



*Webinaire, AIREPME, 19 juin 2020*

# Étudier la performance entrepreneuriale et... en éviter les écueils

**Marcus Dejardin**

Université de Namur & UCLouvain

marcus.dejardin@unamur.be OU marcus.dejardin@uclouvain.be

# Étudier la performance entrepreneuriale et... en éviter les écueils

Le contenu de ce webinaire est basé pour l'essentiel sur:

SIEPEL, J., DEJARDIN, M. (2020), "How Do We Measure Firm Performance? A Review of Issues Facing Entrepreneurship Researchers", in G. Saridakis and M. Cowling, eds., *The Handbook of Quantitative Research Methods in Entrepreneurship*, Cheltenham: Edward Elgar, à paraître.

La version des auteurs est téléchargeable à l'adresse:

<https://ideas.repec.org/e/pde36.html>

RePEc: Research Papers in Economics

NEP: New Economics Papers

nep-ent: NEP report on Entrepreneurship

<http://nep.repec.org/nep-ent.html>

# Objectifs généraux de ce webinar

- Eveiller à quelques questions fondamentales relatives à la recherche empirique quantitative
- Penser la recherche quantitative en entrepreneuriat
- Se donner les moyens d'être ambitieux en recherche

# Plan pour ce webinar

Intention: fournir un aperçu succinct des défis importants auxquels sont confrontés les chercheurs en entrepreneuriat menant des recherches quantitatives sur la performance de la firme, ainsi qu'un aperçu des différentes sources de données qui peuvent être utilisées et de certaines techniques qui peuvent être utilisées pour garantir la fiabilité des données.

- La performance de la firme: un phénomène multifacette.
- La mesure de la performance.
- Les types de données disponibles.
- Les problèmes potentiels en lien avec les données.
- Quelques enjeux supplémentaires.

# La performance de la firme: un phénomène multifacette

- L'entrepreneuriat comme domaine de recherche.
- Quelle unité d'observation?
  - Individu, équipe, firme, région géographique,..
  - >> Différents niveaux d'analyse, différentes questions, différentes données.
  - >> Exemples: impact du profil psychologique du créateur sur la performance en termes d'innovation; survie de l'entreprise.
- La suite de la communication est développé en référence à la firme.
- La firme ? L'agrégat correspondant à la structure organisationnelle productive plutôt que la fonction de production (Mäntysaari, 2012), avec la difficulté que les limites de l'organisation deviennent de plus en plus floues (Zingales, 2000).

# La performance de la firme: un phénomène multifacette

- Ce dont nous n'avons pas l'intention de traiter ici:
  - Le niveau de l'individu (même si indépendant)
    - Revenus et bien-être individuels écartés
  - Les enjeux liés à l'analyse des données (biais de régression)
- Point focal: s'assurer que les données utilisées sont appropriées pour traiter la question identifiée
  - Importance du design de recherche et du choix de données

# La mesure de la performance

- La recherche quantitative au niveau de la firme s'intéresse généralement aux facteurs, toutes choses égales par ailleurs, qui font que les entreprises ont une performance différente - meilleure ou pire que les autres.
- Deux points préalables avant d'aborder quelques mesures:
  - Les mesures que nous utilisons pour observer les performances peuvent différer de celles qui sont importantes pour d'autres chercheurs et pour les entrepreneurs. Les entreprises peuvent apparaître très différentes selon la mesure privilégiée.
  - La performance de la firme a une forte composante temporelle.
    - Les différentes mesures usuelles de la performance sont corrélées, mais la corrélation n'est pas forte (Chandler et al., 2009). Une entreprise rentable peut ne pas avoir un grand nombre d'emplois. Une entreprise très innovante peut ne pas avoir un niveau élevé de ventes.
    - La croissance ne s'observe pas simultanément à travers les différentes mesures mais a tendance à se produire dans une séquence (Coad et al. 2017).
    - Période d'observation vs « instantané ».

# La mesure de la performance

- L'emploi
  - Mesure préférée des économistes et des responsables politiques.
  - La littérature sur le séquençement de la croissance fournit des résultats mitigés sur la croissance de l'emploi. En général, la croissance de l'emploi stimule la croissance des ventes et des bénéfices (Coad et Rao 2009; Coad 2010; Coad et al. 2014). Cependant, dans les entreprises à forte croissance, la croissance de l'emploi est souvent à la traîne et apparaît comme l'une des dernières étapes du processus de croissance (Coad et al., 2017).
- Le chiffre d'affaires
  - Avec ses dérivés, mesure plus usuelle en management.
  - La littérature sur le séquençement de la croissance (voir ci-dessus) suggère que pour la plupart des entreprises, la croissance du chiffre d'affaires suit la croissance de l'emploi, mais pour les entreprises à forte croissance, la croissance du chiffre d'affaires peut précéder la croissance de l'emploi.

# La mesure de la performance

- Le profit / la profitabilité
  - Souvent associé à la reconnaissance de l'opportunité entrepreneuriale.
  - Comptes de l'entreprise; ratios financiers (e.g. ROA, ROE).
- La productivité
  - Mesurée au niveau de l'entreprise, la productivité est un indicateur important de l'(in)efficacité de l'entreprise dans son utilisation des facteurs de production.
    - Productivité du travail; par exemple, la valeur ajoutée par travailleur ou par heure de travail.
    - Moins fréquemment, productivité du capital, comme la valeur ajoutée par unité de stock de capital fixe.
    - Productivité totale des facteurs (TFP) (Gal, 2013).

# La mesure de la performance

- La R&D, les brevets, l'innovation
  - Il existe des liens entre la littérature sur l'entrepreneuriat et l'innovation, et par conséquent, un grand nombre d'études se concentrent sur les activités innovantes des entreprises.
  - L'innovation est un sujet difficile à mesurer. Les études sur l'innovation sont fréquemment limitées aux intrants, tels que la recherche et développement (R&D), le personnel ou les ressources financières engagées; les résultats intermédiaires, tels que les brevets; et le chiffre d'affaires généré par les nouveaux produits (Hopkins et Siepel 2013).
  - Mesures: R&D, brevets, données d'enquête (CIS, par exemple).
- La survie de l'entreprise
  - La durée pendant laquelle une entreprise est active sur un marché concurrentiel peut être interprétée comme une mesure de sa capacité à rivaliser.
  - Durée? Difficultés liées à l'identification de l'entrée (*entry*) et de la sortie (*exit*).
- La mesure de la performance de l'entreprise sociale? De la coopérative d'intérêt général (vs d'intérêt mutuel)?

# Les types de données disponibles

- Une variété énorme dans les bases de données.
- Un bref état des lieux.
  - Données auto-déclarées (*self-report data*)
    - Il s'agit de données collectées par le biais d'un entretien ou d'un questionnaire dans lequel le répondant donne ses réponses à une série de questions.
    - Exemples emblématiques: Global Entrepreneurship Monitor; US Panel Study of Entrepreneurial Dynamics (PSED) I & II.
    - Les données d'auto-évaluation restent une source majeure d'informations en ouvrant très largement la possibilité aux chercheurs de collecter des données originales sur les caractéristiques personnelles ou organisationnelles qui sont à peine, voire pas du tout, disponibles dans les statistiques officielles.
    - Malheureusement, souvent disponibles en coupe transversale, uniquement. Une piste de solution? Combiner les données d'auto-évaluation avec des données provenant d'autres sources de données, telles que sources officielles ou commerciales. Cela nécessite une clé permettant de faire le lien entre les ensembles de données (numéro d'identification de l'entreprise, par exemple).

# Les types de données disponibles

- Un bref état des lieux (suite).
  - Les statistiques officielles
    - Une grande quantité d'informations, collectée par les pouvoirs publics.
    - Les études au niveau de l'entreprise peuvent généralement être classées en trois catégories: enquêtes officielles, obtenues par déclaration obligatoire, et données administratives.
      - Enquêtes officielles: souvent réalisées par les offices nationaux ou régionaux de statistique, avec attention pour la conception d'enquête, l'échantillonnage, la collecte de données, etc.
        - Exemples: CIS, SINE.
        - Largement acceptées comme sources valables.
        - Accessibles au public, bien que les modalités d'accès et les restrictions de sécurité varient selon les pays.
        - Il peut toutefois être difficile de faire correspondre la question de recherche aux données disponibles: par exemple, les chercheurs utilisant les études CIS qui souhaitent étudier en détail la propriété intellectuelle n'y trouveront qu'un ensemble ordinaire de variables binaires.

# Les types de données disponibles

- Un bref état des lieux (suite).
  - Les statistiques officielles
    - Obtenues par déclaration obligatoire.
      - Exemple: Comptes annuels.
    - Données administratives.
      - Exemple: données TVA.
      - Un avantage significatif de ces données est que le risque d'erreur de mesure est généralement faible, car dans de nombreux cas, les répondants pourraient être pénalement responsables des données incorrectes fournies.
      - Les variables disponibles sont limitées à celles collectées par les PP, de sorte que d'autres mesures, que le chercheur pourrait souhaiter, pourraient très bien ne pas être disponibles.
  - Les données commerciales
    - Bien qu'elles soient souvent tirées de données officielles, les bases de données commerciales peuvent être relativement faciles d'accès (au moins pour les chercheurs d'institutions abonnées).
      - Exemple: données du Bureau van Dijk (BvD); Barclays Bank data set (Frankish et al., 2013); « scraped web data » et *APIs*.
      - Service à la clientèle; problème des données manquantes.

# Les types de données disponibles

- Un bref état des lieux (suite).
  - Combinaisons de données; données employeur-employé; base de données à plusieurs niveaux; 'Big Data'
    - La recherche peut bénéficier énormément de la combinaison de différents ensembles de données, ce qui implique l'utilisation de variables clés pour relier les ensembles de données.
    - Il peut arriver que les questions de recherche soulevées rendent cette combinaison tout simplement nécessaire car il n'existe aucun ensemble de données pouvant couvrir toutes les variables impliquées dans l'analyse.
      - Exemple: en utilisant des données employeur-employé pour les entreprises de Nouvelle-Zélande, Kirker et Sanderson (2018) ont pu tester la relation entre la croissance de la productivité et la qualité des nouveaux travailleurs, ainsi qu'avec les connaissances que les nouveaux travailleurs pouvaient avoir acquis chez leurs employeurs antérieurs.

# Les types de données disponibles

- Un bref état des lieux (suite).
  - Combinaisons de données; données employeur-employé; base de données à plusieurs niveaux; 'Big Data'
    - Exemple: vous vous attendez à ce que la dimension culturelle locale, qui diffère au sein d'un État donné, puisse expliquer les performances de l'entreprise. En négligeant l'existence de grappes (ou groupes) d'entreprises partageant le contexte local donné, qui se différencie des autres contextes, le chercheur négligerait le fait que l'hypothèse d'indépendance des résidus pourrait ne pas être respectée.
  - La collecte, le stockage et la fourniture de gigantesques bases de données ouvrent des opportunités insoupçonnées pour la recherche en économie, en gestion et en économétrie (voir Einav et Levin 2014; George et al., 2014; Varian 2014).
  - En ce qui concerne les études empiriques sur la performance des entreprises, 'Big Data ' représente une opportunité d'ouvrir très largement ce qui apparaît encore largement comme une boîte noire et de répondre à des questions de recherche nouvelles, ou négligées par manque de données.
  - Ce qui semble être de nouvelles et formidables voies de recherche ne permet cependant pas d'échapper aux questions d'ordres éthique et légal.

# Les problèmes potentiels en lien avec les données

- Erreur de mesure
  - Une multitude de problèmes peuvent conduire à une mesure incorrecte. Il peut s'agir d'écart entre les mesures utilisées et les «faits», ainsi que de problèmes de sélection et de biais.
  - Une littérature substantielle existe (cf. Bertrand et Mullainathan 2001; Bound et al. 2001), mais signalons ici quelques exemples saillants.
    - La construction de nouvelles mesures est d'une importance considérable pour la validation de nouveaux concepts. Des défis importants existent dans ce domaine (cf. Chandler et Lyon, 2001; Crook et al., 2010).
    - Sélection et biais potentiels.
      - Exemples: le biais de non-réponse (Rutherford et al., 2017); biais incombant aux répondants (Zhang et Cueto, 2017), liés à l'excès de confiance en soi ou au sur-optimisme entrepreneurial, par exemple.

# Les problèmes potentiels en lien avec les données

- Données manquantes

- Un problème fréquent dans de nombreuses recherches empiriques, en particulier dans les recherches sur la performance des entreprises.
  - Une question clé est de savoir si les données manquent au hasard. Si tel est le cas, ces données manquantes peuvent être ignorées car elles n'ont pas d'impact ou de biais substantiels sur les résultats.
  - En revanche, si des données sont manquantes mais que les données manquantes ne sont pas aléatoires, le biais peut être important. Heureusement, il existe un certain nombre de techniques statistiques qui peuvent être utilisées pour résoudre les problèmes de données manquantes (cf. e.g. Schafer et Graham, 2002; Allison, 2001; Honaker et King, 2010)
  - Le problème des données censurées et tronquées peut être envisagé parallèlement au problème des données manquantes. Particulièrement notables dans l'analyse de survie. Censure: la survie d'un échantillon d'entreprises est observée sur une période donnée; quid des entreprises survivantes à la fin de la période alors que toutes mourront un jour? Troncature: certaines des entreprises faisant partie d'une cohorte étaient déjà décédées au début de la période d'étude et ne peuvent pas faire partie de l'échantillon. Des techniques statistiques ont été développées pour résoudre ces problèmes (voir, par exemple, Klein et Moeschberger 2003).

# Les problèmes potentiels en lien avec les données

- Biais de la variance de méthode commune
  - Variance liée à la méthode de mesure plutôt qu'au phénomène étudié; peut prendre plusieurs formes (cf. Podsakoff et al., 2003); assez fréquent dans les données issues d'enquête transversale, auto-déclarées.
  - Il existe des solutions (cf. Conway et Lance, 2010). Cela étant, une bonne conception des instruments et du design de recherche sont des atouts majeurs pour éviter le problème.
- Données incorrectes et erreur du répondant
  - Fraser et al. (2007) ont utilisé des données d'enquête par panel pour explorer le biais de rappel parmi les répondants. Ils ont identifié un certain nombre de sources potentielles de biais. Parmi celles-ci, les entreprises qui ont subi des changements, comme la croissance, la contraction des activités, ou la délocalisation, sont plus susceptibles de faire preuve de biais. Cf. aussi Bloom et al. (2019).

# Les problèmes potentiels en lien avec les données

- Données incorrectes et erreur du répondant (suite)
  - En outre, de simples erreurs de saisie de données ou des instructions mal lues peuvent s'avérer problématiques.
    - Exemple: si un répondant à une enquête est invité à fournir le chiffre d'affaires annuel de son entreprise, en milliers, et est distrait, il pourrait introduire 100 000 CAD dans l'espace réservé, là où la réponse correcte aurait été 100 CAD.
    - Certains de ces problèmes peuvent être résolus lors du nettoyage des données.
- Erreur du chercheur
  - Responsabilité du chercheur d'utiliser les approches les plus appropriées et les plus robustes pour répondre à sa question de recherche. Il existe une grande variété d'erreurs qui peuvent survenir en raison d'une utilisation inappropriée des techniques économétriques ou de gestion des données.
  - Responsabilité éthique (cf. O'Boyle et al. 2017).

# En guise de conclusion...

## Quelques enjeux supplémentaires

- Les progrès en termes de puissance de calcul, de technique économétrique et d'analyse permettent aux chercheurs de répondre à des questions qu'il aurait été pratiquement impossible de traiter il y a même dix ans. Il y a sans aucun doute lieu de **se familiariser avec les techniques les plus avancées**, qui ouvrent d'énormes opportunités.
- Pour rappel: importance du **design de la recherche**. Les mesures doivent correspondre très étroitement aux théories et aux concepts explorés, mais cela reste un problème récurrent.
- Nous n'avons pas discuté des **données d'origine (quasi-)expérimentale**, ou des **régressions sur discontinuité**. Les recherches reposant sur ces développements sont probablement parmi les plus prometteuses, y compris en entrepreneuriat.
- *En conclusion, un alignement minutieux des mesures, des sources de données et des approches empiriques est nécessaire afin de générer des résultats convaincants qui permettront de développer des explications solides en référence à la performance des entreprises.*

Merci de votre attention!

Marcus Dejardin

marcus.dejardin@unamur.be

marcus.dejardin@uclouvain.be

## Références

- Allison, P.D. (2001). *Missing Data: Quantitative Applications in the Social Sciences*. Thousand Oaks, CA: SAGE Publications.
- Bertrand, M. and Mullainathan, S. (2001). Do people mean what they say? Implications for subjective survey data. *American Economic Review*, 91(2), 67–72.
- Bloom, N., Brynjolfsson, E., Foster, L., Jarmin, R., Patnaik, M., Saporta-Eksten, I. and Van Reenen, J. (2019). *American Economic Review*, 109(5), 1648–1683.
- Bound, J., Brown, C. and Mathiowetz, N. (2001). Measurement error in survey data. *Handbook of Econometrics*, 5, 3705–3843.
- Chandler, G.N. and Lyon, D.W. (2001). Issues of research design and construct measurement in entrepreneurship research: the past decade. *Entrepreneurship Theory and Practice*, 25(4), 101–113.
- Chandler, G.N., McKelvie, A. and Davidsson, P. (2009). Asset specificity and behavioral uncertainty as moderators of the sales growth: Employment growth relationship in emerging ventures. *Journal of Business Venturing*, 24(4), 373–387.
- Coad, A. (2010). Exploring the processes of firm growth: Evidence from a vector auto-regression. *Industrial and Corporate Change*, 19(6), 1677–1703.
- Coad, A. and Rao, R. (2009). Firm growth and R&D expenditure. *Economics of Innovation and New Technology*, 19(2), 127–145.
- Coad, A., Daunfeldt, S.O., Holz, W., Johansson, D. and Nightingale, P. (2014). High-growth firms: Introduction to the special issue. *Industrial and Corporate Change*, 23(1), 91–112.
- Coad, A., Siepel, J. and Cowling, M. (2017). Growth processes of high-growth firms as a four-dimensional chicken and egg. *Industrial and Corporate Change*, 26(4), 537–554.
- Conway, J.M. and Lance, C.E. (2010). What reviewers should expect from authors regarding common method bias in organizational research. *Journal of Business and Psychology*, 25(3), 325–334.
- Crook, T.R., Shook, C. Morris, M.L., and Madden, T. (2010). Are we there yet? An assessment of research design and construct measurement practices in entrepreneurship research. *Organizational Research Methods*, 13(1): 192–206.
- Einav, L. and Levin, J. (2014). Economics in the age of big data. *Science*, 346(6210), 715–721.
- Frankish, J.S., Roberts, R.G., Coad, A., Spears, T.C. and Storey, D.J. (2013). Do entrepreneurs really learn? Or do they just tell us that they do? *Industrial and Corporate Change*, 22(1), 73–106.
- Fraser, S., Green F.J. and Mole, K. (2007). Sources of bias in the recall of self-generated data: the role of anchoring. *British Journal of Management*, 18(2), 192–208.
- Gal, P. (2013). Measuring Total Factor Productivity at the Firm Level using OECD-ORBIS, OECD Economics Department Working Papers, No. 1049, OECD Publishing, Paris.
- George, G., Haas, M.R. and Pentland, A. (2014). Big data and management. *Academy of Management Journal*, 57(2), 321–332.
- Honaker, J. and King, G. (2010). What to do about missing values in time-series cross-section data. *American Journal of Political Science*, 54(2), 561–581.
- Hopkins, M.M. and Siepel, J. (2013). Just how difficult can it be counting up R&D funding for emerging technologies (and is tech mining with proxy measures going to be any better)? *Technology Analysis & Strategic Management*, 25(6), 655–685.
- Kirker, M. and Sanderson, L. (2018). Firm Productivity Growth and Its Relationship to the Knowledge of New Workers, 15 February, Mimeo.
- Klein, J.P. and Moeschberger, M.L. (2003). *Survival Analysis: Techniques for Censored and Truncated Data*, 2nd edition. Heidelberg: Springer-Verlag.
- Mäntysaari, P. (2012). *Organising the Firm: Theories of Commercial Law, Corporate Governance and Corporate Law*. Heidelberg: Springer.
- O’Boyle, E., Banks, G. and Gonzalez-Mule, E. (2017). The chrysalis effect: how ugly initial results metamorphosize into beautiful articles. *Journal of Management*, 43(2) 376–399.
- Podsakoff, P.M., MacKenzie, S.B., Lee, J.-Y. and Podsakoff, N.P. (2003). Common method biases in behavioral research: a critical review of the literature and recommended remedies. *Journal of Applied Psychology*, 88(5), 879–903.
- Rutherford, M.W., O’Boyle, E.H., Miao, C., Goering, D. and Coombs, J.E. (2017). Do response rates matter in entrepreneurship research? *Journal of Business Venturing Insights*, 8, 93–98.
- Schafer, J.L. and Graham, J.W. (2002). Missing data: our view of the state of the art. *Psychological Methods*, 7(2), 147–177.
- Siepel, J., Dejardin, M. (2020), “How Do We Measure Firm Performance? A Review of Issues Facing Entrepreneurship Researchers”, in G. Saridakis and M. Cowling, eds., *The Handbook of Quantitative Research Methods in Entrepreneurship*, Cheltenham: Edward Elgar, à paraître.
- Varian, H.R. (2014). Big data: new tricks for econometrics. *Journal of Economic Perspectives*, 28(2), 3–28.
- Zhang, S.X. and Cueto, J. (2017). The study of bias in entrepreneurship. *Entrepreneurship Theory and Practice*, 41(3), 419–454.
- Zingales, L. (2000). In search of new foundations. *The Journal of Finance*, 55, 1623–1653.