise y soient aussi notoirement faibles que dans le reste du monde. L'organisation de la recherche et la production de connaissances sur les entreprises en France est dominée par l'Insee dont sont issus (et qui rassemble) les plus gros fichiers statistiques. D'autres ministères produisent leurs propres enquêtes, comme le ministère de l'Industrie qui dispose de son propre service statistique (le Sessi), ou même d'autres institutions (par exemple, la banque de France a son propre panel d'entreprises). Dans le respect des règles de confidentialité et du cadre réglementaire qui s'impose à tous, ces données ne sont pas facilement mises à disposition ; et lorsqu'elles le sont, c'est sous forme anonymisée, sous convention de recherche – le nombre de conventions étant limité par la volonté de suivre de près les travaux des chercheurs extérieurs. Mais les fichiers principaux, dont *l'Enquête Annuelle d'Entreprises (*EAE*)* et *SIRENE* (le recensement des entreprises), la *Base d'Analyse Longitudinale (*BAL*)*, *Liaisons financières (*LIFI*)*, *SUSE* (données fiscales et données de l'EAE) sont à l'Insee (ou rassemblées par l'Insee). De nombreuses informations sur les organisations participant au système productif français sont dispersées dans des institutions ad hoc, comme par exemple l'Institut national de la propriété intellectuelle (qui vend très cher l'information sur les brevets déposés par les entreprises et organisations de recherche publique). Les universités, contrairement à ce qui se passe aux États-Unis, n'ont qu'un rôle négligeable dans la production de données statistiques sur les entreprises. À ces données publiques, il faut ajouter les fichiers de données privés comme celles de Kompass[3] (dont sont facilement extraites, par exemple, les informations sur les relations interlock entre conseils d'administration).

L'Insee dispose de données administratives, en complément des données d'enquête, qu'il n'a pas le droit de mettre à disposition des chercheurs universitaires. En général, l'accès aux données d'enquête sur les entreprises est difficile – en tout cas pas immédiat – parce que l'anonymat est souvent difficile à conserver[4] ; les fichiers peuvent être mis en concordance et procurer à certains utilisateurs commerciaux des avantages concurrentiels sur d'autres. Pour ces raisons, les fichiers Insee et Sessi ne sont le plus souvent accessibles que par le dépôt d'un projet de recherche et d'une demande d'accès aux données individuelles auprès de la Cnil (à moins de collaborer personnellement avec un chercheur membre de ces institutions). Lorsque le chercheur et universitaire obtient du Comité du secret statistique l'autorisation d'accéder aux données individuelles qui l'intéressent, il paie un droit de mise à disposition. Il ne peut ni les communiquer à des tiers, ni en faire état dans ses relations avec les entreprises ; il s'engage aussi à respecter les règles du secret statistique dans les publications qui seront faites de ses travaux[5] ; enfin il s'engage à détruire les données individuelles une fois l'étude achevée. Un représentant du patronat et un représentant de la Direction des Statistiques d'Entreprises siègent au Comité du secret statistique chargé d'instruire les demandes d'utilisation de bases de données par des universitaires et autres. Il faut noter aussi que la difficulté d'accès n'est pas seulement liée à l'anonymat. L'Insee a maintenant des chercheurs de très haut niveau qui travaillent sur des enquêtes originales et qui ont tendance à ne céder le droit d'accès qu'une fois les fichiers exploités. On pense par exemple aux données sur la concentration verticale et les enquêtes sur les têtes de réseaux d'entreprises dans l'habillement et dans le bricolage menées par la Division Commerce de l'Insee. Notre expérience du Sessi est qu'il

[3]. Des équipes peu nombreuses ont construit et mettent à jour leurs propres fichiers (par exemple celle de François Morin à Toulouse, qui redescend jusqu'aux rapports d'activités des grands groupes français ou européens).

[4]. Il n'est pas difficile de deviner quelle est l'entreprise qui fabrique des pneus à Clermont-Ferrand.

[5]. Par exemple, aucun résultat relatif à un groupe de moins de trois entreprises ou à un groupe d'entreprises dont une seule représente plus de 85 % du résultat total ne peut être diffusé.

délivre gratuitement aux universitaires des extraits de fichiers (moyennant un rapport et la rédaction d'un quatre pages à la fin du délai accordé pour la recherche).

En tant qu'universitaires, nous n'avons pas l'expérience d'une quelconque participation ou consultation en matière de conception d'enquêtes nouvelles. Or ce qui intéresse vraiment la sociologie économique ou la sociologie quantitative des organisations dans l'analyse secondaire des bases de données sur les entreprises en France, en particulier sur les relations inter-entreprises, c'est l'identification de diverses formes de discipline sociale sous-jacentes aux échanges de ressources ayant cours dans l'appareil de production. Les sociologues partent du principe que les entreprises opèrent dans des régimes d'interdépendances de ressources (pour la production) et de marchés contraints (pour les échanges). Avec de bonnes données, pour l'instant très difficiles à exiger des entreprises (ou que l'Insee n'a pas le droit de rétrocéder, comme dans le cas de données dites administratives, qui ne sont pas des données d'enquête), la contribution de cette discipline sociale à la productivité et à la performance de ces entreprises (ou grands groupes d'entreprises) devrait être mesurable. Des enquêtes à grande échelle allant dans ce sens devraient être possibles à mener. En France, pour l'instant, l'Insee ou le Sessi (ou un centre de recherche équivalent) sont apparemment seuls à avoir l'envergure nécessaire pour entreprendre dans ce sens. Ils restent cependant tributaires de la coopération des entreprises pour lesquelles la charge statistique est toujours lourde. Il n'en reste pas moins qu'une telle connaissance est de nature à éclairer, non seulement les chercheurs, mais aussi les institutions sociales et les décideurs publics et privés.

Le manque d'informations quantitatives intéressant le sociologue de l'économie n'est cependant pas seulement dû aux difficultés de rapprochement avec les grandes institutions d'économistes. Il n'est pas nécessaire de faire un sondage spécifique pour se rendre compte que les enquêtes sur les entreprises sont peu utilisées par les sociologues universitaires ou CNRS. Deux raisons au moins contribuent à ce fait. Premièrement, il n'existe pas en France de tradition quantitative forte en sociologie économique procédant par une approche organisationnelle ; or les fichiers entreprises sont souvent complexes. Deuxièmement, ces enquêtes sont pensées et conduites par des économistes soumis à des contraintes administratives et politiques (alléger la charge statistique pesant sur les entreprises en réduisant la taille des questionnaires envoyés par l'Insee). Par exemple, même lorsque les thématiques se rapprochent des préoccupations des sociologues, comme c'est nettement le cas pour l'enquête *Liaisons industrielles* (LI) du SESSI, les données recueillies adoptent comme unité d'analyse les relations contractuelles de l'entreprise focale avec ses sous-traitants « en général » ; ce niveau d'agrégation interdit au sociologue de descendre au niveau d'analyse qui l'intéresse véritablement, c'est-à-dire les relations (contractuelles, sociales) spécifiques entre deux unités bien spécifiées. Autrement dit, l'unité d'analyse reste l'entreprise focale, mais L.I. ne connaît pas les entités avec lesquelles la première est en relation ; ses relations avec les autres entreprises ne sont pas décrites de manière désagrégeable. L'analyse multiniveaux au sens de la statistique des réseaux sociaux n'est donc pas possible sur ce fichier. Il n'est pas possible d'interroger les données pour établir l'existence ou l'absence d'arrangements contractuels d'un type spécifique s'accompagnant d'une discipline sociale, elle-même spécifique, qui offre des garanties pour la prise de risque ou la mise en œuvre des règles contractuelles. Ces fichiers ne recueillent donc pas toujours des informations immédiatement et directement utilisables par des sociologues.

Il n'en reste pas moins que les travaux de recherches menés à l'Insee sur la base de ces données se rapprochent des préoccupations des sociologues de l'économie qui utilisent une entrée intra- et inter-organisationnelle. On pense ici à nouveau essentiellement aux travaux sur les groupes et les relations intra-groupes parus sur la base de l'enquête LIFI et l'accent qu'elle met sur le contrôle capitalistique par une holding tête de groupe. Ou encore aux travaux sur les têtes de réseaux dans l'habillement ou l'enquête sur la commercialisation d'articles de bricolage par la Division Commerce de l'Insee. Il va de soi que d'autres recherches sont menées à l'Insee, notre objectif n'est pas d'en faire un recensement.

Données européennes

À notre connaissance, l'Insee ne communique pas à Eurostat ses bases de données entreprises. Outre les problèmes de confidentialité, la statistique d'entreprise posant des problèmes conceptuels et méthodologiques spécifiques de définition d'unité, la manière dont les pays européens gèrent les répertoires d'entreprises, les sources de statistique sur la démographie des entreprises (créations et cessation), leur maintenance, les usages auxquels ils sont associés varient considérablement. À l'échelle européenne, il y a peu d'homogénéité dans les bases statistiques (bien que sous des appellations différentes peuvent se dissimuler des entités semblables) et dans les politiques de recueil, bien que l'harmonisation comptable soit à l'ordre du jour. Historiquement, chaque pays a son organisation industrielle, sa réglementation sociale et fiscale, qui ont fondé les fichiers administratifs dont ils sont issus. Ce manque d'harmonisation se traduit par une certaine pauvreté des données d'Eurostat et des programmes de comparaisons internationales. Pour l'instant, à notre connaissance, les données Eurostat sont des données économiques agrégées qui ne peuvent pas servir à l'usage intensif auquel les destineraient les sociologues de l'économie.