



DIRECTION GÉNÉRALE DE LA RECHERCHE ÉCONOMIQUE
ET DE L'ANALYSE DES POLITIQUES



Collection Documents de travail

Le capital de risque comme facteur de croissance

Ryan Kelly, Industrie Canada
Hankook Kim, Industrie Canada

Document de travail 2013-01



Autorisation de reproduction

À moins d'indication contraire, l'information contenue dans cette publication peut être reproduite, en tout ou en partie et par quelque moyen que ce soit, sans frais et sans autre permission d'Industrie Canada, pourvu qu'une diligence raisonnable soit exercée afin d'assurer l'exactitude de l'information reproduite, qu'Industrie Canada soit mentionné comme organisme source et que la reproduction ne soit présentée ni comme une version officielle ni comme une copie ayant été faite en collaboration avec Industrie Canada ou avec son consentement.

Pour obtenir l'autorisation de reproduire l'information contenue dans cette publication à des fins commerciales, faire parvenir un courriel à droitdauteur.copyright@tpsgc-pwgsc.gc.ca.

N.B. Dans cette publication, la forme masculine désigne tant les femmes que les hommes.

N° de catalogue lu172-1/2013-1F-PDF

ISBN 978-0-660-21280-7

Le capital de risque comme facteur de croissance[†]

Ryan Kelly*

Direction générale de la recherche économique et de l'analyse des politiques
Industrie Canada
Courriel : ryan.kelly@ic.gc.ca

Hankook Kim

Direction générale de la recherche économique et de l'analyse des politiques
Industrie Canada
Courriel : hankook.kim@ic.gc.ca

Septembre 2013

Résumé. Nous utilisons de nouvelles données liant l'information sur le financement par le capital de risque (CR) de Thomson Reuters et les données administratives au niveau d'entreprise de Statistique Canada pour comparer les entreprises financées par le CR et celles qui ne le sont pas. En particulier, nous évaluons l'impact du CR sur la croissance des variables clés au niveau de l'entreprise. Cette étude est la première du genre sur les entreprises financées par le CR, qui intègre l'information financière officielle des entreprises qui opèrent au Canada. La richesse des données nous permet de construire un groupe de contrôle des entreprises non financées par le CR en utilisant les covariables au niveau d'entreprise qui n'étaient pas disponibles auparavant et qui sont souvent associés au potentiel de croissance et à la probabilité de recevoir un financement du CR. Les nouvelles covariables incluent entre autres: les dépenses de R-D, la participation dans un programme de soutien de recherche industrielle, la taille, l'âge, l'industrie et l'emplacement. En utilisant la méthode économétrique d'appariement, nos résultats suggèrent que la croissance des entreprises financées par le CR est plus grande que celle de leurs homologues non financées par le CR. En particulier, les entreprises financées par le CR connaissent une croissance plus élevée que celles non financées par le CR sur le plan du revenu, salaire, et de l'emploi sur une période de 5 ans. En matière de dépenses de R-D, les entreprises financées par le CR connaissent une croissance plus élevée mais sur une courte période d'une année alors qu'à long terme, les niveaux de dépenses de R-D convergent. Au niveau de la variable de rentabilité, nous ne trouvons aucune différence statistique entre les entreprises financées et non financées par le CR. Dans l'ensemble, cette étude présente des résultats empiriques probants montrant que le financement par le CR est associé à une forte croissance et une accélération des processus d'innovation et de commercialisation dans les entreprises.

[†] Les idées et les opinions exprimées dans cette étude sont celles des auteurs et ne représentent aucunement les idées ou opinions du ministère de l'Industrie ou du gouvernement du Canada.

^{*} Nous tenons à remercier Statistique Canada, l'Association canadienne du capital de risque et d'investissement et la Direction générale de la petite entreprise d'Industrie Canada pour l'aide et les commentaires qu'ils nous ont apportés durant le projet. Nous aimerions souligner en particulier le travail de Jim Valerio et de Shane Dolan, qui ont supervisé l'élaboration des données et l'analyse descriptive, grâce auxquelles cette étude a été rendue possible. De même, nous voulons remercier tout spécialement Anne-Marie Rollin et Younes Errounda, qui ont travaillé sans relâche à mettre au point l'ensemble de données final. Enfin, nous aimerions remercier Thomas Hellmann and Leonard Sebbati pour leurs commentaires.

1. Introduction

Il est incontestable, à la lumière de la littérature, que le capital de risque (CR) est un mode de financement important pour les entreprises à forte croissance. Les fonds de CR s'appuient sur des connaissances spécialisées pour identifier les entreprises en démarrage qui ont un fort potentiel de croissance, et elles prennent une importante participation dans ces entreprises dans le but de les aider à devenir rentables. Dans sa stratégie de croissance, le CR ne se limite pas à apporter un financement; il assure un suivi rigoureux auprès de l'entreprise ayant reçu du financement, il lui apporte son expertise et l'aide à établir des contacts dans l'industrie. Ces soutiens additionnels peuvent accroître grandement les chances de réussite d'une entreprise, étant donné que, en général, les entreprises financées par le CR sont jeunes et possèdent des technologies, des idées ou des actifs ayant un potentiel élevé, sauf qu'elles manquent d'expérience ou de chiffre d'affaires stable pour qu'elles puissent obtenir un niveau de financement suffisant pour assurer leur expansion.

Étant donné le soutien à la gestion offert par le CR, il n'est pas étonnant d'observer l'impact du CR dans différents aspects de l'opération d'une entreprise. En particulier, la littérature souligne le rôle actif du CR dans les entreprises sur le plan de la stratégie marketing ou de la commercialisation (Hellmann et Puri, 2000), des pratiques en ressources humaines (Hellmann et Puri, 2002) et des stratégies d'innovation (Da Rin et Penas, 2007). Ce rôle actif a des répercussions positives qui se mesurent au niveau global. Ainsi, Kortum et Lerner (2000) montrent que le taux de brevetage est plus élevé dans les industries où le CR est plus présent, tandis que Samila et Sorensen (2011) reconnaissent le CR comme un facteur expliquant la hausse des activités entrepreneuriales dans une région métropolitaine.

De façon anecdotique, la proposition de valeur que présentent le CR – à savoir les contacts commerciaux, l'accès aux nouveaux marchés, l'expertise en techniques de marketing, pour ne nommer que ceux-là – cadre parfaitement bien avec l'observation selon laquelle les fonds de CR peuvent favoriser la croissance des entreprises dans lesquelles ils investissent. Cependant, tandis que la littérature démontre de nombreux exemples de l'incidence du CR sur différents aspects de l'opération d'une entreprise, il y a peu d'études empiriques qui s'intéressent à une question importante liée à ce genre de financement: celle de savoir si les entreprises qui bénéficient de ce type de financement affichent une croissance plus rapide que si elles ne l'avaient pas.

Afin d'éclairer la question, nous utilisons un nouvel ensemble de données qui lie les données de Thomson Reuters sur le financement par le CR et les données administratives de Statistique Canada pour examiner la croissance des entreprises financées par le CR. Particulièrement, nous créons un groupe de contrôle qui est similaire à l'échantillon d'entreprises financées par le CR, en tenant compte d'un certain nombre de caractéristiques qui peuvent influencer sur la croissance d'une entreprise, comme l'actif total, les dépenses en R-D et l'âge. Les covariables utilisées dans notre modèle comprennent même une variable binaire qui indique si l'entreprise a participé à un programme de soutien à la R-D, p. ex. le Programme d'aide à la recherche industrielle (PARI), ce qui peut influencer la probabilité qu'une entreprise recherche du financement par le CR ou qu'elle concentre ses efforts sur la croissance.

Nos résultats indiquent que dans la période qui suit l'investissement en CR, les entreprises qui bénéficient de ce type de financement connaissent une croissance plus élevée que celles qui sont financées autrement, et ce sur cinq plans : l'actif, le nombre d'employés, le revenu, les ventes et les salaires. Nous constatons également que les entreprises soutenues par le CR augmentent leurs dépenses en R-D plus rapidement que les entreprises comparables non financées par le CR, mais cela est vrai uniquement pour la période qui suit immédiatement le premier tour de financement.

La présente étude s'apparente à la littérature qui évaluent les résultats des entreprises financées par le CR. Parmi ceux-ci, mentionnons l'étude de Jain et Kini (1995), où les auteurs constatent que le suivi exercé par le CR a un effet positif et significatif sur la performance post introduction en bourse des entreprises ayant bénéficié de CR. Bien que les auteurs montrent que les activités de surveillance menées par le CR peuvent améliorer les résultats des entreprises, l'interprétation des résultats de leur étude se heurte à certaines limites. Premièrement, l'étude porte uniquement sur un sous-ensemble d'entreprises pour lesquelles le financement par le CR a déjà donné d'excellents résultats en arrivant à l'étape de l'introduction en bourse, et elle dépend de l'hypothèse que le CR continue de jouer un rôle actif au sein de l'entreprise après le premier appel public à l'épargne. Ces deux aspects risquent de surestimer l'importance du rôle de CR dans les résultats des entreprises. En outre, par construction, l'étude fait abstraction des résultats des entreprises avant le premier appel public à l'épargne. Or c'est justement à cette période précédant l'introduction en bourse que les activités de surveillance du CR peuvent avoir le plus d'effet sur les résultats de l'entreprise.

Dans une littérature plus récente, Puri et Zarutskie (2009) font un couplage entre les données sur le CR tirées d'une source privée et les données financières d'entreprises tirées de sources officielles aux États-Unis pour la période de 1981 à 2001. Ils observent que les entreprises financées par le CR enregistrent une croissance plus rapide de l'emploi et des ventes que celles qui ne sont pas financées par le CR, mais qu'elles affichent une rentabilité moindre par unité de vente. Toutefois, en raison d'un manque de données, les auteurs utilisent un ensemble limitée de covariables pour former le groupe de contrôle pour la comparaison. Tandis qu'ils tiennent compte de covariables importantes comme la taille et l'âge de l'entreprise, son emplacement et le secteur d'activité, ils omettent des covariables peut-être aussi importantes concernant les caractéristiques financières et la créativité de l'entreprise.

La présente étude bonifie la littérature de trois façons. Premièrement, la richesse des données nous permet d'intégrer dans l'analyse un large ensemble de covariables qui peuvent avoir une incidence sur la croissance des entreprises, notamment des mesures se rapportant aux résultats de l'entreprise en matière d'innovation, p. ex. les dépenses en R-D. Grâce à ces nouvelles covariables, nous sommes en mesure de créer un groupe de contrôle qui ressemble plus à l'échantillon d'entreprises financées par le CR que dans les études antérieures. Deuxièmement, il s'agit de la première étude empirique du genre qui exploite les données financières officielles pour évaluer les résultats des entreprises qui sont financées par le CR et qui exercent des activités au Canada. Troisièmement, la période étudiée dans cette étude nous permet de faire une évaluation plus réaliste des résultats des entreprises ayant reçu du CR, après une période où les investisseurs ont fait preuve d'une exubérance peut-être excessive à l'égard des secteurs de haute technologie.

L'étude est structurée comme suit : dans la section 2, nous décrivons les données; dans la section 3, nous exposons le cadre d'analyse et la méthodologie; la section 4 sert à la présentation des résultats, et la section 5 renferme nos conclusions.

2. Données

L'ensemble de données qui a servi à l'étude est le fruit d'un ambitieux projet qui consistait à coupler plusieurs bases de données externes à un certain nombre de bases de données administratives au niveau d'entreprise de Statistique Canada (données de Statistique Canada), à savoir les déclarations de revenus des sociétés (T2), l'état de la rémunération payée (T4) et le

registre des entreprises (RE)¹. Ces données ont été couplées avec des données externes tirées de la base de données de Thomson Reuters sur le CR (données de Thomson) et de celle du Programme d'aide à la recherche industrielle (données du PARI).

2.1 Couplage des données

Notre échantillon est prélevé parmi les entreprises constituées en société qui emploient au moins une personne². Nous utilisons le terme « entreprise » pour désigner ces entités dans notre analyse. Toutes les bases de données de Statistique Canada utilisent un même identificateur pour chaque entreprise, ce qui permet d'effectuer un couplage déterministe des données contenues dans chacune des bases de données de Statistique Canada. Ces données comprennent des renseignements détaillés sur les caractéristiques de base des entreprises, sur les éléments qui figurent habituellement dans leurs états financiers, et sur leurs employés. Particulièrement, l'actif total, le revenu total, les ventes, le bénéfice net, les bénéfices non répartis, le bénéfice brut, et les dépenses en R-D sont tirés des T2; les salaires et le nombre d'unités individuelles de main-d'œuvre (UIM)³ sont tirés des T4; la date de constitution, la province d'opération et le code principal du Système de classification des industries de l'Amérique du Nord (SCIAN) sont tirés du RE.

Les données de Statistique Canada sont couplées à deux ensembles de données externes, les données de Thomson et les données du PARI. Les données de Thomson contiennent des renseignements détaillés sur la somme et la date de tous les investissements en CR reçus par une entreprise donnée depuis le début des années 1990. Cette information sert à définir le groupe traité et le groupe de contrôle. Les données du PARI renseignent sur la somme et la date de tous les financements que les entreprises ont reçu du PARI. Combinés aux données de Statistique Canada, ces renseignements nous permettent d'identifier les covariables aux fins de l'estimation de l'appariement. En particulier, nous supposons que les entreprises bénéficiaires du PARI ont plus de probabilités de se concentrer sur les secteurs de haute technologies et de chercher des occasions de croissance tout comme les entreprises qui recherchent du CR. En outre, étant donné les similarités entre le soutien offert par le PARI et celui qui est offert par le CR, le fait de tenir compte du PARI diminuera le risque d'attribuer au financement par le CR d'éventuels écarts de croissance découlant du PARI.

Le couplage des données de Statistique Canada aux données externes a été effectué en fonction du nom et de l'adresse ainsi que du code postal et de la ville où ont lieu les activités⁴. Le nom et l'adresse des entreprises dans le RE ne sont pas toujours identiques aux données de l'entreprise qui est présumée être la même dans les données de Thomson et celles du PARI. Certains

¹ L'ensemble de données final comprend aussi des éléments d'information tirés des formulaires de versement de retenues à la source (PD7), qui fournissent des mesures plus précises de l'emploi. Toutefois, comme les données de ces formulaires sont manquantes pour les années 1999 et 2000, ce qui réduit du tiers environ la période de l'échantillon d'entreprises financées par le CR, nous n'utiliserons que les mesures de l'emploi tirées des formulaires T4 pour notre analyse.

² « L'entreprise est une unité statistique qui correspond à l'unité organisationnelle d'une firme qui dirige et contrôle l'affectation de ressources intéressant ses activités nationales et pour laquelle on établit des états financiers et des bilans consolidés à partir desquels il est possible de tirer des renseignements sur les transactions internationales, la situation des investissements internationaux et la situation financière générale de l'entité. » Adresse : <http://www.statcan.gc.ca/concepts/definitions/ent-fra.htm>.

³ L'UIM est une mesure continue du nombre annuel d'employés dans une entreprise. Chaque personne (identifiée par son numéro d'assurance social) qui reçoit un T4 correspond à 1,0 UIM, qui peut être répartie entre plusieurs entreprises selon la fraction du revenu total de la personne qui provient de chacune. Cette mesure permet de tenir partiellement compte des employés embauchés en cours d'année et des personnes qui occupent plus d'un emploi, mais ne tient pas compte des différences dans le nombre d'heures travaillées d'un employé à l'autre.

⁴ Les données du PARI ne comprennent aucune information sur l'adresse ou le code postal où ont lieu les activités, elles ont donc été couplées en fonction du nom de l'entreprise, de la ville et de la province. Même s'il repose sur une information moins approfondie, le couplage des données du PARI a donné des correspondances initiales de meilleure qualité, car le nom des entreprises y était plus souvent exact. Ce n'est pas une surprise, puisque les bénéficiaires du PARI doivent utiliser leur dénomination sociale; il était donc plus probable de la retrouver telle quelle dans le RE.

jugements ont donc été posés en vue d'améliorer le taux de couplage. Par exemple, les entreprises ont été considérées ne faire qu'un lorsqu'il y avait une légère variation dans le nom, mais que l'adresse était la même. En outre, certains couplages ont été établis manuellement par des spécialistes connaissant bien les entreprises concernées, à savoir des analystes de la direction générale de la petite entreprise d'Industrie Canada. À la suite d'un exercice de couplage rigoureux effectué en fonction de lignes directrices favorisant l'exactitude, 1 044 entreprises qui ont reçu un financement par le CR pour la première fois entre 1999 et 2009 ont été repérées dans les données de Statistique Canada. L'annexe A présente de plus amples renseignements sur le couplage des données de Statistique Canada aux données externes.

2.2 Affinement additionnel en vue de comparer les entreprises qui sont financées par le CR et celles qui ne le sont pas

L'ensemble de données couplées comporte des renseignements longitudinaux sur 1 044 entreprises financées par le CR au cours de notre période d'échantillonnage. Cependant, l'ensemble des renseignements utiles à notre analyse n'était pas disponible pour toutes ces entreprises. En particulier, plusieurs observations de ce groupe initial d'entreprises financées par le CR au cours de la période à l'étude n'ont pas émis de T2 ou de T4 dans l'année où elles ont reçu un financement de CR pour la première fois. Notre approche exige de comparer les entreprises financées par le CR à celles du groupe de contrôle au moment de la première injection du financement pour ensuite évaluer les divergences de croissance éventuelles. Nous n'avons donc pas le choix d'exclure de notre échantillon les entreprises pour lesquelles manquent des renseignements financiers au moment du premier financement par le CR. Outre cette contrainte, nous en imposons une autre en limitant le groupe expérimental aux industries à 4 chiffres du SCIAN pour lesquelles au moins trois entreprises ont reçu un financement par le CR entre 1990 et 2009 selon les données de Thomson. Cette limite vise à réduire l'effet potentiellement imputable à l'attribution du mauvais code SCIAN à certaines entreprises⁵. Une fois que les données ont été ajustées de la sorte, nous arrivons à 662 entreprises financées par le CR qui sont susceptibles de faire partie du groupe traité. La taille finale du groupe traité dépendra du nombre d'homologues appropriés dans la population générale d'entreprises selon la méthode d'appariement utilisée.

2.3 Longitudinalisation

L'ensemble initial de données couplées est essentiellement constitué d'une série de données transversales annuelles pour la période de 1999 à 2009. Cette base de données comporte des lacunes potentielles importante qui pourraient nous empêcher d'effectuer une analyse rigoureuse de la croissance des entreprises canadiennes financées par le CR. Toute entreprise se voit attribuée un identificateur dans le RE, lequel permet d'identifier l'entreprise dans toutes les bases de données de Statistique Canada. Cependant, les identificateurs peuvent parfois changer pour d'autres raisons que la disparition ou l'apparition d'une entreprise, p. ex. un changement de dénomination sociale. De plus, les fusions et acquisitions ne sont pas reflétées dans les données de Statistique Canada. En vue de pallier à ces lacunes et de faire en sorte que la croissance soit mesurée correctement et attribuée à la bonne entreprise, nous avons longitudinalisé les données de Statistique Canada au moyen d'une méthode de suivi de la main-d'œuvre. En pratique, il s'agit de suivre les employés à l'aide des T4 émis et des numéros d'assurance sociale y figurant, de

⁵ Même si leur utilisation est répandue en raison de leur commodité, les codes du SCIAN peuvent donner de piètres résultats comme indicateurs du marché dans lequel l'entreprise exerce ses activités, car les entreprises y sont classées en fonction de la production plutôt que de la demande. Comme les investisseurs de CR se concentrent sur des marchés dans lesquels ils ont des connaissances techniques ou une forme quelconque d'expertise, il se pourrait qu'une entreprise qui a reçu du capital de risque sous un code SCIAN où il y a un faible taux d'investissements de ce type ait été classée dans le mauvais code, ou qu'il n'y ait pas d'entreprises lui ressemblant sous ce code SCIAN.

manière à repérer tout changement structurel dans l'entreprise. Par exemple, supposons que l'on observe pour un identificateur du RE donné une cessation d'activité économique alors que le niveau d'emploi correspondait à 50 personnes. Si au même moment un nouvel identificateur du RE apparaît dans l'échantillon et qu'il emploie plus de 25 personnes qui étaient associées à l'identificateur du RE précédent, et que nous ne voyons aucune relation substantielle entre ce premier identificateur du RE et d'autres identificateurs du RE dans la population, nous pouvons conclure que les deux identificateurs du RE correspondent à une seule et même entreprise. Cette procédure nous permet de créer un enregistrement longitudinal unique pour chacune des entreprises financées par le CR et pour leurs homologues du groupe de contrôle, en utilisant comme année de référence l'année où le financement par le CR a été obtenu pour la première fois ou bien l'année où l'appariement a été effectué. L'annexe B présente de plus amples renseignements sur la longitudinalisation.

3. Appariement des entreprises financées par capital de risque aux entreprises qui ne le sont pas

Nous cherchons à estimer l'effet du financement par le CR et de son soutien sur diverses mesures du rendement des entreprises. Plus précisément, nous définissons $Y_i(1)$ comme la mesure du rendement de l'entreprise i (e.g. la croissance du nombre d'employés ou de la R-D) ayant reçu le traitement (i.e. ayant reçu du capital de risque), et $Y_i(0)$ comme la mesure du rendement de l'entreprise i n'ayant pas reçu le traitement. Idéalement, il suffirait de calculer l'effet moyen du traitement sur l'entreprise traitée (ATT)⁶ à l'aide de l'équation 1,

$$ATT = E[Y_i(1) - Y_i(0) | VC = 1] = \frac{1}{N_t} \sum_{i|VC=1} [Y_i(1) - Y_i(0)], \quad 1)$$

où N_t représente le nombre d'entreprises traitées et VC est une variable binaire qui indique si l'entreprise a reçu ou non du CR. Cependant, comme c'est en général le cas dans les études d'évaluation, nous ne pouvons observer qu'une de ces deux possibilités pour la même entreprise i . Par conséquent, la littérature sur l'évaluation des effets du traitement aborde ce problème comme s'il s'agissait d'un problème de données manquantes. La pratique normale consiste à remplacer les données manquantes des situations contrefactuelles (non observables) par les données d'entreprises non traitées qui auraient des caractéristiques similaires aux entreprises traitées. Les différences dans les mesures du rendement entre les deux groupes pouvant alors idéalement être attribuées aux différences du traitement plutôt qu'aux différences de caractéristiques des entreprises. Dans la pratique, il est possible d'appliquer une méthode d'appariement générale et, sous certaines conditions, obtenir des estimations cohérentes de l'ATT.

L'application d'un simple estimateur d'appariement exige deux hypothèses de base. D'abord, il doit exister un ensemble de covariables, X , de sorte que les résultats de l'entreprise sont indépendants du traitement une fois ces caractéristiques prises en considération. Comme nous nous intéressons précisément à l'estimation de l'ATT, nous pouvons nuancer cette hypothèse pour invoquer l'indépendance moyenne. Heckman, Ichimura et Todd (1998) offrent de plus amples renseignements à ce sujet. Ensuite, nous supposons que pour tout X , il existe une probabilité positive de recevoir ou non du financement par le CR. En gros, cette deuxième hypothèse stipule qu'il existe dans la population générale des homologues similaires

⁶ Comme seul un petit sous-ensemble de la population des entreprises canadiennes est considéré comme une cible viable pour un financement de CR, nous n'évaluons pas l'effet moyen du traitement (ATE), puisque le résultat correspondrait à un effet moyen pour la population entière, ce qui présente peu d'intérêt et est difficile à interpréter utilement.

aux entreprises traitées, ce qui permet d'établir un groupe de contrôle convenable. Plus précisément, ces hypothèses s'écrivent comme suit :

Hypothèse 1 : Indépendance conditionnelle ou ignorabilité – $E[Y(0), Y(1)] \perp (VC | X)$

Hypothèse 2 : Région de support commun – $0 < \Pr(VC = 1 | X) < 1$

Si ces hypothèses sont satisfaites, l'équation 2 produit des estimations cohérentes de l'ATT.

$$ATT = \frac{1}{N} \sum_{i=1}^N [(Y_i(1) | VC = 1, X = x) - (Y_i(0) | VC = 0, X = x)], \quad 2)$$

où le deuxième terme compris dans l'opérateur de sommation correspond au résultat moyen d'un groupe de contrôle apparié, en fonction de l'ensemble de covariables X , à l'échantillon d'entreprises financées par le CR. À partir des hypothèses 1 et 2, plusieurs méthodes d'appariement différentes permettent d'établir un groupe de contrôle en fonction de l'ensemble de covariables.

Nous analysons d'abord l'échantillon d'entreprises financées par le CR pour déterminer comment elles se comparent, au moment où elles reçoivent du financement pour la première fois, à la population générale des entreprises canadiennes au cours de la période d'échantillonnage (de 1999 à 2009). Une comparaison directe des moyennes de ces deux groupes ne serait pas très utile. Par rapport à l'ensemble des entreprises, les entreprises financées par le CR sont en majorité concentrées au début de la période d'échantillonnage. En outre, la répartition industrielle de la population générale n'est pas la même que celle des entreprises financées par le CR, ces dernières étant disproportionnellement concentrées dans les services professionnels. Pour surmonter ces problèmes, nous calculons le ratio moyen de la valeur associée aux entreprises financées par le CR sur la valeur moyenne obtenue pour le code à quatre chiffres du SCIAN correspondant, dans la première année où le financement est injecté. Dans le tableau 1, nous présentons ces ratios et indiquons s'ils s'écartent statistiquement de la valeur unitaire. Les résultats révèlent que notre échantillon de 662 entreprises financées par le CR est significativement différent de la population générale des entreprises canadiennes en termes de plusieurs covariables qui pourraient avoir une influence notable sur la croissance. Par conséquent, une méthode d'appariement est nécessaire en vue d'établir un groupe de contrôle approprié.

Tableau 1 – Comparaison entre les entreprises financées par le CR et la population générale des entreprises pour la première année où du financement de CR est injecté

Variable	Multiple relative par rapport à la population générale *	Nombre d'observations	Valeur <i>p</i>
Actif total	2,36	662	0,000
Ventes	1,38	662	0,002
Emploi	2,43	662	0,000
Salaires	1,16	662	0,000
Âge	0,61	662	0,000
Bénéfice brut	1,59	662	0,000
Marge bénéficiaire brute**	0,77	482	0,000
Bénéfice brut par emploi**	0,59	485	0,000
Dépenses en R-D***	1,82	432	0,000

* moyenne de{ $X(GE_{f,i,t})$ / moyenne de{ $X(GC_{f,i,t})$ } }

f – entreprise; i – code à quatre chiffres du SCIAN;

t – année

La population générale est l'ensemble des entreprises actives au Canada ayant déclaré les renseignements financiers nécessaires et appartenant au code à quatre chiffres du SCIAN au cours de la période d'échantillonnage, de 1999 à 2009 (2 573 663 observations).

** Moyenne des entreprises affichant un bénéfice (et non une perte)

*** Moyenne des entreprises ayant effectué de la R-D

3.1 Appariement par correspondance exacte

La méthode d'appariement la plus simple et exigeant le plus de données est celle de la *correspondance exacte*. En pratique, elle consiste à établir des paires composées d'une observation traitée et d'une observation non traitée pour lesquelles les covariables sont rigoureusement identiques. Même si la méthode de correspondance exacte produit des appariements plus robustes, son application, notamment à un grand ensemble de covariables ou à des covariables continues, nécessite un très grand nombre de données et n'est souvent pas une option viable pour le traitement de données empiriques. Les chercheurs ont la possibilité de réduire le nombre de covariables pris en considération ou de les regrouper pour utiliser une variante de la méthode de correspondance exacte. Mais ces opérations risquent d'accroître la probabilité de violation de l'hypothèse d'ignorabilité.

Dans la littérature sur le rendement des entreprises financées par le CR, Puri et Zarutskie (2009) ont utilisé la méthode de correspondance exacte dans leur stratégie empirique. Toutefois, ils l'ont fait en appariant les entreprises en fonction d'un petit nombre de covariables, pour la plupart discrètes, c.-à-d. l'âge de l'entreprise, le code d'industrie, la région géographique et le nombre d'employés. Nous appliquons une méthode de correspondance exacte dans la présente étude de façon à permettre les comparaisons avec Puri et Zarutskie (2009). Cependant, nos résultats donnent à penser que la correspondance exacte n'est peut-être pas la meilleure façon de traiter nos données. En particulier, étant donné l'objectif de nos recherches, nous choisissons d'apparier les entreprises en fonction de variables financières au début de l'étude, et nous examinons l'évolution de ces variables une fois l'investissement effectué. Selon cette méthode, il faut prendre en considération un ensemble important de covariables continues, et nous n'avons simplement pas assez de données pour tenir compte de toutes les caractéristiques pertinentes des entreprises tout en effectuant des correspondances exactes. Nos résultats mettent notamment en lumière le fait que des entreprises appariées peuvent être pratiquement identiques sur certains aspects, et pourtant très différentes selon des mesures connexes pouvant avoir une grande incidence. La

section 4, où sont présentés les résultats empiriques, offre de plus amples renseignements à ce sujet.

3.2 Appariement par score de propension

Étant donné nos objectifs de recherche et la richesse des données dont nous disposons, nous avons besoin d'une méthode d'appariement mieux adaptée à un large ensemble de covariables continues que celle de la correspondance exacte. Nous utilisons l'appariement par score de propension, qui a d'abord été décrite par Rosenbaum et Rubin (1983). Cette méthode consiste à estimer la probabilité de recevoir le traitement (le score de propension) et à appairer des entreprises traitées à des entreprises susceptibles de faire partie du groupe de contrôle qui ont un score de propension très similaire. En pratique, le score de propension résume la distribution des covariables pour chacune des observations en tenant compte du rôle de chaque covariable dans la possibilité de recevoir le traitement. Il s'ensuit que l'appariement par score de propension nous permet de tenir compte d'un nombre accru de caractéristiques, car nous pouvons établir un groupe de contrôle au sein duquel la distribution des covariables voulues est semblable, sans pour autant que la correspondance de chacune des covariables soit exacte (ou presque exacte).

Notre estimateur d'appariement par score de propension fonctionne en trois étapes.

D'abord, nous ajustons un modèle logit en vue de calculer la probabilité qu'une entreprise de la population générale reçoive du CR, comme dans l'équation 3.

$$\Pr(VC_{it} = 1) = \alpha_c + \alpha_t + \alpha_{ind} + \beta X_{it} + \varepsilon_{it}, \quad 3)$$

où X_{it} représente l'ensemble des covariables étudiées, α_c est la constante, α_t et α_{ind} représentent les ordonnées à l'origine propres au temps et à l'industrie, et ε_{it} correspond au terme d'erreur normal en économétrie.

Ensuite, nous définissons des seuils de tolérance pour les scores de propension calculés à la première étape, en vue de déterminer pour chacune des entreprises traitées un ensemble de partenaires potentiels dont sont exclues les mauvaises correspondances. Même si le choix de ce seuil est par définition assez arbitraire, les chercheurs peuvent se fier à certaines lignes directrices. En particulier, Rosenbaum et Rubin (1985) suggèrent comme point de départ général un écart de 0,25 multiplié par l'écart-type de score de propension linéaire, ce que nous appliquons dans nos appariements, comme le montre l'équation 4. S'il n'existe aucune correspondance satisfaisante pour une entreprise financée par le CR, celle-ci est éliminée de l'échantillon puisqu'il n'y a pas d'homologue convenable avec lequel la comparer parmi les entreprises n'ayant pas reçu de CR.

$$\delta = 0.25 * \sigma \left\{ \ln \left(\frac{\Pr(VC = 1)}{(1 + \Pr(VC = 1))} \right) \right\}. \quad 4)$$

Enfin, nous établissons le groupe de contrôle en appariant chaque entreprise traitée à l'entreprise qui, au sein du groupe d'homologues potentiels cerné à la deuxième étape, a le score de propension linéaire le plus proche. Plus précisément, supposons que p correspond au score de propension linéaire et I_0 à l'ensemble d'entreprises non financées par le CR classées sous le même code à quatre chiffres du SCIAN et opérant dans la même province à l'année où le CR est injecté pour la première fois. L'appariement pour l'entreprise i , M_i , est déterminé par :

$$M_i = \left\{ \min_j \left\| \text{abs}(p_i - p_j) \right\|, j \in I_0 \mid \text{abs}(p_i - p_j) < \delta \right\}. \quad 5)$$

Une fois deux entreprises appariées (une financée par le CR et l'autre non), nous les retirons du bassin d'observations à appairer. Selon la terminologie d'appariement, nous utilisons une méthode d'appariement par recherche du plus proche voisin selon le score de propension linéaire sans remplacement. Comme certaines des entreprises n'ayant pas reçu de CR pourraient représenter la correspondance la plus proche pour plusieurs entreprises financées par le CR, l'ordre dans lequel les entreprises sont appariées peut avoir une incidence sur les entreprises qui seront ajoutées au groupe de contrôle. Étant donné qu'il existe un bon nombre de correspondances potentielles pour plusieurs des entreprises du groupe traité, la méthode d'appariement optimale visant à minimiser une mesure globale de la distance exigerait des calculs excessivement lourds qui dépassent la portée de la présente étude. Par conséquent, nous appariions les entreprises du groupe traité selon l'ordre dans lequel elles ont reçu un financement par le CR pour la première fois, de manière à maximiser la durée des enregistrements longitudinaux utilisés dans l'étude⁷.

Il est important de souligner que la méthode ci-dessus ne peut tenir compte que des différences observables entre le groupe traité et le groupe de contrôle, comme le veut l'hypothèse d'ignorabilité. Cependant, comme le note Lerner (2010), l'investisseur de CR moyen consacre souvent plus de 100 heures à l'évaluation d'un investissement potentiel. Les efforts visant à atténuer tous les écarts entre l'échantillon d'entreprises financées par le CR et celui d'entreprises non financées par le CR en utilisant les données administratives seulement sont sans doute voués à l'échec peu importe la méthode choisie. Cependant, même si nous ne pouvons pas entièrement faire la distinction entre les résultats qui sont attribuables à des caractéristiques de sélection et ceux qui sont attribuables à l'effet des investissements de CR et aux conseils qui les accompagnent, l'ampleur de notre exercice d'appariement et l'utilisation de nouvelles covariables dans notre étude font en sorte que nos résultats devraient offrir des perspectives économiques pertinentes sur la croissance des entreprises financées par le CR.

4. Résultats

4.1 Appariement par correspondance exacte

Nous utilisons un estimateur de correspondance exacte semblable à celui de Puri et Zarutskie (2009). Quatre covariables sont prises en considération : le code à quatre chiffres du SCIAN, la province d'activité, l'âge en années depuis la constitution et la taille, mesurée par le nombre d'UIM déterminée à partir des T4. Puri et Zarutskie (2009) ont utilisé des catégories légèrement différentes d'industries et d'emplacements. En outre, ils ont utilisé un dénombrement annuel de l'effectif comme mesure de la taille des entreprises, alors que nous utilisons une mesure plus précise de l'emploi qui tient compte des travailleurs à temps partiel et de ceux qui sont embauchés en cours d'année. Comme notre mesure d'UIM est continue et comporte des décimales, nous appliquons un seuil de 10 %, ce qui signifie que les entreprises sont signalées comme correspondantes possibles lorsque leur UIM est à moins de 10 % de celle de l'entreprise du groupe traité, ou à moins d'une UIM de celle de l'entreprise du groupe traité lorsque cette dernière compte moins de 10 UIM. Cette méthode a produit 377 paires d'entreprises, ce qui représente un taux d'appariement de 57 %.

⁷ Bien qu'il n'en soit pas question ici, l'appariement avec remplacement donnait des résultats similaires, ce qui laisse croire que peu de gains auraient pu être effectués en ayant recours à une méthode d'appariement optimal.

Le tableau 2 montre une comparaison entre le groupe traité et le groupe de contrôle selon plusieurs covariables. Comme la méthode de correspondance exacte signifie, par définition, que les entreprises seront situées dans la même province et dans la même catégorie d'industries, il est inutile d'inclure ces covariables aux fins de la comparaison. La partie supérieure du tableau établit les comparaisons en fonction des covariables prises en considération pour l'appariement, et comme on est en droit de s'y attendre, les entreprises appariées affichent des profils d'emploi très similaires et ont le même âge. La partie inférieure du tableau montre la comparaison selon d'autres variables dignes d'intérêt qui pourraient influencer sur les décisions d'investissement sous forme de CR et sur la croissance des entreprises. Bien que la méthode de correspondance exacte se traduise par des correspondances très proches sur un petit nombre de covariables, les résultats à la partie inférieure du tableau montrent l'existence d'écarts substantiels entre les entreprises du groupe traité et les entreprises du groupe de contrôle avec lesquelles elles sont appariées en ce qui concerne toutes les autres variables. Les écarts sont considérables d'un point de vue statistique, puisque la valeur p est à zéro au niveau de trois décimales pour toutes les mesures. Il s'agit d'une indication statistique forte que les entreprises appariées du groupe de contrôle sont différentes des entreprises du groupe traité sur ces aspects.

Ces résultats permettent de douter sérieusement que l'approche par correspondance exacte fondée sur ces covariables soit la meilleure approche empirique pour la présente étude. Certaines des variables pour lesquelles de grands écarts sont observables entre entreprises appariées pourraient jouer un rôle crucial dans la croissance des entreprises financées par le CR. En moyenne, comparativement à celles du groupe de contrôle, les entreprises du groupe traité ont un actif plus important et des travailleurs plus qualifiés (à en juger par les salaires), mais un revenu et des ventes moindres. Étant donné ces différences initiales, il pourrait être difficile d'interpréter la croissance subséquente des entreprises. Par exemple, le fait que les entreprises financées par le CR aient initialement des ventes moins importantes que les entreprises du groupe de contrôle, malgré un actif et des salaires supérieurs, pourrait indiquer que les entreprises du groupe de contrôle ont généralement un bassin de clientèle relativement mature, ce qui limiterait leur potentiel de forte croissance à court terme.

Tableau 2 – Comparaison entre les groupes – Appariement par correspondance exacte

Covariable	Moyenne (Financée par le CR)	Moyenne (Groupe de contrôle)	Écart	Valeur p de l'écart	Écart normalisé entre les moyennes
Covariables utilisées					
Ln emploi	2,193	2,185	0,008	0,931	0,007
Âge	3,694	3,694	0,000	1,000	0,000
Autres mesures pertinentes					
Ln actif total	14,081	12,949	1,132	0,000	0,749
Argsinh ventes	9,877	13,141	-3,264	0,000	-0,531
Ln salaires	10,754	10,426	0,328	0,000	0,620
Argsinh bénéfices non répartis	-10,825	0,704	-11,529	0,000	-1,247
Argsinh revenu	12,389	13,953	-1,563	0,000	-0,338
Argsinh bénéfice net	-9,926	1,156	-11,082	0,000	-1,159
Argsinh dépenses en R-D	9,435	3,212	6,223	0,000	1,034
Argsinh bénéfice brut	8,346	11,851	-3,506	0,000	-0,457

Écart normalisé entre les moyennes = (moyenne_X[GE] – moyenne_X[GC]) / Éc. norm._X[GE]

Argsinh signifie « argument sinus hyperbolique ».

Observations du groupe traité = 377; Observations du groupe de contrôle = 377.

Dans le tableau 3, nous nous intéressons particulièrement à l'équilibre entre le nombre d'entreprises qui, dans le groupe traité et dans le groupe de contrôle, affichent des valeurs positives pour deux des indicateurs de l'innovation. Les dépenses en R-D constituent un intrant du processus d'innovation et sont souvent un facteur de croissance, d'innovation et de croissance de la productivité. Dans les résultats obtenus par correspondance exacte, on constate un déséquilibre entre les deux groupes quant aux dépenses en R-D moyennes (tableau 2) et au nombre d'entreprises effectuant de la R-D (tableau 3).

Que l'entreprise obtienne ou non un financement du PARI peut aussi servir d'indicateur en vue de repérer les entreprises dont la stratégie est axée sur l'innovation. Pour obtenir un financement du PARI, les entreprises passent par un processus de sélection au cours duquel elles doivent démontrer que leur projet ou leur plan d'affaires est commercialement viable et fait une place suffisante à l'innovation. Outre les aspects techniques du projet, le processus de diligence raisonnable du PARI prend aussi en considération les capacités opérationnelles et de gestion de l'entreprise. Même si le soutien offert aux entreprises dans le cadre du PARI diffère de celui qui accompagne les investissements de CR (le PARI offre un financement direct plutôt que de prendre une participation dans l'entreprise), il comprend des services de mentorat semblables et vise le même type d'entreprises. Par conséquent, l'obtention de financement du PARI peut être interprété comme l'une des caractéristiques les plus révélatrices en vue d'évaluer la similarité entre une entreprise financée par le CR et une entreprise pouvant servir de contrôle. Cependant, comme dans le cas des dépenses en R-D, le groupe de contrôle issu de l'appariement par correspondance exacte ne comporte pas un nombre d'entreprises financées par le PARI comparable à celui du groupe traité. Il y a presque quatre fois plus de bénéficiaires du PARI dans le groupe traité que dans le groupe de contrôle, ce qui laisse croire que les entreprises du groupe traité mettent davantage l'accent sur le développement et la commercialisation de produits innovants. Ces résultats alimentent les doutes quant à la similitude des perspectives de croissance des deux groupes au moment de l'appariement.

Tableau 3 – Comparaison entre les groupes – Appariement par correspondance exacte, R-D et financement du PARI

	Nombre		Pourcentage	
	Groupe traité (CR)	Groupe de contrôle	Groupe traité (CR)	Groupe de contrôle
Activités de R-D à l'année d'appariement	242	87	64	23
Activités de R-D sur la période d'échantillonnage	277	123	73	33
Financement du PARI à l'année d'appariement	76	18	20	5
Financement du PARI sur la période d'échantillonnage	116	23	31	6

Observations du groupe traité = 377; Observations du groupe de contrôle = 377.

Pour permettre des comparaisons robustes, il faut que toutes les covariables ayant une influence potentielle sur la croissance soient similaires entre le groupe traité et le groupe de contrôle. Cependant, la méthode d'appariement par correspondance exacte a produit un groupe traité et un groupe de contrôle qui ont des profils nettement différents en ce qui a trait à plusieurs de ces covariables. Il n'est pas possible de résoudre ce problème en tenant simplement compte d'un plus grand nombre de covariables dans le processus d'appariement, car pour établir des correspondances exactes selon toutes les caractéristiques dont nous disposons tout en conservant un taux d'appariement raisonnable, la quantité de données à traiter serait beaucoup trop grande. Nous abandonnons donc cette méthode, étant donné que les données ne nous permettent pas d'être certains que le groupe de contrôle établi satisfait l'hypothèse d'ignorabilité. Or, si cette dernière est violée, les estimations de l'effet moyen du traitement sur les traités ne seront pas cohérentes.

4.2 Appariement par score de propension

L'utilisation de l'appariement par score de propension nous impose d'abord de préparer un modèle permettant d'estimer la probabilité de recevoir le traitement. Les résultats de l'estimation logit du score de propension figurent au tableau C1 de l'annexe C. Pour augmenter la probabilité que l'hypothèse d'ignorabilité soit satisfaite, nous tenons compte de toutes les variables connues qui pourraient avoir un effet sur la croissance de l'entreprise ou sur la probabilité qu'elle reçoive un financement par le CR. Nous prenons aussi en considération le carré de chaque variable continue, de façon à tenir compte des relations non linéaires. Les variables utilisées dans l'estimation logit sont l'actif total, les ventes, le nombre d'UIM, les salaires, les bénéfices non répartis, le revenu, le bénéfice net, l'âge, les dépenses en R-D, une variable binaire indiquant si l'entreprise a reçu ou non du financement du PARI, une variable binaire pour les entreprises financées par le CR en 1999 qui ont effectué des activités de R-D en 2000⁸, les effets fixes attribuables à l'industrie telle qu'elle est identifiée par le code à quatre chiffres du SCIAN, et les effets fixes attribuables à l'année.

Cette méthode d'appariement par score de propension donne un taux d'appariement de 82 %, soit 544 paires. Le tableau 4 montre les écarts entre le groupe traité et le groupe de contrôle. Ils sont très similaires en tous points, et les valeurs de p des écarts sont assez élevées pour suggérer qu'il n'existe pas de différence statistiquement significative entre les deux groupes. Outre la

⁸ Les données sur les dépenses en R-D ne sont pas accessibles pour l'année 1999. Il convient de souligner que dans la section suivante, le calcul du taux de croissance des dépenses en R-D ne tient pas compte des données de 1999. La variable binaire a pour unique fonction d'améliorer les appariements pour les entreprises innovantes ayant reçu du CR en 1999.

vérification normale des hypothèses, nous présentons les résultats d'une autre mesure de comparaison qui prend en considération le déséquilibre de l'échantillon, soit l'écart normalisé entre les moyennes (ÉNM). Cette mesure est souvent jugée particulièrement fiable pour l'évaluation des différences entre deux échantillons en présence du déséquilibre, car les vérifications normales présentent un risque élevé d'erreurs de type II lorsque l'échantillon est de petite taille. Rubin (2001) suggère qu'il peut y avoir des problèmes lorsque les valeurs d'ÉNM sont supérieures à 0,5. Comme le montre le tableau 4, les valeurs d'ÉNM de toutes les covariables dont nous tenons compte et du score de propension linéaire se situent en dessous de ce seuil, ce qui indique que les deux groupes sont bien équilibrés.

Tableau 4 – Comparaison entre les groupes — Appariement par score de propension

Covariable	Moyenne (Financée par le CR)	Moyenne (Groupe de contrôle)	Écart	Valeur de <i>p</i> de l'écart	Écart normalisé entre les moyennes
Ln actif total	14,485	14,453	0,032	0,735	0,020
Argsinh ventes	11,618	11,835	-0,217	0,539	-0,037
Ln emploi	2,717	2,765	-0,049	0,544	-0,036
Ln salaires	10,683	10,674	0,009	0,787	0,016
Argsinh bénéfiques non répartis	-8,089	-7,835	-0,254	0,727	-0,021
Argsinh revenu	13,507	13,454	0,053	0,850	0,012
Argsinh bénéfice net	-7,572	-7,305	-0,267	0,707	-0,023
Âge	4,952	5,042	-0,090	0,796	-0,016
Argsinh dépenses en R-D	8,228	8,456	-0,227	0,566	-0,035
Argsinh bénéfice brut	9,617	9,201	0,416	0,410	0,052
Score de propension linéaire	-4,729	-4,754	0,025	0,860	0,011

Écart normalisé entre les moyennes = (moyenne_X[GE] – moyenne_X[GC]) / Éc. norm._X[GE].

Argsinh signifie « argument sinus hyperbolique ».

Observations du groupe traité = 544; Observations du groupe de contrôle = 544.

L'amélioration de l'équilibre des covariables par l'appariement par score de propension est aussi mise en lumière par une comparaison du nombre d'entreprises effectuant de la R-D et du nombre d'entreprises bénéficiaires du PARI. Le tableau 5 présente ces résultats, qui renforcent notre confiance à l'égard de la qualité du groupe de contrôle. Le nombre d'entreprises effectuant de la R-D et le nombre d'entreprises bénéficiaires du PARI sont presque identiques entre les deux groupes, ce qui laisse croire que les entreprises de chacun des groupes mettent un accent comparable sur les stratégies d'innovation.

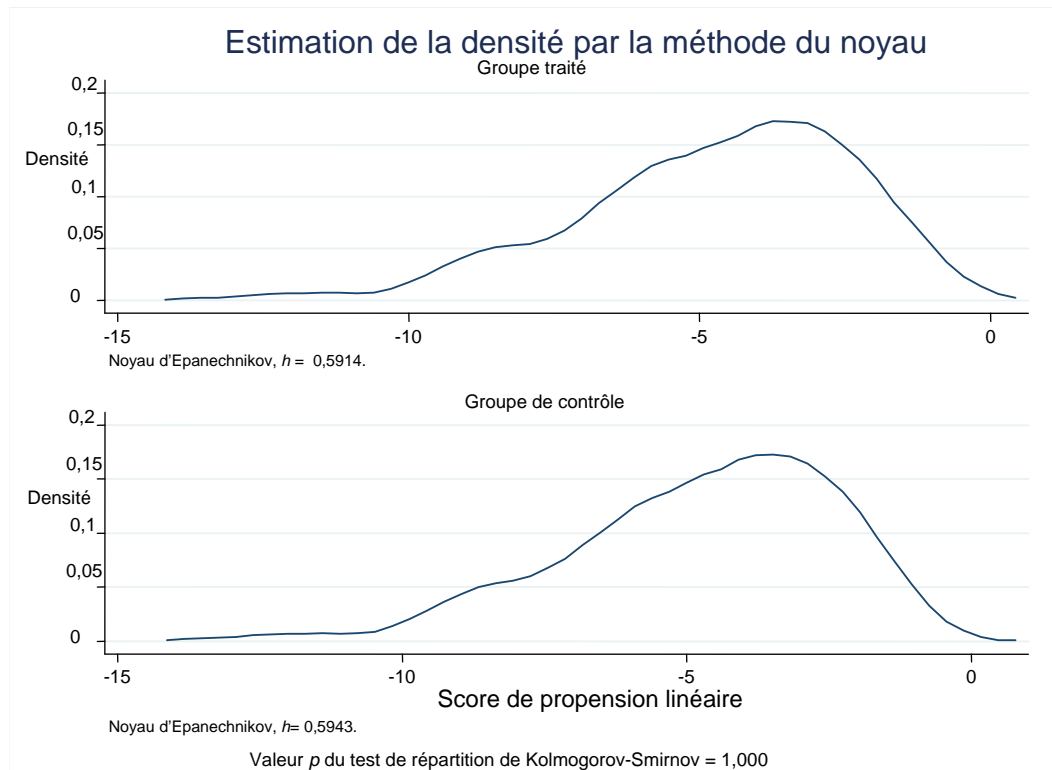
Tableau 5 – Comparaison entre les groupes — Appariement par score de propension, R-D et PARI

	Nombre		Pourcentage	
	Groupe traité (CR)	Groupe de contrôle	Groupe traité (CR)	Groupe de contrôle
Activités de R-D, à l'année d'appariement	336	347	62	64
Activités de R-D sur la période d'échantillonnage	412	396	76	73
Financement du PARI à l'année d'appariement	91	88	17	16
Financement du PARI sur la période d'échantillonnage	152	156	28	29

Observations du groupe traité = 544; Observations du groupe de contrôle = 544.

Au-delà de la comparaison des moyennes et du simple dénombrement des entreprises, nous comparons la distribution réelle du score de propension linéaire du groupe traité et du groupe de contrôle à la figure 1⁹. L'estimation des coefficients de propension par la méthode de densité du noyau montre que la distribution du score de propension linéaire est pratiquement identique entre les deux groupes. Cette évaluation visuelle est corroborée par le test de distribution de Kolmogorov-Smirnov, dont la valeur p est très près de 1, ce qui signifie que les deux distributions sont identiques en termes statistiques.

Figure 1 – Tracé de l'estimation des coefficients de propension par la méthode de densité du noyau



⁹ Malheureusement, il nous est impossible de reproduire cet exercice sur d'autres variables sans enfreindre les règles de Statistique Canada en matière de confidentialité.

Les résultats présentés aux tableaux 4 et 5 et à la figure 1 montrent que notre estimateur d'appariement est robuste, car les principales caractéristiques des deux échantillons sont similaires, p. ex. dépenses en R-D, taille, âge et rentabilité. Contrairement à nos échantillons appariés par correspondance exacte, les entreprises sont assez similaires à l'égard de toutes les covariables, en particulier celles qui sont importantes dans le cadre de notre question de recherche, comme la taille et les activités de R-D.

4.3 Comparaisons de croissance

Forts d'un groupe traité et d'un groupe de contrôle convenablement appariés, nous pouvons comparer le taux de croissance de chaque groupe sur différentes périodes. Le calcul du taux de croissance en question correspond à la formule suivante :

$$\text{Croissance sur } y \text{ années} = \frac{\sum_{i=1}^n (\ln X_{i,t+y} - \ln X_{i,t})}{n}, \quad 6)$$

où y représente le nombre d'années sur lequel le taux de croissance est calculé. Nous calculons le taux de croissance par intervalles d'un, de trois et de cinq ans. Pour ce faire, nous ne tenons compte que des entreprises qui affichent des valeurs positives au début et à la fin de la période étudiée. En outre, nous éliminons les entreprises lorsqu'il y a sortie (fusion ou acquisition) avant la fin de la période. L'annexe B donne des détails sur la identification de ces sorties.

Toutes les mesures de rendement exprimées en dollars sont des valeurs nominales, car les déflateurs propres à l'industrie et à l'année ne figurent pas parmi nos données. Ainsi, tous les taux de croissance, sauf ceux qui se rapportent à l'emploi ou à la marge bénéficiaire brute, sont nominaux. Malgré cette limite, les comparaisons de la croissance entre les entreprises financées par le CR et les entreprises qui ne l'ont pas été restent valides puisque, par construction, les entreprises du groupe traité et celles du groupe de contrôle appartenaient à la même industrie et avaient le même âge au moment de l'appariement, de sorte qu'elles devraient avoir été soumises à la même inflation au cours de la période d'échantillonnage.

Cependant, deux complications potentielles se posent. D'abord, comme nous mesurons le taux de croissance jusqu'à cinq ans, l'attrition au sein de l'échantillon pourrait perturber l'équilibre initial établi en fonction de l'âge et de l'industrie si certaines entreprises restent dans l'échantillon après la sortie de son homologue. Ensuite, les comparaisons de taux de croissance pourraient être faussées par un biais du survivant si, par exemple, les entreprises à faible croissance ont plus de chances de survie dans un groupe que dans l'autre. Pour tenir compte de ces problèmes, nous reproduisons les résultats en ne tenant compte que des paires où les données nécessaires au calcul du taux de croissance sont disponibles pour les deux entreprises. L'exercice a donné lieu à des résultats qualitativement et quantitativement similaires à nos résultats principaux, en dépit d'une signification statistique légèrement moindre pour une minorité de comparaisons de croissance. Le tableau D1 de l'annexe D présente ces résultats.

Dans l'esprit de notre question de recherche principale, nous comparons le taux de croissance des entreprises du groupe traité et celui des entreprises du groupe de contrôle selon quatre mesures de la taille de l'entreprise : l'emploi, l'actif total, les revenus et les ventes. Comme le montre le tableau 6, après l'appariement, la croissance a été beaucoup plus forte dans le groupe qui a reçu du financement par le CR. Les taux de croissance annuelle les plus élevés s'observent dans l'année suivant l'injection de capital; ils varient, selon l'indicateur, de 20 % dans le cas de l'actif total à 50 % dans le cas des revenus totaux. Cette croissance initiale semble se limiter à la

première année dans le cas de certaines mesures. Par exemple, les taux cumulatifs de croissance de l'emploi pour les entreprises financées par le CR, sur 3 ans et sur 5 ans, sont respectivement de 41,5 % et de 50,6 %, soit une faible hausse par rapport au taux de croissance sur un an, qui est de 34,4 %. La croissance de l'emploi est réelle, mais les autres taux de croissance sont nominaux et comprennent l'inflation. Par conséquent, il faut évaluer l'écart entre la croissance des entreprises financées par le CR et la croissance des entreprises n'ayant pas reçu de CR pour déterminer l'effet de CR sur la croissance.

Tableau 6 – Taux de croissance dans le groupe traité et le groupe de contrôle, mesures de la taille des entreprises

	Moyenne (Financée par le CR)	Moyenne (Groupe de contrôle)	Écart	Valeur <i>p</i> de l'écart	Observations (Financée par le CR)	Observations (Groupe de contrôle)
Actif total						
<i>Croissance (1 an)</i>	19,6 %	0,1 %	19,5 %	0,001	487	493
<i>Croissance (3 ans)</i>	38,7 %	-10,2 %	48,8 %	0,000	334	360
<i>Croissance (5 ans)</i>	53,5 %	1,1 %	52,4 %	0,001	215	254
Emploi						
<i>Croissance (1 an)</i>	34,4 %	6,6 %	27,7 %	0,000	496	485
<i>Croissance (3 ans)</i>	41,5 %	-3,5 %	45,0 %	0,000	322	337
<i>Croissance (5 ans)</i>	50,6 %	3,6 %	47,0 %	0,000	193	225
Revenus totaux						
<i>Croissance (1 an)</i>	52,9 %	28,6 %	24,2 %	0,002	470	468
<i>Croissance (3 ans)</i>	93,7 %	36,7 %	57,0 %	0,000	323	333
<i>Croissance (5 ans)</i>	107,3 %	40,3 %	67,0 %	0,001	203	232
Ventes						
<i>Croissance (1 an)</i>	48,7 %	27,2 %	21,6 %	0,022	383	403
<i>Croissance (3 ans)</i>	100,1 %	47,2 %	53,0 %	0,001	262	273
<i>Croissance (5 ans)</i>	137,4 %	56,0 %	81,4 %	0,001	156	181

Les résultats de notre comparaison du taux de croissance entre le groupe de contrôle et le groupe traité indiquent que les entreprises financées par le CR croissent plus rapidement que les entreprises qui ne le sont pas, peu importe la mesure de comparaison ou la période sur laquelle la croissance est étudiée. Selon toutes les mesures de rendement, l'avantage des entreprises financées par le CR est statistiquement significatif à $p = 0,01$, sauf dans un cas, où l'écart est significatif à $p = 0,05$. L'ampleur de la supériorité de croissance estimée pour les entreprises financées par le CR est similaire entre les différentes mesures. Dans les comparaisons du taux de croissance cumulatif sur cinq ans, l'avantage varie entre 47 % et 81,4 %, et l'effet le plus élevé est observé sur les ventes. La cohérence des résultats parmi ces mesures de la croissance renforce la conclusion selon laquelle les entreprises financées par le CR croissent plus rapidement que les entreprises qui ne le sont pas. Les écarts statistiquement significatifs entre les taux de croissance permettent de penser que le lien entre le financement de CR et la croissance n'est pas attribuable à l'inflation. En outre, les résultats liés aux revenus et aux ventes semblent indiquer que les entreprises financées par le CR ne font pas qu'acheter des actifs ou embaucher des employés avec

l'argent reçu : elles réussissent généralement à agrandir et développer leurs activités. En moyenne, les entreprises financées par le CR qui survivent voient leurs ventes augmenter, que ce soit en accroissant leur offre pour satisfaire la demande dans leur marché actuel ou en entrant sur d'autres marchés pour y trouver de nouvelles occasions de ventes.

Les entreprises qui cherchent à innover et à commercialiser de nouveaux produits et services sont une cible de choix pour les investisseurs de CR. Les conseils et mentorat offerts avec les investissements en CR sont souvent dirigés vers ces activités d'innovation. Ainsi, nous examinons la croissance en termes de deux mesures liées à ce processus : les salaires et les dépenses en R-D.

À l'égard des salaires, le développement et la commercialisation de produits exigent souvent une main-d'œuvre très qualifiée. Le recours à de tels employés à forte valeur ajoutée en vue de réussir à commercialiser de nouveaux produits devrait se traduire par une augmentation des salaires dans l'entreprise. Cependant, cet effet pourrait être supplanté par d'autres changements dans le portrait de la main-d'œuvre des entreprises, par exemple, l'importante augmentation générale du nombre d'emplois découlant des stratégies d'expansion des entreprises qu'on peut constater dans le tableau 6.

La comparaison de la croissance des salaires entre le groupe traité et le groupe de contrôle figure au tableau 7. Dans l'année suivant l'injection de fonds, la croissance des salaires dans les entreprises financées par le CR a été de 3,6 % supérieure à la croissance des salaires dans les entreprises qui n'ont pas reçu de CR; sur cinq ans, la croissance des salaires dans les entreprises du groupe traité est plus élevée de 10,1 %. Ces résultats sont statistiquement significatifs à $p = 0,05$ sur cinq ans, et à $p = 0,1$ sur un an. Sur trois ans, l'écart de la croissance n'atteint pas tout à fait un seuil de signification de 0,1. La signification statistique de l'écart sur le long terme (cinq ans) donne à penser que la différence dans la croissance des salaires n'est pas uniquement due à l'injection de nouveaux fonds sous forme de CR, et que l'entreprise met à profit sa main-d'œuvre pour créer de la valeur réelle après avoir reçu du CR.

Tableau 7 – Taux de croissance dans le groupe traité et dans le groupe de contrôle, salaires et dépenses en R-D

	Moyenne (Financée par le CR)	Moyenne (Groupe de contrôle)	Écart	Valeur p de l'écart	Observations (Financée par le CR)	Observations (Groupe de contrôle)
Salaires						
<i>Croissance (1 an)</i>	8,2 %	4,6 %	3,6 %	0,061	496	485
<i>Croissance (3 ans)</i>	16,7 %	11,8 %	4,9 %	0,106	322	337
<i>Croissance (5 ans)</i>	29,4 %	19,3 %	10,1 %	0,029	193	225
Dépenses en R-D						
<i>Croissance (1 an)</i>	25,3 %	9,2 %	16,1 %	0,012	292	294
<i>Croissance (3 ans)</i>	24,9 %	8,9 %	16,0 %	0,230	178	172
<i>Croissance (5 ans)</i>	48,9 %	29,4 %	19,5 %	0,230	99	103

En ce qui concerne la croissance des dépenses en R-D, les entreprises ayant reçu du CR affichent un fort écart positif, statistiquement significatif, d'environ 16 % pour la première année.

Cependant, l'écart de croissance en faveur des entreprises financées par le CR n'est pas statistiquement significatif sur trois ans ni sur cinq ans. Cette absence d'écart statistiquement significatif à long terme pourrait refléter la plus grande importance que joue la R-D aux étapes préliminaires du développement et de la commercialisation d'un produit afin de permettre aux entreprises d'atteindre leurs objectifs de croissance. En particulier, compte tenu des résultats montrant une croissance des ventes supérieure dans les entreprises financées par le CR, ces constatations sur les dépenses en R-D pourraient signifier que le CR accélérerait le processus de commercialisation, sans toutefois nécessairement se traduire par une plus grande innovation à long terme. Cette interprétation est cohérente avec les résultats de Hellmann et Puri (2000) qui ont constaté qu'un financement par le CR est associé à une mise en marché plus rapide des produits, ainsi que ceux d'Engel et Keilbach (2007) qui ont révélé qu'une fois neutralisé le nombre de demandes de brevets à la constitution de l'entreprise, le CR n'a pas d'effet significatif sur les activités de brevetage des entreprises. Il convient de souligner que cette interprétation de nos résultats ne contredit pas nécessairement ceux de Kortum et Lerner (2000) qui ont observé un effet significatif du CR sur le taux de brevetage. Leurs résultats concernaient chaque industrie dans son ensemble, et l'effet du CR sur le taux de brevetage dans une industrie pourrait dépendre d'autres processus, comme l'augmentation des activités entrepreneuriales qu'ont relevée Samila et Sorensen (2011).

Enfin, au tableau 8, nous comparons la rentabilité des entreprises financées par le CR et des entreprises qui ne le sont pas selon deux mesures : la marge bénéficiaire brute et le bénéfice brut par emploi. Nous ne décelons aucun écart statistiquement significatif entre le groupe traité et le groupe de contrôle, peu importe la période. Le fait que le bénéfice brut par emploi est statistiquement identique semble indiquer que, au moins chez les entreprises financées par le CR qui survivent, l'augmentation rapide de la taille a peu d'effets négatifs sur la capacité à générer des profits à long terme. Ce résultat donne des renseignements additionnels sur l'effet de CR sur la rentabilité qui ne contredit pas nécessairement les résultats de Puri et Zarutskie (2009). En particulier, à partir de données américaines, ils avaient estimé que les entreprises financées par le CR croissent plus rapidement selon l'emploi et les ventes mais que leur rentabilité chutait après l'injection des fonds. Cela dit, nos résultats n'offrent aucun élément de preuve convaincant en faveur d'effets positifs du CR sur la rentabilité, et concordent donc avec l'hypothèse dominante dans la littérature selon laquelle le CR a davantage d'effets sur la croissance que sur la rentabilité.

Il faut souligner que ces résultats ne visent pas à alimenter le débat sur le rendement des fonds de CR au Canada. Le rendement d'un investissement en CR dans une entreprise dépend de la valeur de sortie de la dite entreprise, et comme les investisseurs de CR ciblent de petites entreprises en début de croissance, la valeur de sortie dépend probablement plus de l'évolution de la taille de l'entreprise que de l'évolution de sa rentabilité mesurée par rapport aux ventes ou aux nombres d'employés.

Tableau 8 – Taux de croissance dans le groupe traité et dans le groupe de contrôle, mesures de rentabilité

	Moyenne (Financée par le CR)	Moyenne (Groupe de contrôle)	Écart	Valeur <i>p</i> de l'écart	Observations (Financée par le CR)	Observations (Groupe de contrôle)
Marge bénéficiaire brute						
<i>Croissance (1 an)</i>	-1,2 %	-0,5 %	-0,7 %	0,717	341	338
<i>Croissance (3 ans)</i>	0,5 %	6,4 %	-5,9 %	0,211	240	243
<i>Croissance (5 ans)</i>	0,8 %	5,7 %	-5,0 %	0,255	138	159
Bénéfice brut par emploi						
<i>Croissance (1 an)</i>	14,3 %	26,7 %	-12,5 %	0,183	341	333
<i>Croissance (3 ans)</i>	54,6 %	48,1 %	6,4 %	0,668	235	233
<i>Croissance (5 ans)</i>	70,2 %	71,3 %	-1,1 %	0,958	135	152

5. Conclusion

Dans la présente étude, nous comparons les entreprises financées par le CR et celles qui ne le sont pas selon plusieurs mesures de rendement tirées d'un riche ensemble de données qui nous permet de tenir compte de caractéristiques jusque là inexplorées dans les analyses empiriques précédentes. Outre un large éventail de renseignements pertinents sur les employés et sur les aspects financiers des entreprises, nous tenons compte de mesures de l'innovation comme les dépenses en R-D et le soutien direct des gouvernements à l'innovation et à la commercialisation. Même si nous ne parvenons pas à éliminer entièrement (aucune technique économétrique ne peut le faire) le biais de sélection attribuable au fait que les investisseurs de CR choisissent des entreprises à fort potentiel, la variété des caractéristiques prises en considération et l'étendue des estimations permettent de penser que nos résultats offrent des perspectives économiques utiles quant à l'effet potentiel du CR la croissance.

En particulier, nos résultats suggèrent que les activités de soutien du CR mettent l'accent sur l'augmentation de la taille de l'entreprise ayant reçu du financement. Nous comparons les entreprises financées par le CR et celles qui ne le sont pas selon quatre mesures de taille : l'actif total, l'emploi, les revenus et les ventes. Peu importe la mesure, les entreprises financées par le CR affichent une croissance statistiquement plus élevée par rapport aux entreprises qui ne reçoivent pas de CR. Nos résultats montrent également que les entreprises financées par le CR enregistrent un taux supérieur de croissance des salaires, ce qui porte à croire qu'elles créent plus d'emplois à forte valeur ajoutée que les entreprises qui ne reçoivent pas de CR. Nos résultats montrent que les dépenses en R-D croissent plus rapidement dans les entreprises financées par le CR que dans celles qui ne le sont pas, mais cet écart n'est statistiquement significatif que pour l'année suivant immédiatement la première injection de CR.

Enfin, nos résultats indiquent que les entreprises financées par le CR ne sont pas plus rentables que celles non financées par le CR. Cette observation concorde avec les avis dominants sur le milieu à l'égard de l'industrie de CR et avec la littérature qui veulent que les investisseurs de CR attachent plus d'importance à la croissance qu'à la rentabilité.

Dans l'ensemble, cette étude présente des résultats empiriques probants montrant que le financement par le CR est associé à la fois à une forte croissance et à une accélération du processus d'innovation et de commercialisation dans les entreprises.

Bibliographie

ABADIE, A., et G. IMBENS. 2006. « Large Sample Properties of Matching Estimators for Average Treatment Effects », *Econometrica*, vol. 74, n° 1, p. 235-267.

ANGRIST, J., et V. LAVY. 2001. « Does Teacher Training Affect Pupil Learning? Evidence from Matched Comparisons in Jerusalem Public Schools », *Journal of Labor Economics*, vol. 19, n° 2, p. 343-369.

BRANDER, J., R. AMIT, et W. ANTWEILER. 2002. « Venture-Capital Syndication: Improved Venture Selection vs. the Value-Added Hypothesis », *Journal of Economics and Management Strategy*, vol. 11, n° 3, p. 423-452.

BRANDER, J., E. EGAN, et T. HELLMAN. 2010. « Government Sponsored versus Private Venture Capital: Canadian Evidence », dans J. Lerner et A. Schoar (dir.), *International Differences in Entrepreneurship*, University of Chicago Press, p. 275-320. Rapport de conférence du NBER.

BERUBE, C., et P. MOHNEN. 2009. « Are Firms that Receive R&D Subsidies More Innovative? », *Canadian Journal of Economics*, vol. 42, n° 1, p. 206-225.

COCHRAN, W., et D. B. RUBIN. 1973. « Controlling Bias in Observational Studies », *Sankhya: The Indian Journal of Statistics*, vol. 35, p. 417-446.

DA RIN, M., et M. PENAS. 2007. « The Effect of Venture Capital on Innovation Strategies », Document de travail n° w13636 du NBER.

ENGEL, D., et M. KEILBACK. 2007. « Firm Level Implication of Early Stage Venture Capital Investments: an Empirical Investigation », *Journal of Empirical Finance*, vol. 14, p. 150-167.

HECKMAN, J., H. ICHIMURA, et P. TODD. 1997. « Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Program », *Review of Economic Studies*, vol. 64, n° 2, p. 605-654.

HECKMAN, J., H. ICHIMURA, et P. TODD. 1997. « Matching as an Econometric Evaluation Estimator », *Review of Economic Studies*, vol. 65, n° 4, p. 261-294.

HELLMAN, T., et M. PURI. 2000. « The Interaction between Product Market and Financing Strategy: The Role of Venture Capital », *Review of Financial Studies*, vol. 13, p. 959-984.

HELLMAN, T., et M. PURI. 2002. « Venture Capital and the Professionalization of Start-up Firms: Empirical Evidence », *Journal of Finance*, vol. 57, p. 169-197.

JAIN, B. A., et O. KINI. 1995. « Venture Capitalist Participation and the Post-Issue Operating Performance of IPO Firms », *Managerial and Decision Economics*, vol. 16, n° 6, p. 596-606.

KORTUM, S., et J. LERNER. 2000. « Assessing the Contribution of Venture Capital and Innovation: Which is First? », *RAND Journal of Economics*, vol. 31, n° 4, p. 674-692.

LAVY, V. 2009. « Performance Pay and Teachers' Effort, Productivity, and Grading Ethics », *American Economic Review*, vol. 99, n° 5, p. 1979-2011.

- LERNER, J. 2010. *Boulevard of Broken Dreams: Why Public Efforts to Boost Entrepreneurship and Venture Capital Have Failed - and What to Do about It*, Princeton University Press, Princeton (N. J.), « Kauffman Foundation Series on Innovation and Entrepreneurship ».
- PURI, M., et R. ZARUTSKIE. 2009. « On the Lifecycle Dynamics of Venture-Capital- and Non-Venture-Capital-Financed Firms », Document de travail n° w14250 du NBER.
- ROSENBAUM, P., et D. RUBIN. 1983. « The Central Role of the Propensity Score in Observational Studies for Causal Effects », *Biometrika*, vol. 70, n° 1, p. 41-55.
- ROSENBAUM, P., et D. RUBIN. 1985. « Constructing a control group using multivariate matched sampling incorporating the propensity score », *The American Statistician*, vol. 39, p. 33-38.
- RUBIN, D. B. 1976. « Multivariate matching methods that are equal percent bias reducing, I: some examples », *Biometrics*, vol. 32, p. 109-120.
- RUBIN, D. B. 1980. « Bias Reduction using Mahalanobis Metric Matching », *Biometrics*, vol. 36, p. 293-298.
- RUBIN, D. B. 2001. « Using Propensity Scores to Help Design Observational Studies: Application of the Tobacco Litigation », *Health Services & Outcomes Research Methodology*, vol. 2, p. 169-188.
- SAMILA, S., et O. SORENSON. 2011. « Venture Capital, Entrepreneurship, and Economic Growth », *Review of Economics and Statistics*, vol. 93, p. 338-349.
- WOOLDRIDGE, J. M. 2010. *Econometric Analysis of Cross Section and Panel Data, Second Edition*, The MIT Press, Cambridge (Mass.).

Annexe A : Couplage des données

La présente étude utilise les nouvelles données couplant des bases de données administratives de Statistique Canada à des renseignements sur les activités d'investissement en CR au Canada fournis par Thomson Reuters et des renseignements sur les bénéficiaires du PARI. La méthode initiale a consisté à coupler les entreprises de ces différents ensembles à celles du RE en fonction de leur nom et de leur adresse, ce qui a donné des résultats de qualité variable en ce qui a trait à l'exactitude des correspondances. Le taux de couplage et la qualité des couplages pour les données sur le CR et sur le PARI figurent respectivement au tableau A1 et au tableau A2.

Tableau A1 – Couplage des données de Thomson Reuters sur les entreprises financées par le CR aux données du RE

Nom	Adresse	Code postal	Ville	Cote de qualité	Nombre	Pourcentage	Pourcentage cumulatif
Identique	Identique			D1	208	8 %	8 %
	Identique			D2	60	2 %	10 %
		Identique		D3	812	29 %	39 %
			Identique	D4	227	8 %	47 %
Similaire	Identique			M1	32	1 %	48 %
	Identique			M2	69	2 %	51 %
		Identique		M3	96	3 %	54 %
			Identique	M4	86	3 %	58 %
Couplage manuel					463	17 %	74 %
Non couplées					709	26 %	100 %
Total					2 762	100 %	

Dans les données sur les investissements en CR, un certain nombre d'entreprises financées par le CR étaient présumées correspondre à des entreprises du RE, mais leur nom ou leur adresse n'étaient pas exactement les mêmes. Ce n'est pas entièrement surprenant, car les méthodes de collecte de données de Thomson Reuters n'ont pas nécessairement enregistré les entreprises par leur dénomination sociale. En outre, il est possible que certaines de ces entreprises financées par du capital de risque aient changé de nom depuis leur apparition dans chacun des ensembles de données. Dans ces cas, des analystes de Statistique Canada et de la direction générale de la petite entreprise d'Industrie Canada ont couplé les données manuellement, mettant à profit leur expertise et les sources publiques de renseignements sur ces entreprises.

Après le couplage initial, des renseignements tirés d'autres bases de données de Statistique Canada, y compris des renseignements sur les T2, les T4 et les PD7, ont servi à évaluer plus en profondeur la qualité des couplages initiaux. Conformément aux lignes directrices favorisant l'exactitude, les couplages douteux ont été cernés et éliminés en suivant un ensemble de règles prudentielles. En particulier, les couples de données ont été éliminés dans les cas suivants : l'entreprise financée par le CR n'a pas paru au RE dans les deux ans suivant la première injection de CR¹⁰; l'entreprise n'a fait état d'aucune activité économique l'année du financement ou

¹⁰ Les entreprises ayant reçu une première injection de CR avant 1999 ont été éliminées lorsqu'elles ne figuraient pas au RE en 1999 ou en 2000.

l'année suivante¹¹; l'entreprise a déclaré des revenus dépassant 50 millions de dollars l'année de la première injection de capital; les données administratives de Statistique Canada et les données de sortie sur les investissements en CR fournies par Thomson Reuters ne concordent pas. Une fois ces critères appliqués, il reste un échantillon de 1 545 entreprises financées par le CR correctement couplées aux données de Statistique Canada. Le tableau A2 résume l'attrition de l'échantillon attribuable à l'application séquentielle de chacun des critères énoncés ci-dessus.

Tableau A2 – Filtrage et attrition attribuables à l'établissement de la base de données de recherche et du groupe traité

	Filtre	Attrition	Taille de l'échantillon
Entreprises financées par le CR selon Thompson Reuters (1990-2009)			2 762
	Couplages aux données du Registre des entreprises (RE) de Statistique Canada	709	2 053
	Correspondance dans le RE à l'année de l'injection ou à l'année suivante	372	1 681
	Activité économique (T2, T4 ou PD7) à l'année de l'injection ou à l'année suivante	81	1 600
	Entreprises hors normes à l'année de l'injection	11	1 589
Base de données de recherche regroupant 1 545 entreprises financées par le CR (1999-2009)	Analyse des données de sortie	44	1 545
	Première injection de CR (1999-2009)	501	1 044
	Information financière suffisante pour l'étude dans les industries considérées	382	662
Groupe traité des entreprises financées par le CR	Appariement à une entreprise non financée par le CR	118	544

Le processus de couplage des données sur les investissements en CR a été conçu de manière à permettre divers types de recherche et d'analyse. Il y avait 1 545 entreprises financées par capital de risque en activité au Canada de 1999 à 2009, mais beaucoup d'entre elles avaient obtenu ce capital avant le début de la période d'échantillonnage, et ne cadrent donc pas dans l'étude. La base de données de recherche comprend 1 044 entreprises qui ont reçu une première injection de CR au cours de la période d'échantillonnage (de 1999 à 2009).

¹¹ L'activité économique est définie comme la production de formulaires de retenues à la source T2, T4 ou PD7. Les entreprises ayant reçu une première injection de financement par le CR avant 1999 ont été éliminées si elles n'ont fait état d'aucune activité économique en 1999 et en 2000.

Tableau A3 – Couplage des données du PARI à celles du RE

Nom	Ville	Province	Cote de qualité	Nombre	Pourcentage	Pourcentage cumulatif
Identique	Identique		D1	5 187	67 %	67 %
	Identique		D2	961	12 %	79 %
		Identique	D3	6	0 %	80 %
Similiare	Identique		M1	192	2 %	82 %
		Identique	M3	122	0 %	84 %
Non couplées				1 268	15 %	100 %
Total				7 736	100 %	

Les données du PARI contenaient moins de renseignements sur l'adresse d'opération de l'entreprise, aussi nous avons appliqué un critère plus strict pour déterminer les correspondances. Cependant, le couplage des données du PARI a été de qualité exceptionnellement élevée : près de 80 % des données de l'échantillon faisaient état d'un nom identique à celui figurant dans le RE. Cette congruence relativement élevée des données du PARI et de celles de Statistique Canada pourrait s'expliquer par les méthodes respectives de collecte de données. Alors que les entreprises donnent leur dénomination sociale exacte au conseil national de recherches lorsqu'elles présentent une demande de financement dans le cadre du PARI, les données sur les investissements en CR de Thomson sont collectées indirectement auprès des investisseurs de CR et des associés commanditaires. Étant donné la grande qualité du couplage des données de Statistique Canada à celles du PARI, nous n'avons pas cherché à améliorer davantage le couplage.

Annexe B : Suivi de la main-d'œuvre

Comme il a été mentionné à la section 2.2, l'identificateur assigné aux entreprises dans le RE n'est pas conçu pour suivre une entreprise au fil du temps. Notamment, l'identificateur d'une entreprise peut changer pour des raisons relativement triviales, comme un changement de dénomination sociale. À l'inverse, une entreprise peut conserver le même identificateur du RE même après des changements structurels substantiels, y compris une fusion ou une acquisition.

En vue de limiter les biais potentiels dans les estimations de croissance subséquemment calculées à partir de ces données, les analystes de Statistique Canada ont utilisé une méthode de suivi de la main-d'œuvre en collaboration avec des analystes d'Industrie Canada connaissant bien les entreprises financées par le CR. La procédure consiste à suivre les masses d'employés (identifiés au moyen de leur numéro d'assurance sociale dans les relevés T4) se déplaçant d'un identificateur du RE à l'autre. Selon la nature des relations observées, nous avons modifié les enregistrements correspondant aux deux identificateurs respectifs pour créer un enregistrement longitudinal unique pour chaque entreprise du groupe traité et du groupe de contrôle, en utilisant l'année d'appariement comme point de départ.

Dans la méthode de suivi de la main-d'œuvre, les relations entre identificateurs ne sont prises en considération que lorsque les renseignements des relevés T4 associés à un identificateur commencent à être enregistrés ou cessent de l'être, soit l'apparition de T4 ou la disparition de T4. À l'apparition ou à la disparition de T4, c'est la taille de l'entreprise et la proportion d'employés que l'ancien identificateur et le nouvel identificateur ont en commun pour la détermination de l'existence d'une relation. Les seuils fixés à cet égard sont présentés au tableau B1. Environ le tiers des identificateurs correspondant à des entreprises du groupe traité ou du groupe de contrôle ont été associés à une ou plusieurs relations de suivi de la main-d'œuvre répondant à ces critères.

Tableau B1 – Seuils d'employés en commun aux fins de la détermination des relations de suivi de la main-d'œuvre

	Nombre d'employés visés par l'apparition ou la disparition de T4					
	> 250 employés	> 50 employés	> 15 employés	> 7 employés	> 5 employés	5 employés
Proportion d'employés en commun	25%	30%	50%	50 % en cas d'apparition ou de disparition	70%	100%
				60 % en cas de continuité		

On trouve trois principaux types de relations de suivi de la main-d'œuvre: une disparition menant à une apparition, une disparition menant à une entreprise existante (continuité), et une entreprise existante menant à une apparition. Même si notre méthode ne permet pas de déduire avec certitude les événements précis du cycle de la vie de l'entreprise auxquels correspondent ces relations, elles représentent essentiellement de fausses disparitions (les données d'une entreprise

sont attachées à un nouvel identificateur), des acquisitions et des dessaisissements (entreprises dérivées). Par ailleurs, un identificateur peut être engagé dans des relations multiples, par exemple deux disparitions liées à une apparition correspondent en gros à une fusion. Ces relations peuvent devenir fort complexes lorsqu'un identificateur est engagé dans une variété de relations touchant plusieurs autres identificateurs, une situation qui n'est pas entièrement inusitée parmi les identificateurs correspondant aux entreprises d'une certaine taille.

Dans le cadre de notre étude, nous ne souhaitons pas examiner uniquement la croissance «organique» des entreprises¹². Notre objectif consiste à mesurer le rendement des entreprises financées par le CR, et la croissance au moyen d'une acquisition pourrait s'inscrire dans la stratégie d'expansion de ces entreprises. Or si l'entreprise financée par le CR fait l'objet d'une fusion ou d'une acquisition, cela se traduit sans doute par la sortie du CR, et idéalement par la fin de l'enregistrement longitudinal. Dans le cas des entreprises dérivées, il est difficile de savoir si le nouvel identificateur représente une entité distincte ou un changement dans les pratiques de déclaration de l'entreprise. Par conséquent, nous couplons les enregistrements lorsque le suivi de la main-d'œuvre semble indiquer une fausse disparition, et nous y mettons fin lorsqu'il porte à croire que l'entreprise a été acquise ou a donné naissance à une entreprise dérivée substantielle.

Nous utilisons un seuil de 50 % d'employés en commun pour évaluer toutes les relations de suivi de la main-d'œuvre¹³. En particulier, dans l'éventualité d'une fausse disparition, l'ancien et le nouvel identificateur sont couplés si au moins 50 % des employés du nouvel identificateur travaillaient pour l'ancien auparavant. Dans les cas où il semble que l'entreprise du groupe traité ou du groupe de contrôle en ait acquis une autre, nous mettons fin à l'enregistrement si l'acquéreur fournit moins de 50 % du total des employés de la nouvelle entreprise. Dans les cas de dessaisissement, nous mettons fin à l'enregistrement si l'entreprise dérivée emploie plus de 50 % des employés de l'entreprise du groupe traité ou du groupe de contrôle. Enfin, dans les cas complexes où un identificateur est concerné par de multiples relations, nous appliquons le seuil de 50 % de manière cumulative, c'est-à-dire que nous mettons fin à l'enregistrement si l'ensemble des emplois attribuables à l'acquisition ou au dessaisissement représente plus de 50 % des emplois de l'entreprise du groupe traité ou du groupe de contrôle.

Les données sur les investissements en CR de Thomson comprennent aussi des renseignements sur les sorties, comme les appels publics à l'épargne, les fusions et acquisitions et les fermetures d'entreprise. Ces données sont utilisées en complément des sorties repérées au moyen du suivi de la main-d'œuvre dans le groupe traité.

¹² La croissance «organique», sur le plan de l'emploi, fait généralement référence à la création ou à l'élimination nette d'emploi, indépendamment de la redistribution des postes. Cependant, dans nos données, l'évolution de l'emploi dans une entreprise rend aussi compte de changements non «organiques», comme les fusions, les acquisitions et les dessaisissements. Même si la croissance non «organique» peut se résumer à des déplacements de postes à l'échelle de l'industrie, cet aspect de la croissance témoigne d'une évolution réelle de la capacité de l'entreprise.

¹³ Bien qu'il soit plutôt arbitraire, ce seuil a été choisi après consultation des analystes de Statistique Canada et d'Industrie Canada en vue de confirmer que l'entreprise du groupe traité ou du groupe de contrôle est l'élément dominant dans la relation de suivi de l'emploi. Un seuil dépassant 50 % a été jugé trop strict étant donné le taux d'attrition et les changements normaux au sein des entreprises.

Annexe C : Estimation de la probabilité de recevoir un financement de CR

Comme il a été expliqué à la section 3, nous estimons cette probabilité au moyen de plusieurs caractéristiques des entreprises qui pourraient influencer sur les décisions d'investissement sous forme de CR, et de la valeur au carré de ces covariables, de façon à tenir compte des relations non linéaires. Les résultats de l'estimation figurent au tableau C1. Les scores de propension de nos appariements sont calculés à l'aide de ces estimations de régression logit.

Les coefficients des mesures de la taille de l'entreprise (actif total, emploi, ventes et revenu) portent à croire que la taille de l'entreprise est un facteur qui joue dans l'obtention d'un financement de CR, mais jusqu'à un certain seuil seulement, car les coefficients associés aux valeurs au carré sont tous négatifs. Fait digne de mention, les ventes et le revenu semblent avoir moins d'influence sur le score de propension que les autres mesures de taille, et sont aussi moins significatifs statistiquement. Ce résultat pourrait être dû à un degré de colinéarité élevé entre les mesures de taille ou simplement signifier des étapes différentes du cycle de vie des entreprises lors de l'investissement en CR, certaines pouvant être caractérisées par de faibles ventes.

Comme on est en droit de s'y attendre, étant donné les ratios donnés au tableau 1, les résultats révèlent aussi que le CR choisit généralement de jeunes entreprises. En outre, lorsque l'effet de la taille de l'entreprise est neutralisé, le revenu net a un effet négatif sur la probabilité d'obtenir du financement de CR. De la même manière, les bénéfiques non répartis ont aussi une influence négative sur le score de propension. Ces résultats négatifs associés aux mesures de rentabilité s'expliquent probablement par la tendance du CR à s'intéresser à des entreprises en début de croissance, de même que par un biais de sélection qui fait que les entreprises les plus rentables peuvent s'autofinancer ou trouver plus facilement du financement de types traditionnels.

Pour ce qui est de nos mesures d'innovation des entreprises, nous constatons qu'elles contribuent de façon significative et substantielle à la probabilité d'obtenir un financement de CR. Le fait d'être bénéficiaire du PARI augmente le score de propension d'une entreprise donnée de plus de 300 %. Même si des différences marginales dans les dépenses en R-D n'ont pas une influence particulièrement marquée sur la probabilité d'obtenir du CR, le fait d'effectuer de la R-D (par opposition au fait de ne pas en effectuer) augmente considérablement les chances de recevoir le traitement. Par exemple, parmi les entreprises sans activités de R-D du groupe traité et du groupe de contrôle, le fait de faire passer de 0 \$ à 15 000 \$ les dépenses en R-D (une somme relativement faible) a entraîné une augmentation de 460 % de la probabilité de recevoir du CR.

Tableau C1 – Logit – Pr(première injection de capital de risque = 1)

Variables indépendantes	Coefficient	Erreur-type	Valeur <i>p</i>
Ln actif total	3,527	0,384	0,000
H ventes	0,021	0,038	0,583
Ln emploi	0,538	0,107	0,000
Ln salaires	6,270	1,736	0,000
H bénéfices non répartis	-0,043	0,005	0,000
H revenu	0,018	0,014	0,200
H bénéfice net	-0,029	0,005	0,000
Âge	-0,076	0,012	0,000
H dépenses en R-D	0,270	0,045	0,000
Ln actif total au carré	-0,107	0,013	0,000
H ventes au carré	-0,002	0,003	0,415
Ln emploi au carré	-0,056	0,019	0,003
Ln salaires au carré	-0,282	0,082	0,001
H bénéfices non répartis au carré	-0,008	0,001	0,000
H revenu au carré	-0,009	0,002	0,000
H bénéfice net au carré	0,014	0,002	0,000
Âge au carré	0,001	0,000	0,021
H dépenses en R-D au carré	-0,009	0,003	0,004
Variable binaire PARI	1,154	0,097	0,000
Variable binaire R-D en 2000	1,973	0,253	0,000
Effets de l'industrie	Oui		
Effets de l'année	Oui		
Observations	2 573 663	Pr>chi carré	0,000
Test du rapport des vraisemblances	5 059,62	LL	-3 603,92
Pseudo R2	0,412		

L'échantillon comprend toutes les entreprises ayant émis un T2 et un T4 de 1999 à 2009 et classées sous un code à quatre chiffres du SCIAN dans lequel au moins trois entreprises ont reçu du financement par le CR à un moment ou à un autre entre 1990 et 2009.

Toutes les valeurs sont présentées soit sous la forme d'un logarithme népérien (Ln), soit sous la forme d'un argument sinus hyperbolique (H).

Annexe D : Vérification de la robustesse

Tableau D1 : Écart de taux de croissance entre les entreprises appariées du groupe traité et du groupe de contrôle, paires où chacune a déclaré ses résultats financiers après 1, 3 et 5 ans

	Moyenne (Financée par le CR)	Moyenne (Groupe de contrôle)	Écart	Valeur p de l'écart	Observations (Financée par le CR)	Observations (Groupe de contrôle)
Actif total						
<i>Croissance (1 an)</i>	0,208	0,016	0,192	0,001	455	455
<i>Croissance (3 ans)</i>	0,369	-0,101	0,470	0,000	255	255
<i>Croissance (5 ans)</i>	0,472	0,035	0,437	0,046	133	133
Revenus totaux						
<i>Croissance (1 an)</i>	0,490	0,302	0,189	0,017	415	415
<i>Croissance (3 ans)</i>	0,858	0,418	0,440	0,005	226	226
<i>Croissance (5 ans)</i>	0,967	0,566	0,401	0,149	114	114
Ventes						
<i>Croissance (1 an)</i>	0,478	0,285	0,192	0,062	304	304
<i>Croissance (3 ans)</i>	0,961	0,470	0,492	0,014	159	159
<i>Croissance (5 ans)</i>	1,412	0,589	0,823	0,015	72	72
Emploi						
<i>Croissance (1 an)</i>	0,346	0,067	0,278	0,000	458	458
<i>Croissance (3 ans)</i>	0,436	-0,058	0,494	0,000	233	233
<i>Croissance (5 ans)</i>	0,474	0,052	0,422	0,008	106	106
Salaires						
<i>Croissance (1 an)</i>	0,084	0,053	0,031	0,120	458	458
<i>Croissance (3 ans)</i>	0,165	0,112	0,053	0,155	233	233
<i>Croissance (5 ans)</i>	0,339	0,143	0,196	0,006	106	106
Dépenses en R-D						
<i>Croissance (1 an)</i>	0,245	0,065	0,179	0,009	234	234
<i>Croissance (3 ans)</i>	0,176	-0,007	0,182	0,334	101	101
<i>Croissance (5 ans)</i>	0,367	0,079	0,288	0,348	37	37
Marge bénéficiaire brute						
<i>Croissance (1 an)</i>	-0,005	-0,009	0,004	0,864	236	236
<i>Croissance (3 ans)</i>	0,018	0,112	-0,094	0,237	129	129
<i>Croissance (5 ans)</i>	-0,001	0,049	-0,050	0,437	58	58
Bénéfice brut par emploi						
<i>Croissance (1 an)</i>	0,187	0,276	-0,088	0,409	231	231
<i>Croissance (3 ans)</i>	0,638	0,411	0,227	0,213	120	120
<i>Croissance (5 ans)</i>	0,657	0,794	-0,136	0,660	55	55