

N° 11-633-X au catalogue — N° 063
ISSN 2371-3437
ISBN 978-0-660-99771-1

Études analytiques : méthodes et références

Analyse d'impact quantitative appliquée : un examen des méthodes contemporaines d'évaluation des programmes destinés aux entreprises

par Manassé Drabo, Landry Kuate et Michael Willox

Date de diffusion : le 27 mai 2026



Statistique
Canada

Statistics
Canada

Canada

Comment obtenir d'autres renseignements

Pour toute demande de renseignements au sujet de ce produit ou sur l'ensemble des données et des services de Statistique Canada, visiter notre site Web à www.statcan.gc.ca.

Vous pouvez également communiquer avec nous par :

Courriel à infostats@statcan.gc.ca

Téléphone entre 8 h 30 et 16 h 30 du lundi au vendredi aux numéros suivants :

- | | |
|---|----------------|
| • Service de renseignements statistiques | 1-800-263-1136 |
| • Service national d'appareils de télécommunications pour les malentendants | 1-800-363-7629 |
| • Télécopieur | 1-514-283-9350 |

Normes de service à la clientèle

Statistique Canada s'engage à fournir à ses clients des services rapides, fiables et courtois. À cet égard, notre organisme s'est doté de normes de service à la clientèle que les employés observent. Pour obtenir une copie de ces normes de service, veuillez communiquer avec Statistique Canada au numéro sans frais 1-800-263-1136. Les normes de service sont aussi publiées sur le site www.statcan.gc.ca sous « Contactez-nous » > « [Normes de service à la clientèle](#) ».

Note de reconnaissance

Le succès du système statistique du Canada repose sur un partenariat bien établi entre Statistique Canada et la population du Canada, les entreprises, les administrations et les autres organismes. Sans cette collaboration et cette bonne volonté, il serait impossible de produire des statistiques exactes et actuelles.

Publication autorisée par la ministre responsable de Statistique Canada

© Sa Majesté le Roi du chef du Canada, représenté par la ministre de l'Industrie, 2026

L'utilisation de la présente publication est assujettie aux modalités de l'[entente de licence ouverte](#) de Statistique Canada.

Une [version HTML](#) est aussi disponible.

This publication is also available in English.

Analyse d'impact quantitative appliquée : un examen des méthodes contemporaines d'évaluations des programmes destinés aux entreprises

par

Manassé Drabo, Landry Kuate et Michael Willox

Division de l'analyse et de la modélisation économiques et sociales, Direction des études analytiques et de la modélisation, Statistique Canada

11-633 063

2026003

ISSN 2371-3429

ISBN 978-0-660-99771-1

27 mai 2026

Études analytiques : méthodes et références

Les documents de cette série traitent des méthodes utilisées pour produire des données qui seront employées pour effectuer des études analytiques à Statistique Canada sur l'économie, la santé et la société. Ils visent à fournir aux lecteurs des renseignements sur les méthodes statistiques, les normes et les définitions utilisées pour élaborer des bases de données à des fins de recherche. Tous les documents de cette série ont fait l'objet d'un examen par les pairs et d'une révision institutionnelle, afin de garantir qu'ils sont conformes au mandat de Statistique Canada et qu'ils respectent les normes généralement reconnues régissant les bonnes pratiques professionnelles.

Les documents peuvent être téléchargés gratuitement à l'adresse www.statcan.gc.ca.

Table des matières

Analyse d'impact quantitative appliquée : un examen des méthodes contemporaines d'évaluations des programmes destinés aux entreprises	4
Table des matières.....	2
Résumé.....	3
1. Introduction	4
2. Contexte de l'évaluation du programme et justification méthodologique	4
3. Cadre méthodologique et exigences en matière de données.....	7
3.1 Couplage et nettoyage	8
3.2 Fusion.....	9
3.3 Appariement	9
3.4 Modèles de régression.....	11
4. Limites	15
4.1 Limites de la méthode des doubles différences qui repose sur des effets fixes bidirectionnels	15
4.2 Solution de rechange aux effets fixes bidirectionnels : l'estimateur des doubles différences échelonnées	16
5. Application pratique : évaluation des programmes de soutien.....	17
6. Conclusion	22
Annexe.....	24
A1. Comprendre les données de Statistique Canada.....	24
Registre des entreprises	24
Quelles données les programmes devraient-ils recueillir pour garantir des couplages de qualité avec le registre des entreprises?	24
Fichier de microdonnées longitudinales des comptes nationaux.....	25
Base de données sur la diversité et les compétences.....	25
A2. Renseignements supplémentaires sur les modèles de régression	25
A3. Comparaison des méthodes.....	26
Bibliographie.....	28

Résumé

Le présent article présente un cadre méthodologique pour évaluer les répercussions économiques des programmes publics, appliqué à des organismes comme les sociétés d'État (p. ex. Banque de développement du Canada [BDC], Financement agricole Canada [FAC], et Exportation et développement Canada [EDC]). En utilisant la méthode des doubles différences, les études d'événements et l'analyse de survie, il permet d'estimer les effets quasi causaux sur certains résultats, tels que le chiffre d'affaires, l'emploi et la pérennité de l'entreprise. Le document décrit la théorie, la mise en œuvre et les besoins en données de chaque méthode en soulignant la valeur des ensembles de données couplés de Statistique Canada. Il aborde également des hypothèses clés selon une étude de cas sur la BDC qui illustre l'application pratique. Le travail favorise la collaboration avec Statistique Canada pour l'évaluation des politiques fondée sur des données probantes.

Mots-clés : doubles différences, étude d'événement, analyse de survie, évaluation d'impact, analyse des politiques, rendement des entreprises, efficacité des programmes publics.

1. Introduction

Statistique Canada a une longue histoire de collaboration avec d'autres ministères gouvernementaux et sociétés d'État pour les aider à évaluer l'incidence de leurs politiques et programmes (p. ex. assurance-emploi et compétitivité fiscale) dans le but de fournir aux Canadiens des renseignements impartiaux, précis et pertinents sur les enjeux qui les touchent le plus. Récemment, Statistique Canada a pris en charge un rôle plus important, aidant les organisations canadiennes, principalement les organismes du gouvernement fédéral et les sociétés d'État, à mesurer l'incidence de leurs services sur leurs clients. Le soutien à l'évaluation d'impact en tant que service analytique a connu une croissance rapide au cours de la dernière décennie grâce aux améliorations continues de Statistique Canada dans l'intégration de ses actifs vastes et diversifiés de microdonnées aux renseignements propres à chaque client des organisations partenaires dans un environnement sécurisé qui protège la confidentialité des renseignements des personnes et des organisations.

Le présent article offre un aperçu du processus que Statistique Canada suit pour soutenir les organisations dans le cadre de la réalisation d'analyses d'impact et décrit les méthodes statistiques couramment utilisées. Pour comprendre les types de renseignements que Statistique Canada inclut dans ses rapports, les résultats d'une étude récente réalisée pour la Banque de développement du Canada (BDC), une société d'État financée par le gouvernement fédéral, sont présentés. Les programmes de la BDC fournissent des services financiers principalement aux petites et moyennes entreprises (PME) canadiennes. En général, leurs services visent à accroître les ventes, l'emploi et la pérennité, en particulier lorsque la situation des entreprises est plus difficile, comme pendant la récession de 2008-2009 ou la pandémie de COVID-19.

L'article est structuré de la manière suivante : la section 2 décrit le contexte de l'étude et expose les principaux arguments justifiant le choix de l'approche méthodologique. La section 3 présente le cadre méthodologique et précise les exigences en matière de données, la section 4 aborde les limites, la section 5 applique le cadre aux programmes des sociétés d'État et la section 6 comprend la conclusion.

2. Contexte de l'évaluation du programme et justification méthodologique

Les gouvernements cherchent de plus en plus à évaluer l'efficacité des programmes mis en œuvre par les ministères et les entreprises privées, afin d'assurer l'imputabilité, d'optimiser l'affectation des ressources et d'améliorer les résultats. Toutefois, l'évaluation causale rigoureuse se heurte souvent à des contraintes pratiques, notamment dans les environnements stratégiques réels où la randomisation est impossible, les données sont fragmentées et les interventions sont déployées de manière inégale. De plus, des contraintes techniques et financières empêchent la réalisation d'une évaluation avant la mise en œuvre. Dans de nombreux cas, les programmes sont évalués après leur mise en œuvre, ce qui peut être très difficile (Frenette, Chan et Handler, 2025). Dans ce contexte, des organismes et certains ministères gouvernementaux ont établi une collaboration avec Statistique Canada. En tant qu'organisme national de statistique, Statistique Canada offre un accès sécurisé à des données administratives de haute qualité, à des données longitudinales et à des mesures normalisées qui permettent une analyse économétrique robuste, évolutive et pertinente sur le plan des politiques. Son infrastructure soutient le travail empirique qui assure l'équilibre entre la rigueur méthodologique et la faisabilité opérationnelle.

Lorsqu'on évalue si un programme permet d'améliorer le rendement d'une entreprise, il est important de comprendre la façon dont divers facteurs externes peuvent influencer la relation entre le programme et les changements dans les résultats (p. ex. emploi et chiffre d'affaires). Dans le cadre d'une analyse d'impact, certaines variables, telles que les facteurs de confusion,

les médiateurs et les modérateurs jouent des rôles distincts (figure 1). Les facteurs de confusion sont des variables (p. ex. l'âge de l'entreprise, sa taille et l'expérience en gestion) qui influencent la décision d'une entreprise de participer à un programme (attribution du traitement) et ses résultats finaux. Il est important de tenir compte des facteurs de confusion, car ils peuvent créer l'illusion d'un effet là où il n'en existe pas. Par exemple, une entreprise recevant un prêt peut afficher une croissance de son chiffre d'affaires, mais cette croissance pourrait découler de conditions de marché favorables plutôt que du programme lui-même (c'est-à-dire le prêt, qui est « l'intervention »). Par exemple, lorsque les conditions du marché aident différemment les entreprises récentes et anciennes ou leur nuisent, des facteurs de confusion, comme l'âge de l'entreprise, peuvent masquer la véritable incidence du programme.

Les médiateurs, tout comme les facteurs de confusion, influencent les résultats, mais ils sont mieux décrits comme un résultat du programme qui, à son tour, provoque un changement dans le résultat. Ils n'influencent pas la décision des entreprises de participer au programme qui offre un « traitement », comme un prêt. Par conséquent, on dit qu'ils s'inscrivent dans l'enchaînement de causalité. Par exemple, un prêt peut renforcer la capacité de commercialisation d'une entreprise, ce qui entraîne ensuite une augmentation des ventes; l'investissement dans la commercialisation est le médiateur. Enfin, les modérateurs influencent la force de l'effet sans influencer la décision de participer au programme. Par exemple, un programme peut profiter davantage aux entreprises technologiques qu'aux fabricants, faisant du type d'industrie un modérateur¹.

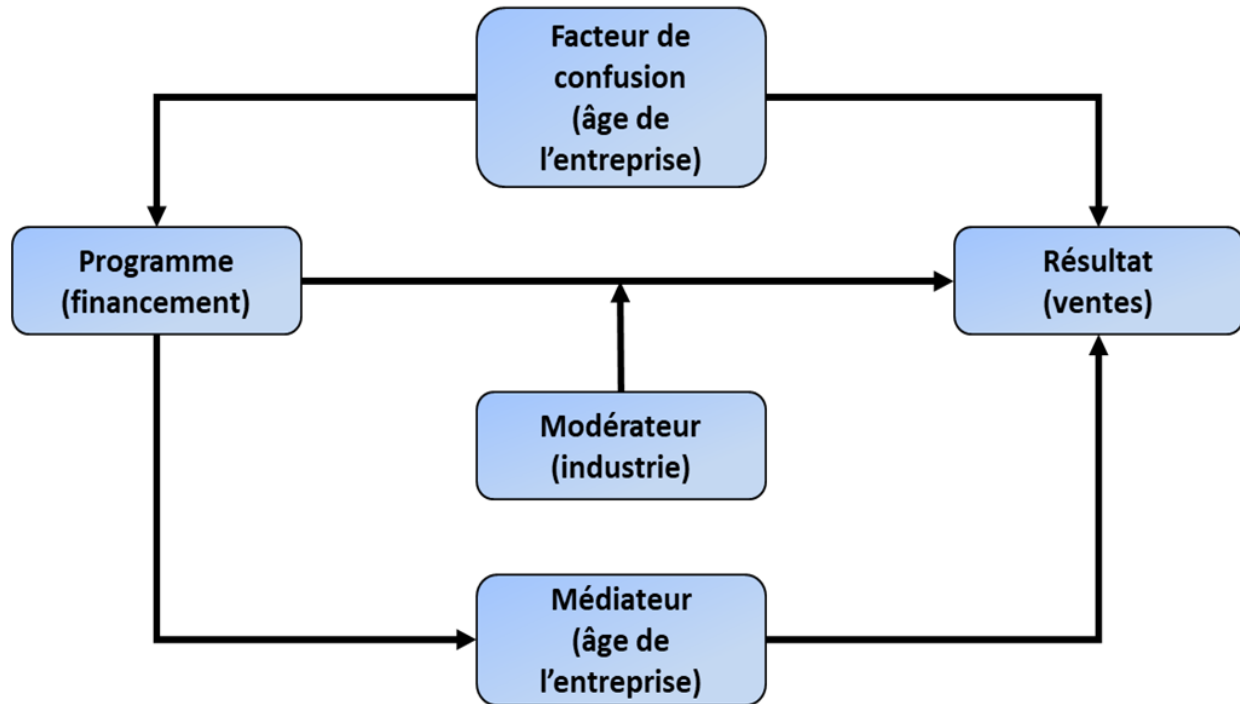
Comprendre ces relations est important, car l'évaluation de l'incidence causale de tels programmes est complexe en raison des facteurs de confusion (p. ex. différences dans les caractéristiques des entreprises, dynamique sectorielle ou tendances macroéconomiques) ce qui peut empêcher de déterminer si les résultats observés découlent des programmes eux-mêmes. Par conséquent, pour fournir des données probantes fiables aux décideurs, aux intervenants et aux organismes de financement, des méthodologies robustes sont essentielles pour isoler l'impact des programmes des autres influences sur les résultats.

Distinguer les facteurs de confusion, les modérateurs et les médiateurs n'est pas qu'un détail technique : c'est essentiel pour comprendre la façon dont les programmes génèrent le changement. Déterminer les facteurs de confusion permet aux analystes d'isoler l'impact du programme des influences externes, tandis que reconnaître les médiateurs aide à révéler les mécanismes au moyen desquels un programme fonctionne. De même, connaître les modérateurs qui touchent positivement ou négativement la force de l'incidence permet aux analystes d'évaluer à qui s'adresse un programme ou dans quelles conditions il est le plus efficace.

1. Sur un plan plus technique, un médiateur se situe sur la voie causale, tandis qu'un modérateur interagit avec le traitement pour modifier l'ampleur de l'effet sur le résultat.

Figure 1

Enchaînements de causalité entre les programmes et les résultats influencés par des facteurs externes : facteurs de confusion, médiateurs et modérateurs avec un exemple illustratif



Source : Calculs de l'auteur.

Ces distinctions permettent d'obtenir des estimations plus précises des effets directs (ceux attribuables uniquement au programme) et des effets indirects (ceux canalisés par l'entremise de variables intermédiaires). De plus, les gestionnaires de programme ne sont pas des acteurs passifs dans ce processus. Leur connaissance approfondie de la manière dont les programmes sont conçus, mis en œuvre et ciblés est essentielle pour déterminer les facteurs d'influence pertinents et pour structurer adéquatement une évaluation de l'incidence. En consignait et en partageant des renseignements sur les critères d'admissibilité, les pratiques de mise en œuvre et les trajectoires de changement prévues, les organismes fédéraux et les organisations renforcent la validité de l'analyse causale et contribuent directement au succès de l'évaluation fondée sur des données probantes.

Bien que les approches qui reposent sur des essais contrôlés randomisés (ECR), des plans expérimentaux de discontinuité de la régression et des variables instrumentales soient considérées parmi les meilleures pour l'inférence causale, elles ne conviennent pas toujours aux évaluations de programmes. Les ECR exigent des affectations aléatoires, ce qui peut être éthiquement ou politiquement inacceptable dans le cadre de la prestation de services publics. Les plans expérimentaux de discontinuité de la régression dépendent de seuils d'admissibilité précis que de nombreux programmes n'ont pas, et les méthodes à variables instrumentales reposent sur la recherche d'instruments valides, qui sont souvent difficiles à trouver dans des contextes institutionnels complexes. De plus, ces méthodes sont généralement axées sur des populations restreintes ou des effets à court terme, ce qui limite leur généralisabilité et leur pertinence pour les politiques. En revanche, les modèles quasi expérimentaux, tels que la méthode des doubles différences, les études d'événements et l'analyse de survie offrent des cadres flexibles pouvant accommoder une mise en œuvre échelonnée, des effets de traitement hétérogènes et des dynamiques à long terme.

Le présent article permet d'examiner ces trois méthodes quantitatives avancées, soit la méthode des doubles différences, l'étude d'événement et l'analyse de survie, pour évaluer et isoler les effets quasi causaux des organismes fédéraux et des organisations en comparant des résultats précis des entreprises clientes qui participent aux programmes des organismes ou qui en bénéficient. Les résultats de ces clients, souvent appelés « les entreprises ayant bénéficié d'un programme » ou membres du « groupe de traitement », sont comparés aux résultats au fil du temps d'un groupe de comparaison composé de non-clients afin de fournir une solution de rechange solide aux statistiques descriptives.

Les objectifs du présent article comporte trois volets. En premier lieu, ils décrivent les fondements théoriques, la mise en œuvre pratique et les exigences en matière de données de ces méthodes de manière claire et accessible. En deuxième lieu, ils démontrent leur applicabilité à un large éventail d'évaluations de politiques au-delà des sociétés d'État. En troisième lieu, ils illustrent leur utilisation dans le cadre de l'évaluation des programmes des sociétés d'État, avec des résultats détaillés pour la BDC qui présentent une croissance importante des ventes et de l'emploi. Ce faisant, le présent article s'appuie sur les travaux récents de Frenette, Chan et Handler (2025) concernant la possibilité d'exploiter les données de Statistique Canada pour l'analyse d'impact et expose les renseignements du cadre méthodologique. En démontrant la faisabilité et la valeur de ces méthodes, l'article vise à promouvoir une culture de responsabilité et d'amélioration continue dans le cadre de la prestation des programmes publics.

3. Cadre méthodologique et exigences en matière de données

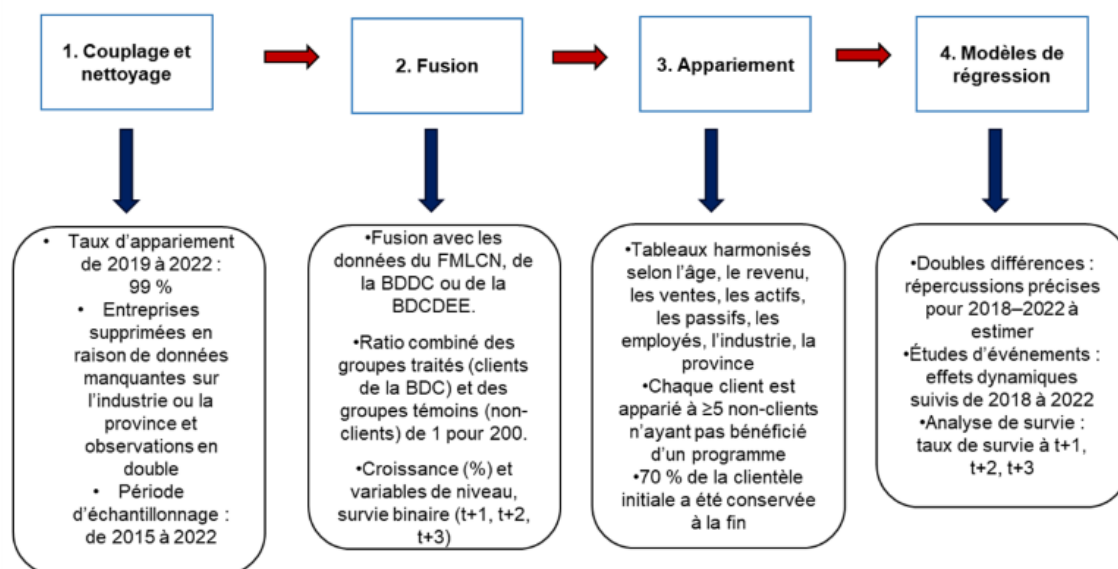
L'évaluation des programmes fédéraux exige des méthodes et des données capables d'isoler les effets quasi causaux des facteurs externes, tels que les caractéristiques des entreprises, les tendances de l'industrie ou les chocs macroéconomiques. La Division de l'analyse et de la modélisation économiques et sociales de Statistique Canada a élaboré un cadre analytique solide qui peut être résumé en quatre étapes, comme il est décrit dans la figure 2.

L'étape 1 consiste à coupler les données du client ou du programme au Registre des entreprises (RE) de Statistique Canada². L'ensemble de données couplé qui en résulte est ensuite nettoyé en supprimant les entreprises dont les renseignements requis pour le couplage sont manquants, et en supprimant les renseignements en double ou les couplages de faible qualité avec le RE³. À l'étape 2, les données couplées de l'étape 1 sont fusionnées avec d'autres ensembles de données d'entreprise de Statistique Canada, tels que le Fichier de microdonnées longitudinales des comptes nationaux (FMLCN) et la Base de données sur la diversité et les compétences (BDDC), des éléments de la Base de données canadienne sur la dynamique employeurs-employés (BDCDEE). Une brève description de ces données est fournie à la section A1 de l'annexe.

2. Le couplage des microdonnées est effectué conformément à la [Directive sur le couplage de microdonnées](#) de Statistique Canada.

3. La qualité du couplage a été évaluée à l'aide d'un système de notation pondéré qui privilégiait la similitude des noms, la cohérence géographique et d'autres identifiants, tels que le numéro de téléphone et l'adresse. Les appariements ont été classés en cinq catégories : A (excellent) pour les correspondances presque parfaites; B (très bon) pour les correspondances solides; C (bon) pour les correspondances modérées; D (acceptable) pour une confiance minimale; E (à utiliser avec prudence) pour les correspondances faibles. Le meilleur appariement pour chaque enregistrement a été signalé, même si d'autres appariements ont été conservés pour examen.

Figure 2
Cadre méthodologique de Statistique Canada pour l'analyse des études d'impact, exemple illustratif



Source : Statistique Canada, Division de l'analyse et de la modélisation économiques et sociales.

La sélection des entreprises qui composent le groupe de comparaison pour l'analyse est effectuée à l'étape 3 en appariant les entreprises ayant bénéficié d'un programme (clients) avec des entreprises non traitées ayant des caractéristiques similaires et qui ne reçoivent pas le produit ou ne participent pas au programme. L'étape 4 de la méthodologie décrit les modèles de régression qui s'appuient sur trois approches économétriques complémentaires (la méthode des doubles différences, l'étude d'événement et l'analyse de survie) pour évaluer l'incidence des services sur les résultats des entreprises, y compris les ventes, l'emploi et la longévité. Des renseignements supplémentaires sur le modèle de régression sont fournis à la section A2 de l'annexe. Le reste de cette section décrit chaque étape en détail.

Des exemples de résultats des programmes de financement de la BDC pour les PME sont utilisés pour illustrer les renseignements et les lignes directrices sur l'interprétation des résultats que les organisations reçoivent généralement dans un rapport d'analyse d'impact. Dans la plupart des cas, les rapports d'analyse d'impact de Statistique Canada visent à fournir aux gestionnaires de programme des outils et des connaissances nécessaires à la compréhension des résultats et à la communication de ces résultats à leurs intervenants, le cas échéant. Cela donne à l'organisme le contrôle sur la manière de diffuser les renseignements contenus dans ces rapports ainsi que les résultats statistiques qui les accompagnent.

3.1 Couplage et nettoyage

L'efficacité de la méthode des doubles différences, de l'étude d'événement et de l'analyse de survie dépend de données longitudinales de haute qualité. La BDCDEE est un ensemble de données administratives exhaustif et couplé qui permet de saisir les relations changeantes entre employeurs et employés dans l'économie canadienne.

Les organismes qui demandent des évaluations d'impact fournissent généralement à Statistique Canada les données des entreprises-clients, ce qui permet d'intégrer ces clients dans la BDCDEE. Les données des clients de l'organisme peuvent également inclure des renseignements sur la nature des services qu'ils ont reçus. Par exemple, l'évaluation de la BDC comprenait des données sur les clients recevant uniquement un financement, ceux recevant uniquement des services consultatifs ou les deux, selon des années de début précises pour

définir les périodes de traitement. Ces données de clients sont couplées au RE, une base de données exhaustive sur les entreprises canadiennes recueillant des données annuelles sur les revenus, l'emploi, l'industrie (à l'aide des codes du Système de classification des industries de l'Amérique du Nord), la taille de l'entreprise (p. ex. nombre d'employés et tranches de revenus) et l'emplacement (p. ex. province et code postal). Une fois les entreprises clientes intégrées dans le RE, leurs renseignements correspondants peuvent être trouvés dans la BDCDEE, comme il est décrit plus en détail à la section 3.2.

Après le couplage, les données sont nettoyées pour éliminer les entreprises dont les renseignements pertinents sont manquants (p. ex. identifiant de l'entreprise, industrie ou province) ainsi que les observations en double.

3.2 Fusion

Pour la méthode des doubles différences et les études d'événements, les organismes ont besoin de données longitudinales sur les résultats pour les groupes ayant bénéficié d'un programme et les groupes témoins, couvrant des indicateurs de rendement clés, tels que les ventes et les valeurs d'exportation (en dollars canadiens) ainsi que l'emploi. À l'aide de l'identifiant d'entreprise qui se trouve dans le RE, Statistique Canada fusionne l'ensemble de données résultant de l'étape 1 avec une ou plusieurs bases de données d'entreprises de la BDCDEE (voir Liu et Zhang, 2025, pour une description détaillée de la BDCDEE et de sa structure). La principale base de données de la BDCDEE pour l'analyse d'impact à Statistique Canada est le FMLCN. Il couvre les entreprises constituées en société ayant des employés, mais exclut les entreprises non constituées en société sans personnel. Pour les analyses comportant des entreprises individuelles, des données provenant du fichier de déclaration financière et de l'annexe 50 de la déclaration T2 peuvent être nécessaires. Une autre base de données importante est la BDDC, qui contient des renseignements démographiques sur les propriétaires d'entreprise et les employés selon les aspects suivants : sexe, âge, état matrimonial, statut d'immigrant, situation vis-à-vis de l'incapacité, l'identité autochtone et le groupe racisé⁴.

Une fois les ensembles de données intégrés, les variables d'intérêt sont créées pour soutenir l'analyse d'impact. Ces variables comprennent des indicateurs de traitement en fonction de la période et de l'intensité de la participation au programme; des covariables au niveau de l'entreprise, telles que la taille, l'industrie et les tendances de croissance; des variables de résultat, telles que la croissance de l'emploi, la survie et les indicateurs de la productivité. Le cas échéant, les caractéristiques de l'entreprise sont également agrégées ou décalées afin de minimiser l'endogénéité. Les variables au niveau de la main-d'œuvre provenant de la BDDC permettent la prise en compte des aspects liés à l'équité et à la diversité, enrichissant ainsi l'interprétation causale des répercussions des programmes.

3.3 Appariement

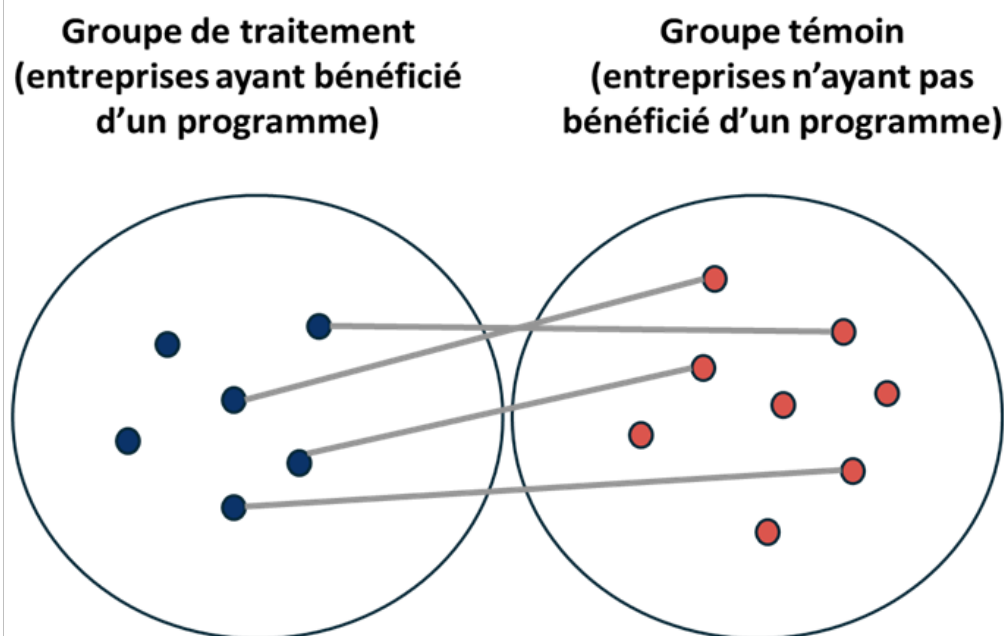
Pour vérifier si un programme (p. ex. prêts ou subventions du gouvernement) permet à une entreprise de croître ou de survivre plus longtemps, il est important de comparer les entreprises participantes ou bénéficiaires avec des entreprises similaires qui n'ont pas reçu ce soutien : un groupe de référence. C'est là que l'appariement entre en jeu. L'appariement consiste à trouver des entreprises non traitées qui ressemblent le plus possible aux entreprises ayant bénéficié d'un programme en fonction de caractéristiques, telles que l'industrie, la taille, le chiffre d'affaires, l'emplacement et le nombre d'employés. En rendant les groupes comparables, on peut réduire l'influence des facteurs externes, tels que l'âge de l'entreprise ou les conditions du marché, qui

4. Pour obtenir de plus amples renseignements sur la manière dont les peuples autochtones et les groupes racisés sont pris en compte dans la BDCDEE, voir Gueye, Lafrance-Cooke et Oyarzun (2022); Gueye (2023); et Gueye (à venir).

pourraient autrement fausser les résultats. Les méthodes d'appariement sont couramment utilisées dans le cadre de l'évaluation des programmes pour constituer des groupes de comparaison qui ressemblent étroitement au groupe recevant l'intervention (Stuart, 2010). Les estimateurs d'appariement incluent l'appariement des scores de propension (ASP), l'appariement du plus proche voisin, l'appariement avec caliper, l'appariement par noyau, entre autres. Les analyses récentes d'études d'impact ont mis l'accent sur l'estimateur de l'appariement par noyau, qui consiste à pondérer les entreprises non traitées en fonction de leur similarité avec les entreprises ayant bénéficié d'un programme à l'aide d'une fonction noyau. L'estimateur d'appariement par noyau repose sur toutes les entreprises non traitées et exige une sélection minutieuse de la largeur de la bande pour son efficacité.

La figure 3 montre un processus d'appariement biunivoque utilisé dans une analyse d'impact, comme l'ASP, pour créer un groupe témoin comparable. Le grand cercle à gauche représente le groupe ayant bénéficié d'un programme (entreprises ayant reçu du soutien dans le cadre du programme ou de l'intervention, représentées par des points bleus), et le grand cercle à droite représente le groupe non traité ou groupe témoin (entreprises n'ayant pas reçu de soutien dans le cadre du programme ou de l'intervention, représentées par des points rouges). Une ligne grise reliant un point bleu à un point rouge indique un appariement réussi, où les deux entreprises présentent des caractéristiques observables très similaires précédant la participation au programme (p. ex. taille, âge et industrie). L'objectif est de s'assurer que la seule différence systématique entre les entreprises appariées est le traitement lui-même, isolant ainsi l'effet causal du traitement. Comme on peut le voir, toutes les entreprises dans l'image ne sont pas appariées. Certaines entreprises ayant bénéficié d'un programme et celles n'ayant pas bénéficié d'un programme demeurent non appariées (points non reliés) et ont été exclues de l'analyse, car aucune contrepartie suffisamment similaire n'a pu être trouvée, ce qui est un résultat courant dans les procédures d'appariement en conditions réelles.

Figure 3
Appariement selon les caractéristiques observables de l'entreprise



Source : Statistique Canada.

Bien que les estimateurs d'appariement puissent améliorer l'inférence quasi causale dans le cadre de l'analyse de la méthode des doubles différences en améliorant l'équilibre des

covariables et en réduisant le biais provenant des facteurs de confusion observables, leur utilisation soulève plusieurs grandes préoccupations qui peuvent l'emporter sur ces avantages. Parmi ces préoccupations figure le risque de violation de l'hypothèse des tendances parallèles : l'appariement harmonise les groupes selon des caractéristiques observées, mais risque de ne pas tenir compte des facteurs non observés et variables au fil du temps qui peuvent entraîner des tendances différentes. De plus, l'appariement réduit souvent la taille de l'échantillon et l'efficacité statistique en éliminant les unités non appariées, et les résultats peuvent être très sensibles au choix des algorithmes d'appariement, ce qui entraîne une instabilité dans les estimations. Ces limites soulignent l'importance de diagnostics rigoureux, de vérifications de sensibilité et de méthodes d'inférence robustes pendant l'intégration de l'appariement dans les cadres de doubles différences.

L'application des estimateurs d'appariement permet la comparaison entre les entreprises appariées ayant bénéficié d'un programme et les entreprises non traitées comparables visées par un appariement subséquent. Cette comparaison aide à déterminer l'effet réel du programme. Pour cette raison, il est important de suivre les entreprises avant et après qu'elles reçoivent du soutien afin de voir la façon dont leurs résultats évoluent au fil du temps. Par exemple, en utilisant les données de la BDCDEE pour suivre l'activité des entreprises, l'analyse de survie vise à déterminer si les entreprises sont toujours actives quelques années après avoir reçu un soutien.

3.4 Modèles de régression

Le présent article porte sur trois méthodes couramment utilisées pour l'analyse d'impact afin de mesurer les effets quasi causaux des programmes des organismes fédéraux et des organisations : méthode des doubles différences, études d'événements et analyse de survie. Chaque approche compare les résultats des entreprises traitées à un groupe soigneusement apparié d'entreprises similaires non traitées, ce qui aide à déterminer l'impact réel du programme. Leur utilisation conjointe de la régression et de l'appariement en fait des outils accessibles et puissants pour l'évaluation des politiques. Un résumé comparatif des méthodes est fourni dans le tableau 1 de la section A3 de l'annexe.

3.4.1 Doubles différences

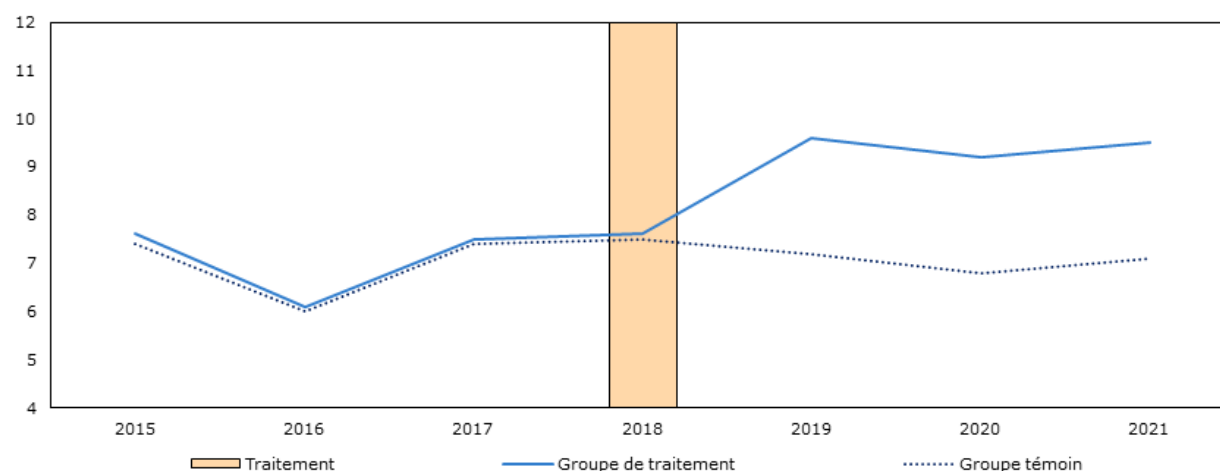
L'une des méthodes les plus efficaces pour évaluer si un programme de soutien aux entreprises, comme un prêt, une subvention ou un service de conseil, entraîne des changements importants dans les résultats, tels que l'emploi ou les ventes, est une méthode connue sous le nom de doubles différences. Cette approche compare l'évolution de ces résultats au fil du temps entre deux groupes : les entreprises ayant bénéficié du programme (groupe de traitement) et des entreprises similaires qui n'en ont pas profité (groupe témoin) (Angrist et Pischke, 2009). La logique est simple : si les deux groupes suivaient des trajectoires similaires avant le programme, toute différence constante qui ressort par la suite peut être attribuée au programme. Le graphique 1 montre un exemple hypothétique des niveaux d'emploi avant et après l'année de traitement (2018). Si les entreprises ayant bénéficié d'un programme commencent à surpasser les entreprises témoins seulement après avoir reçu un soutien, cette divergence peut indiquer l'ampleur de l'effet du programme sur les niveaux d'emploi des clients. Tout au long de l'article, l'impact est appelé l'effet moyen du traitement sur l'entreprise ayant bénéficié du programme plutôt que d'autres estimations de l'effet du traitement, telles que l'intention de traiter ou les effets marginaux du traitement.

Pour que cette comparaison soit valide, le groupe de traitement et le groupe témoin doivent être comparables avant le début du programme. C'est là que l'appariement joue un rôle important. En sélectionnant des entreprises non clientes qui ressemblent étroitement aux entreprises clientes en matière de taille, d'industrie, de région et d'autres caractéristiques, l'analyse garantit que les deux groupes fonctionnaient dans des conditions similaires avant le traitement. Ce processus

aide à éliminer la possibilité que des différences préexistantes, sans lien avec le programme, soient à l'origine des résultats observés.

Graphique 1
Doubles différences pour un programme visant à augmenter l'emploi, exemple illustratif

emploi (pourcentage)



Source : Statistique Canada.

Une hypothèse fondamentale de la méthode des doubles différences est l'hypothèse des tendances parallèles. Cela contribue à affirmer qu'en l'absence du programme, les deux groupes auraient continué à suivre la même trajectoire temporelle qu'avant le traitement. Mettre à l'essai l'hypothèse des tendances parallèles dans un cadre de doubles différences est intrinsèquement limité, car le contrefactuel, ce qui serait arrivé au groupe traité en l'absence de l'intervention, ne peut pas être observé. Bien que l'examen des tendances avant l'intervention entre le groupe ayant bénéficié d'un programme et le groupe témoin puisse fournir des éléments de preuves révélateurs, cela ne constitue pas un essai définitif. Des tendances similaires avant la participation au programme peuvent indiquer une vraisemblance, mais elles ne garantissent pas que ces tendances se seraient poursuivies après le traitement sans intervention.

Cette limite souligne l'importance de disposer d'une période longue et stable avant l'intervention. Plus la similitude des tendances est prolongée et cohérente, plus l'argument est solide en faveur de l'hypothèse selon laquelle les trajectoires parallèles auraient persisté. Toutefois, les chercheurs doivent rester prudents, car même des trajectoires bien alignées pour la période précédant la participation au programme ne peuvent pas valider pleinement l'hypothèse, et toute divergence après la participation au programme pourrait refléter des changements dans les différences sous-jacentes plutôt que des effets de la participation au programme. Par conséquent, bien que l'analyse des tendances préalables soit utile, elle doit être considérée comme une preuve à l'appui, et non comme une preuve déterminante.

Le graphique 1 montre un cas où les tendances précédant la participation au programme sont parallèles en ce qui a trait aux niveaux d'emploi. Dans cet exemple, l'écart dans les niveaux d'emploi qui survient au cours de la période qui suit la participation au programme entre le groupe de traitement et le groupe témoin peut représenter l'incidence du programme ou l'effet moyen du traitement sur l'entreprise ayant bénéficié d'un programme. Cette méthode est particulièrement utile, car elle prend en compte à la fois les tendances temporelles et les différences de base en matière d'emploi entre les entreprises ayant profité d'un programme et celles qui n'en ont pas profité, contrairement aux comparaisons descriptives, qui ne tiennent souvent pas compte de ces facteurs.

En combinant la méthode des doubles différences avec l'appariement et en utilisant des données détaillées au niveau de l'entreprise, telles que les registres de paie ou les déclarations de revenus

des sociétés, il est possible d'isoler l'effet estimé d'un programme avec une plus grande confiance. Cette approche renforce la crédibilité des résultats d'évaluation et fournit aux décideurs des renseignements plus précis sur la manière dont leurs programmes touchent leurs clients.

3.4.2 Études d'événements

Les études d'événements prolongent l'approche de la méthode des doubles différences en montrant la façon dont les effets d'un programme évoluent au fil du temps (Khandker, Koolwal et Samad, 2010). Au lieu de comparer les résultats moyens avant et après un programme, les études d'événements assurent le suivi des changements chaque année avant et après le traitement initial. Cela produit une chronologie détaillée de la manière dont les résultats de l'entreprise, tels que la productivité du travail, les ventes ou l'emploi, réagissent au programme. Le graphique 2 illustre cette approche en traçant la variation moyenne de la productivité du travail pour les entreprises ayant bénéficié d'un programme par rapport à une période de référence. L'axe horizontal montre le temps en termes relatifs, où t est l'année où le client commence à participer au programme. Les années précédant le programme ($t-2$, $t-1$) servent de vérification avant la participation au programme, tandis que les années suivant le programme ($t+1$, $t+2$, $t+3$) révèlent la façon dont les effets de la participation au programme évoluent au fil du temps.

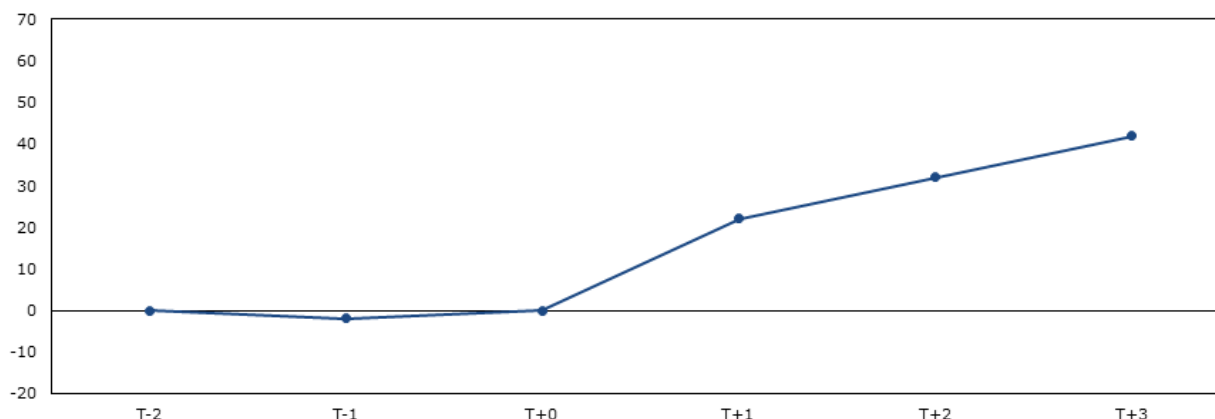
Comme pour la méthode des doubles différences, la crédibilité d'une étude d'événement dépend de la comparaison des entreprises ayant bénéficié d'un programme avec des entreprises non traitées similaires, identifiées grâce au jumelage des caractéristiques, telles que la taille, l'industrie et la région. Cela garantit que les deux groupes ont été soumis à des conditions similaires avant le programme, ce qui permet à l'analyse d'isoler les changements probablement causés par le programme lui-même.

Les études d'événements fournissent également des renseignements précieux sur le moment et la durée des répercussions du programme. Par exemple, le graphique 2 ne montre aucune amélioration mesurable de la productivité au début du programme ($t+0$), mais des gains importants et croissants se manifestent dans les années qui suivent. Cette tendance laisse entendre que les avantages du programme s'accumulent au fil du temps, possiblement à mesure que les entreprises s'adaptent ou réinvestissent. Ces effets dynamiques passeraient inaperçus par des méthodes plus simples, comme les statistiques descriptives, qui ne déclarent généralement qu'un seul résultat moyen. De plus, parce que les études d'événements affichent les résultats pour chaque période, elles aident les analystes à déterminer si les effets s'estompent, se stabilisent ou continuent d'augmenter.

Graphique 2

Étude d'événement mesurant les répercussions des programmes sur les changements de la productivité du travail par rapport à un groupe témoin

changement de la productivité du travail (pourcentage)



Source : Statistique Canada.

La réalisation d'études d'événements nécessite des données couvrant plusieurs années avant et après le début de la participation des entreprises au programme. Bien que cette réalisation permette une bonne compréhension des tendances en matière d'incidence, elle présente également des défis. Par exemple, certaines entreprises pourraient disparaître de l'ensemble de données au cours des années suivantes, soit parce qu'elles ont fermé, soit parce que leurs données ne sont plus déclarées. Cette attrition peut influencer les résultats si les entreprises qui se retirent de l'ensemble de données sont systématiquement différentes. Pour y remédier, les analystes s'appuient sur de grands échantillons et sur un nettoyage minutieux des données, et peuvent appliquer des techniques comme la pondération par probabilité inverse pour tenir compte des retraits⁵.

3.4.3 Analyse de survie

L'analyse de survie est une autre façon d'évaluer si un programme d'affaires aide les entreprises à réussir au fil du temps, surtout pendant les périodes économiques difficiles. Dans ce contexte, « survie » renvoie simplement à la question de savoir si une entreprise est toujours active un, deux ou trois ans après avoir reçu le soutien du programme. Au lieu d'utiliser des modèles statistiques complexes, cette approche se concentre sur une question plus simple : les entreprises qui faisaient partie du programme étaient-elles plus susceptibles de rester en affaires que des entreprises similaires qui n'en faisaient pas partie?

Pour répondre à cette question, la méthode compare les taux de survie entre les entreprises ayant bénéficié d'un programme (celles qui ont reçu un soutien) et un groupe apparié d'entreprises non traitées similaires en matière de taille, d'industrie, de région et d'autres caractéristiques. L'appariement aide à garantir que les différences dans les taux de survie peuvent être liées au programme lui-même plutôt qu'à d'autres facteurs. Cette approche s'inscrit naturellement dans le même cadre de méthode des doubles différences utilisée pour étudier certains résultats, tels que les ventes ou l'emploi, ce qui permet aux organismes qui connaissent déjà ces méthodes de les mettre plus facilement en œuvre pour la survie des entreprises.

Les entreprises sont considérées comme actives si elles continuent de déclarer des valeurs positives pour le chiffre d'affaires et l'emploi. Si ces renseignements disparaissent, l'entreprise est considérée comme s'étant retirée du marché. La méthode permet ensuite de comparer, entre le groupe ayant bénéficié d'un programme et le groupe témoin, les pourcentages d'entreprises toujours actives après un, deux ou trois ans. Par exemple, si 80 % des entreprises traitées sont toujours actives après deux ans, comparativement à 70 % de leurs pairs appariés, les résultats laissent entendre, mais ne prouvent pas nécessairement que le programme a augmenté le taux de survie de 10 points de pourcentage. Il est important de mentionner que la principale hypothèse dans le cas présent est différente de celle d'une analyse par la méthode des doubles différences. Étant donné que toutes les entreprises commencent dans un seul état (existence), l'hypothèse des tendances parallèles n'est pas applicable. En revanche, l'hypothèse la plus importante est que les raisons pour lesquelles une entreprise se retire de l'ensemble de données ne sont pas liées au programme lui-même, un concept connu sous le nom de « censure non informative ».

Bien que cette méthode ne permette pas de déterminer exactement la durée de vie de chaque entreprise ni les raisons pour lesquelles certaines ferment ou disparaissent plus tôt que d'autres, elle fournit un aperçu clair et facile à communiquer de l'incidence du programme. Cette méthode est particulièrement utile pour évaluer la façon dont les programmes soutiennent la résilience des entreprises en période de crise, notamment en période de récession ou lors de la pandémie de COVID-19. Bien que des modèles de survie plus avancés existent, cette approche simplifiée offre

5. La pondération par probabilité inverse est une méthode utilisée pour tenir compte des différences dans la probabilité de participation des entreprises au programme. Elle s'appuie sur le processus d'appariement en accordant plus de pondération aux entreprises qui se ressemblent, mais qui étaient moins susceptibles de participer au programme, ce qui aide à réduire le biais causé par le retrait des entreprises.

aux organismes un moyen pratique et intuitif d'évaluer si leurs programmes aident les entreprises à rester en affaires.

4. Limites

4.1 Limites de la méthode des doubles différences qui repose sur des effets fixes bidirectionnels

La méthode des doubles différences s'appuie fortement sur l'hypothèse des tendances parallèles selon laquelle le groupe ayant bénéficié d'un programme et le groupe témoin auraient suivi des trajectoires de résultats similaires (p. ex. ventes ou emploi) en l'absence du programme. Toutefois, cette hypothèse ne peut pas être directement mise à l'essai, car on n'observe jamais la tendance contrefactuelle du groupe ayant bénéficié d'un programme. Même si les tendances précédant la participation au programme semblent similaires, cela ne garantit pas qu'elles seraient restées parallèles après le traitement. Si l'hypothèse n'est pas respectée, en raison de facteurs de confusion non observés, de l'anticipation de la politique ou de chocs différenciés, l'estimation par la méthode des doubles différences sera biaisée.

D'autres limites importantes comprennent :

- I. période de traitement hétérogène : lorsque les unités reçoivent le traitement à des moments différents, la méthode normalisée des doubles différences (en particulier les modèles à effets fixes bidirectionnels) peut produire des estimations trompeuses causées par une pondération négative ou une contamination provenant d'unités déjà traitées;
- II. facteurs de confusion variant dans le temps : si les facteurs qui influencent le résultat évoluent différemment selon les groupes, ils peuvent fausser l'estimation de l'effet du traitement;
- III. données limitées pour la période précédant la participation au programme : sans une longue période précédant la participation au programme, il est difficile d'évaluer si les tendances sont réellement parallèles;
- IV. retombées : si le traitement touche les unités témoins (p. ex. par l'entremise des marchés ou des réseaux), le groupe témoin n'est plus une hypothèse valable.

Les études d'événements sont sensibles au biais de survie, où les entreprises qui ferment sont retirées de l'échantillon, ce qui peut gonfler les coefficients des années ultérieures. Les clients ayant des taux de survie plus élevés en raison du soutien reçu grâce au programme pourraient présenter des écarts plus importants dans le chiffre d'affaires dans les années ultérieures, reflétant la survie plutôt que les effets directs du programme. La perte de données, attribuable au retrait des entreprises, réduit la taille de l'échantillon et élargit les intervalles de confiance, en particulier pour le long terme. Elle peut être atténuée grâce à l'utilisation d'un grand échantillon, mais l'attrition demeure une préoccupation. Contrairement aux statistiques descriptives, qui évitent l'attrition en étant axées sur un moment précis dans le temps, les études d'événements démontrent la manière dont l'influence des programmes sur les résultats évolue au fil du temps.

L'analyse de survie fondée sur la méthode des doubles différences, qui repose sur un résultat binaire de survie (en activité ou qui n'existe plus à $t+1$, $t+2$ ou $t+3$), se limite à comparer si les entreprises ayant participé à un programme ont plus de chances de survivre que les entreprises du groupe témoin apparié. Cette approche fournit des estimations de la probabilité de survie qui reposent sur la participation au programme à des moments précis après le traitement. Toutefois, elle ne modélise pas le calendrier complet de retrait ni la probabilité de survie au fil du temps comme le font les modèles de survie traditionnels. Autrement dit, cette approche ne permet pas d'estimer le risque changeant de retrait au fil du temps. Par exemple, elle peut démontrer que les

clients ont une probabilité supérieure de 10 points de pourcentage d'être en activité à un moment donné après la participation initiale au programme, mais ne pas mesurer la durée de survie avant la fermeture ni la probabilité que les entreprises survivent pendant plusieurs années consécutives.

Un facteur important à prendre en considération lors de l'interprétation des résultats des analyses qui reposent sur la méthode des doubles différences et des études d'événements est leur validité externe, c'est-à-dire la mesure dans laquelle les résultats peuvent être généralisés à des entreprises autres que celles qui ont participé au programme. Les plans de recherche, comme la méthode des doubles différences et les études d'événements, reposent sur des méthodes d'appariement pour constituer des groupes témoins, précisément pour estimer l'effet moyen du traitement sur l'entreprise ayant bénéficié du programme. Cette approche améliore la validité interne en comparant les entreprises ayant participé au programme avec un ensemble d'entreprises visiblement similaires qui n'y ont pas participé.

Toutefois, cet accent mis sur les entreprises ayant bénéficié d'un programme limite la généralisabilité des résultats. Les entreprises qui ont participé au programme ne constituent pas, par définition, un échantillon aléatoire de toutes les entreprises; elles peuvent être plus motivées, faire face à des défis uniques ou posséder un potentiel de croissance particulier qui les a rendues admissibles ou susceptibles de présenter une demande. Étant donné que les groupes témoins appariés sont créés pour refléter le groupe qui participe au programme, les résultats représentent l'effet des programmes sur des ensembles particuliers d'entreprises (clients) dans un contexte économique particulier. Par conséquent, il convient de faire preuve de prudence pendant l'extrapolation des estimations de l'effet moyen du traitement sur l'entreprise ayant bénéficié du programme pour prédire les effets sur des entreprises ayant des caractéristiques différentes ou étant dans des contextes économiques différents.

4.2 Solution de rechange aux effets fixes bidirectionnels : l'estimateur des doubles différences échelonnées

L'estimateur de Callaway et Sant'Anna (2021) représente une avancée pour les analyses selon la méthode des doubles différences dans les contextes d'adoption échelonnée, où les unités sont traitées à des moments différents. Au lieu de s'appuyer sur une seule régression à effets fixes bidirectionnels, leur approche décompose le problème en une série de comparaisons plus simples de doubles différences en 2×2 : chaque groupe-temps de l'effet moyen du traitement sur l'entreprise ayant bénéficié du programme (g, t) est estimé séparément pour une cohorte précise (groupe g traité pour la première fois au temps g) et la période t après la participation, reposant uniquement sur des comparaisons valides (généralement des unités qui n'ont jamais été traitées ou des unités qui n'ont pas encore été traitées comme témoins). Ces effets groupe-temps sont ensuite agrégés en paramètres sommaires (effet moyen du traitement sur l'entreprise ayant bénéficié du programme dans l'ensemble, effets dynamiques ou études événementielles ou effets de calendrier) à l'aide de pondérations explicites choisies par le chercheur.

Comparativement aux modèles à effets fixes bidirectionnels traditionnels, cette méthodologie offre plusieurs avantages clés. En premier lieu, elle permet d'éviter les problèmes bien documentés du modèle à effets fixes bidirectionnels dans les modèles échelonnés, tels que la pondération négative des effets de traitement et la contamination par des unités déjà traitées servant de témoins, ce qui peut entraîner des estimations biaisées ou non interprétables même dans le cadre de l'hypothèse des tendances parallèles. En deuxième lieu, elle fournit des paramètres clairs et interprétables qui tiennent directement compte de l'hétérogénéité des effets du traitement entre les cohortes et au fil du temps. En troisième lieu, elle prend en charge les tendances parallèles conditionnelles (après les covariables) grâce à des méthodes flexibles, telles que la pondération par probabilité inverse, l'ajustement par régression ou l'estimation

doublement robuste, et permet l'utilisation d'unités qui ne sont pas encore traitées comme témoins supplémentaires pour une meilleure comparabilité⁶.

Toutefois, cette approche n'est pas sans inconvénient. Elle peut être moins efficace sur le plan statistique que certaines solutions de rechange dans certains contextes, exigeant des échantillons plus importants pour obtenir des estimations précises, en particulier lorsque de nombreuses cohortes existent ou que les périodes avant la participation au programme sont limitées. Les cohortes de petites tailles peuvent entraîner des effets « groupe-temps » omis ou imprécis. De plus, comme toutes les méthodes des doubles différences, elle repose encore sur une hypothèse des tendances parallèles (conditionnelle) qui reste fondamentalement invérifiable, bien que les essais préalables aux tendances soient plus simples.

Cet estimateur devient une norme en économétrie appliquée pour l'évaluation de programmes et est actuellement utilisé dans des projets d'analyse d'impact en cours de Statistique Canada afin de fournir des données probantes causales plus solides dans des contextes de traitement échelonné. Toutefois, cet accent mis sur les entreprises ayant profité du programme limite la généralisabilité des résultats. Les entreprises qui ont participé au programme ne constituent pas, par définition, un échantillon aléatoire de toutes les entreprises; elles peuvent être plus motivées, faire face à des défis uniques ou posséder un potentiel de croissance précis qui les a rendues admissibles ou susceptibles de présenter une demande. Étant donné que les groupes témoins appariés sont créés pour refléter le groupe de traitement, les résultats représentent l'effet des programmes sur des ensembles d'entreprises (clients) dans un contexte économique particulier. Par conséquent, il convient de faire preuve de prudence lors de l'extrapolation des estimations de l'effet moyen du traitement sur l'entreprise ayant profité du programme pour prédire les effets sur des entreprises présentant des caractéristiques différentes ou évoluant dans des contextes économiques différents.

5. Application pratique : évaluation des programmes de soutien

Imaginons un programme qui soutient les PME partout au Canada en offrant une aide financière (subventions ou bourses), des prêts ou des garanties de prêt, ainsi que des conseils ou des séances de formation. L'administrateur du programme souhaite savoir si ce genre de programme aide ces entreprises à prendre de l'expansion, par exemple, en créant plus d'emplois ou en augmentant les ventes. Pour le savoir, on examinerait les entreprises qui ont reçu de l'aide au cours d'une année particulière, disons 2018, et on les suivrait de 2015 (avant que l'aide ne soit offerte) à 2021 (quelques années après avoir reçu de l'aide). Cela nous permettrait de comparer ce qui s'est passé avant, pendant et après l'aide. Comme indiqué dans les sections précédentes, l'analyse permettrait de comparer les entreprises ayant reçu un soutien avec des entreprises similaires n'ayant pas reçu de soutien afin de mettre en évidence les effets réels du programme. On tiendrait également compte de certains facteurs, tels que le moment de la participation au programme (le moment de la participation initiale de l'entreprise), l'âge de l'entreprise, sa taille, son emplacement et son secteur d'activité afin d'éviter d'être induit en erreur par des facteurs externes.

Statistique Canada compare généralement la liste des entreprises aidées par le programme à sa base de données des entreprises canadiennes. Dans la plupart des études, environ 75 à 90 % des entreprises qui reçoivent de l'aide d'un programme sont appariées avec succès à une ou

6. Comme le soulignent Callaway et Sant'Anna (2021), l'utilisation des unités n'ayant pas encore participé à un programme est motivée par l'identification plutôt que par le nombre de contrôles disponibles. Même lorsque de nombreuses unités n'ayant pas participé au programme existent, seules celles qui continuent à ne pas participer au programme au moment t fournissent des contrefactuels valides pour la cohorte ayant participé au programme. Nous clarifions ce point dans le cas présent aux fins d'exhaustivité.

plusieurs entreprises qui n'ont pas participé au programme. Le pourcentage de participants au programme appariés avec succès dépend de la qualité des données pour les entreprises du programme (identification unique, adresse et secteur d'activité) ainsi que des variables d'intérêt. L'appariement est effectué à l'aide d'une méthode d'appariement par noyau, chaque entreprise recevant de l'aide étant appariée avec au moins deux entreprises similaires n'ayant pas reçu d'aide, selon certains critères, tels que l'âge, le revenu, le nombre d'employés et l'emplacement.

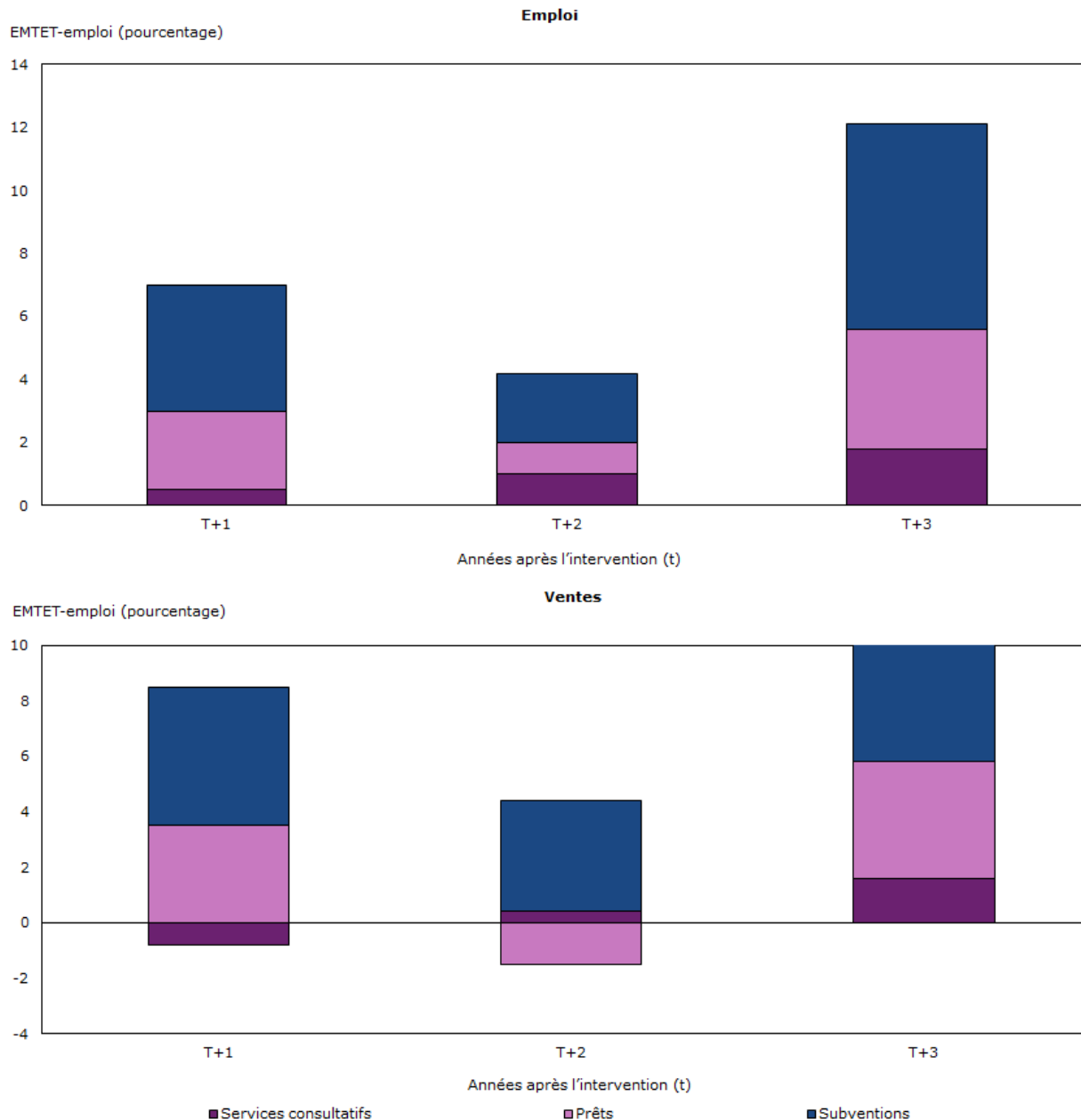
Les estimations de l'effet moyen du traitement sur l'entreprise ayant profité du programme pour différents traitements ne sont pas strictement comparables lorsque les effets du traitement sont estimés séparément pour chaque type de traitement et que l'échantillon apparié d'entreprises inclus dans chacune des analyses n'est pas identique. Par exemple, un groupe d'entreprises ayant choisi de recevoir une subvention pourrait être fondamentalement différent (p. ex. plus grandes ou plus productives) du groupe d'entreprises ayant choisi un prêt ou un service consultatif. Étant donné que chaque effet moyen du traitement sur l'entreprise ayant profité du programme est adapté à un ensemble unique d'entreprises ayant profité d'un programme, comparer directement les changements dans certains résultats, tels que l'emploi ou les ventes, entre les trois traitements ne mesure pas précisément, par exemple, si les subventions sont généralement meilleures que les prêts.

Malgré cette limite technique, comparer les effets moyens estimés du traitement sur les entreprises ayant participé au programme entre les programmes offre toujours une utilité considérable, surtout lorsque les différences dans les échantillons appariés sont relativement faibles. Par exemple, comparer les résultats met en évidence les forces relatives de chaque traitement. Comparer les effets moyens du traitement sur les entreprises ayant participé au programme peut être particulièrement utile pour orienter les projets de recherche à venir, surtout lorsque l'objectif est de mieux comprendre les interactions entre plusieurs traitements⁷. De plus, déterminer les différences dans l'efficacité des programmes sur différentes périodes révèle souvent une hétérogénéité chez les entreprises les plus sensibles à chaque intervention. Bien que ces comparaisons doivent être faites avec prudence, leurs analyses permettent souvent aux administrateurs de programmes d'optimiser la combinaison des soutiens disponibles et de cibler plus efficacement des interventions précises auprès de différents types d'entreprises.

Bien que l'analyse d'impact tende généralement à suivre la méthode décrite dans Callaway et Sant'Anna (2021), ces exemples d'études d'impact menées par la Division de l'analyse et de la modélisation économiques et sociales reposent sur des modèles selon la méthode des doubles différences, d'études d'événements et d'analyse de survie qui incluent une spécification à effets fixes bidirectionnels. Indépendamment de la spécification utilisée, les résultats de l'étude d'impact suivante fournissent un exemple utile de la manière dont les conclusions de l'analyse d'impact sont interprétées et communiquées à un public plus large.

7. Il est fortement recommandé que les organisations qui découvrent l'analyse d'impact quantitative commencent à plus petite échelle en évaluant chaque programme individuellement, comme des traitements distincts avant de s'engager dans des conceptions d'inférence causale plus complexes, comme celles qui estiment les effets des traitements pour les interactions de programmes ou les groupes de programmes.

Graphique 3
Doubles différences de l'emploi des clients en 2020 par rapport aux non-clients



Note : Cette figure montre l'effet moyen du traitement sur l'entreprise ayant bénéficié d'un programme (EMTET) à t+1, t+2 et t+3 pour l'emploi et les ventes, selon le type de service. Les estimations tiennent compte des caractéristiques de l'entreprise et des effets temporels. Les barres sont superposées, et non empilées. Chaque barre représente l'ampleur de l'effet du traitement à partir de zéro, où les barres plus petites sont tracées devant les plus grandes.

Source : Calculs des auteurs.

Le graphique 3 montre qu'en ce qui a trait à l'emploi, les subventions semblent être le traitement le plus efficace, avec des gains clairs pendant la première année et l'incidence la plus importante au cours de la dernière année, les effets passant d'environ quatre points de pourcentage au départ à plus de six points après trois ans. Les prêts sont également liés à des gains en emploi, bien que leur incidence soit plus modérée et s'affaiblit temporairement durant la deuxième année, passant d'environ deux points et demi pendant la première année à moins de quatre

points par la suite⁸. Les services consultatifs n'apportent que des améliorations marginales, généralement de moins de deux points sur l'ensemble de la période.

Le modèle des ventes est similaire, mais avec des effets constamment plus importants liés aux subventions, qui commencent à près de cinq points, diminuent légèrement en 2020 et rebondissent fortement au cours de la troisième année. Les prêts ont fluctué de manière plus marquée, avec une augmentation initiale, suivie d'une baisse temporaire, puis d'une reprise, tandis que les services consultatifs sont restés modestes tout au long de la période, ne dépassant jamais deux points.

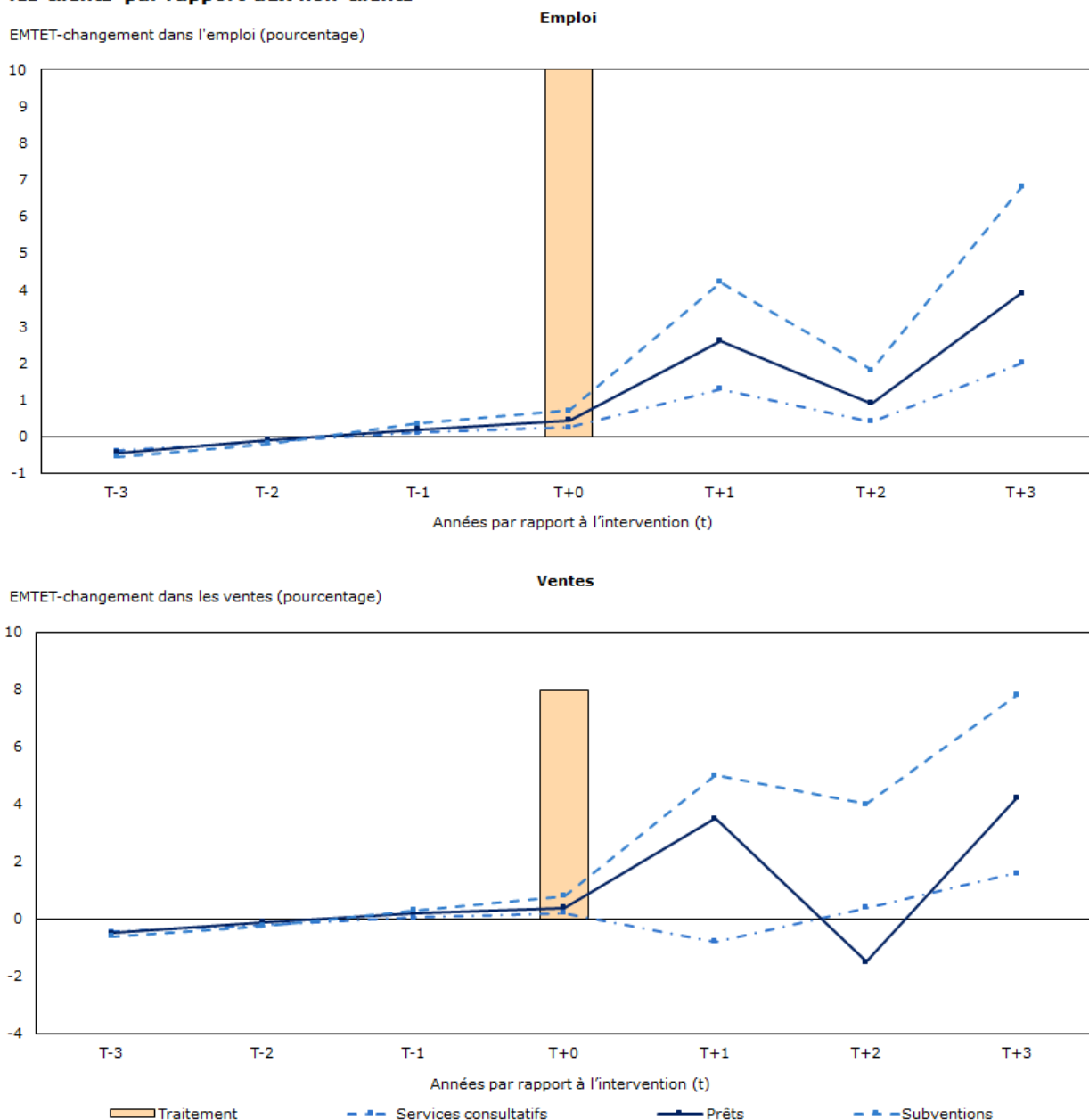
La méthode des doubles différences permet de résumer l'effet moyen sur différentes périodes et est facile à interpréter pour l'élaboration des politiques, mais elle ne révèle pas la rapidité avec laquelle les effets se manifestent, s'ils s'accumulent et la façon dont ils réagissent aux chocs au fil du temps. Le graphique 4 présente les estimations de l'étude d'événement qui répondent à ces questions de délai, tout en fournissant un diagnostic pratique pour l'hypothèse des tendances parallèles.

Les coefficients de prétraitement de t-3 à t+0 restent proches de zéro avec seulement de légères fluctuations, ce qui laisse entendre que les entreprises ayant participé à un programme et les entreprises appariées ont suivi des trajectoires similaires avant l'intervention. Après la participation à un programme, les trajectoires divergent d'une manière qui confirme les tendances observées dans les résultats de la méthode des doubles différences et révèle une dynamique plus riche. En ce qui a trait à l'emploi, les subventions ont l'effet le plus prononcé. L'emploi a augmenté dans les entreprises clientes par rapport au groupe témoin pendant la première année, a diminué en 2020, puis a rebondi au cours de la troisième année, passant d'un peu plus de quatre points de pourcentage à près de sept points de pourcentage de plus. L'emploi dans les entreprises recevant un prêt suit une trajectoire plus modérée, avec des gains au cours de la première année par rapport au groupe témoin, un léger ralentissement au cours de la deuxième année et une reprise au cours de la troisième année. L'emploi dans les entreprises recevant des services consultatifs montre des changements plus modestes, mais positifs tout au long de la période par rapport au groupe témoin, ne dépassant jamais deux points.

En ce qui a trait aux ventes, la tendance générale est similaire, bien que les entreprises ayant reçu un prêt aient connu une baisse plus marquée par rapport au groupe témoin pendant la pandémie avant de repartir à la hausse. Les résultats de cette étude d'événement complètent les estimations de la méthode des doubles différences en confirmant que les tendances antérieures sont probablement parallèles et en illustrant la façon dont les répercussions du programme varient au fil du temps par rapport à la participation initiale, y compris la baisse temporaire pendant la pandémie et la reprise subséquente, en particulier pour les subventions.

8. Cette baisse à t+2 est cohérente avec une hypothèse de réoutillage ou de courbe en J. Si les prêts sont principalement utilisés pour financer des machines et de l'équipement, les perturbations liées à l'installation et aux courbes d'apprentissage peuvent temporairement faire baisser les ventes, tandis que la main-d'œuvre est maintenue ou réaffectée pour la mise en œuvre, suivie d'un rebond à t+3 à mesure que les gains d'efficacité sont réalisés. Mettre à l'essai l'hypothèse de courbe en J exigerait une analyse supplémentaire.

Graphique 4
Étude d'événement mesurant l'incidence des programmes sur les changements dans l'emploi chez les clients par rapport aux non-clients



Note : Cette figure présente les estimations dynamiques des groupes ayant bénéficié d'un programme (EMTET) de t-3 à t+3 pour l'emploi et les ventes. Les coefficients avant le traitement évaluent les tendances parallèles, les trajectoires après le traitement montrent la persistance et la période.

Source : Calculs des auteurs.

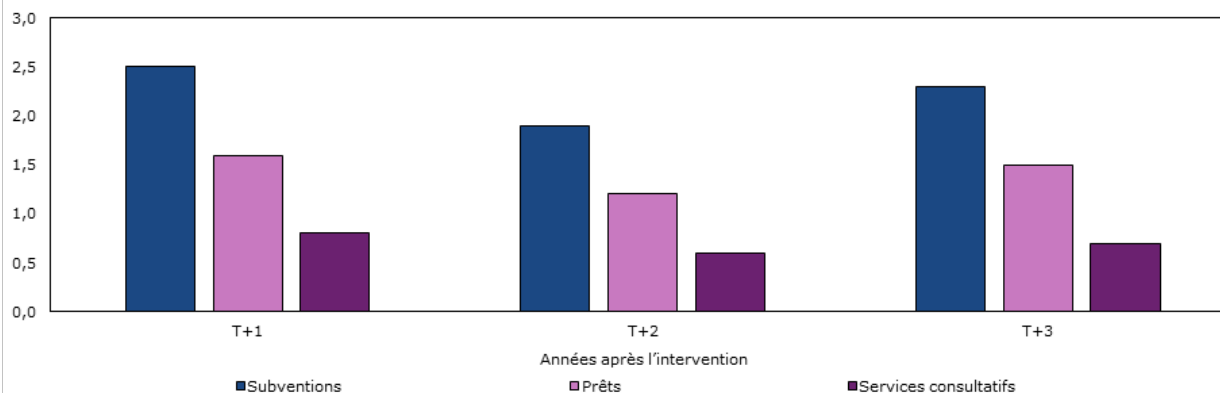
Le graphique 5 présente l'effet moyen du traitement sur l'entreprise ayant bénéficié du programme pour la survie de l'entreprise, qui est définie comme une entreprise restant active et déclarant un taux d'emploi ou un revenu positif dans les données chaque période. Ces estimations de l'effet moyen du traitement sur l'entreprise ayant bénéficié du programme reflètent l'avantage supplémentaire par rapport aux non-clients appariés, ce qui signifie que les valeurs sont modestes et ne doivent pas être confondues avec les taux bruts de survie. Les entreprises recevant des subventions présentent le plus grand avantage supplémentaire, améliorant la survie d'environ deux à trois points de pourcentage sur une période de trois ans, avec une légère baisse en 2020. Les entreprises recevant des prêts présentent des gains plus modestes, mais toujours

considérables, généralement dans une fourchette d'un à un et demi point de pourcentage, tandis que les entreprises profitant de services consultatifs ne présentent que des améliorations marginales par rapport au groupe témoin, généralement inférieures à un point de pourcentage.

Ces tendances reflètent les résultats de l'étude de la méthode des doubles différences et de l'étude d'événement : les effets plus prononcés des subventions en matière de rendement se traduisent par des avantages plus importants en matière de survie des entreprises, les effets des prêts restent positifs, mais sont quelque peu sensibles aux chocs, et l'effet des services consultatifs est plus atténué.

Graphique 5
Analyse de survie (un an plus tard) en matière d'emploi parmi les clients en 2020 par rapport aux non-clients

EMTTET sur le taux de survie (pourcentage)



Note : Cette figure présente un indicateur de survie (entreprise active si l'emploi ou le revenu est positif) pour les clients ayant bénéficié d'un programme (EMTTET) par rapport aux non-clients appariés, selon le type de service, aux temps t+1, t+2 et t+3.

Source : Calculs des auteurs.

Ces estimations sont quasi causales et reposent sur l'hypothèse des tendances parallèles, qui ne peut pas être vérifiée directement. L'appariement réduit le biais lié aux différences observables, mais n'élimine pas les différences non observées dans les données ni l'hétérogénéité latente. L'attrition et les limites des données peuvent également influencer les résultats, en particulier pour les dernières années. Malgré ces mises en garde, l'analyse démontre la valeur des données administratives couplées et des méthodes économétriques robustes pour l'évaluation de programmes fondée sur des données probantes.

6. Conclusion

Les organisations cherchent de plus en plus à évaluer l'efficacité des programmes afin d'assurer la responsabilité, d'optimiser l'affectation des ressources et d'améliorer les résultats. Statistique Canada collabore avec des organisations externes (p. ex. BDC, EDC et FAC) pour appliquer des cadres d'analyse d'impact afin d'évaluer la façon dont leurs services influencent les résultats des clients⁹.

Il est important de mentionner que l'analyse d'impact constitue une entrée à l'évaluation et que l'estimation causale rigoureuse fait souvent face à des contraintes pratiques, notamment lorsque la randomisation n'est pas possible et que les programmes sont évalués après leur mise en œuvre. La présente étude montre la façon dont un cadre d'évaluation bien conçu, combinant la

9. Ratté, S. (2025). [Mesurer l'incidence de BDC sur sa clientèle pendant la pandémie de COVID-19](#). Banque de développement du Canada.

Ratté, S. (2022). [Mesurer l'incidence de BDC sur ses client-es \(2014-2018\) \(ST-CLIENTIMPACT-F2206\)](#). Banque de développement du Canada.

méthode des doubles différences, les études d'évaluations et l'analyse de survie, peut être appliqué pour mesurer l'incidence économique des programmes. Les résultats illustrent la façon dont plusieurs méthodes quantitatives peuvent se compléter. La méthode des doubles différences résume les effets moyens à des périodes clés, les études d'événements valident les hypothèses et révèlent des tendances dynamiques, et l'analyse de survie permet de faire la relation entre les gains de rendement à la continuité des affaires. Ensemble, ils fournissent un complément solide à d'autres approches, telles que les entrevues ciblées ou les études de cas qui aident à isoler les répercussions du programme des facteurs de confusion. Bien que ces modèles ne puissent pas éliminer l'hétérogénéité non observée, ils offrent une base pratique et crédible pour soutenir l'évaluation fondée sur des données probantes.

Auteurs

Manassé Drabo, Landry Kuate et Michael Willox travaillent à la Division de l'analyse et de la modélisation économiques et sociales au sein de la Direction des études analytiques et de la modélisation de Statistique Canada.

Annexe

A1. Comprendre les données de Statistique Canada

Les principaux ensembles de données utilisés pour l'évaluation des programmes du gouvernement fédéral au service des entreprises comprennent les suivants :

- Registre des entreprises (RE) : Répertoire central de renseignements de base de Statistique Canada sur les entreprises et les établissements qui exercent des activités au Canada, continuellement mis à jour.
- Base de données canadienne sur la dynamique employeurs-employés (BDCDEE) : Ensemble de fichiers administratifs couplés et élaborés par la Division de l'analyse microéconomique. Les principaux fichiers utilisés pour l'évaluation d'impact quantitative comprennent :
 - Le Fichier de microdonnées longitudinales des comptes nationaux (FMLCN);
 - La Base de données sur la diversité et les compétences (BDDC).

Registre des entreprises

- Le RE est un répertoire de renseignements reflétant la population des entreprises au Canada et qui sert principalement à fournir des bases de sondage pour toutes les enquêtes économiques de Statistique Canada.
 - Il comprend les sociétés, les entreprises non constituées en société (travailleurs autonomes), les propriétaires uniques, les entités gouvernementales, les organismes à but non lucratif, les partenariats et les fonds financiers.
- Le RE est mis à jour par l'entremise de diverses sources :
 - fichiers T2, T1, TPS, PD7 et T4 de l'Agence du revenu du Canada;
 - données provenant d'autres enquêtes de Statistique Canada ou d'autres sources;
 - données recueillies directement auprès des répondants aux enquêtes.
- Certaines variables utilisées pour les projets proviennent directement du RE, telles que l'industrie (Système de classification des industries de l'Amérique du Nord) et la géographie.

Quelles données les programmes devraient-ils recueillir pour garantir des couplages de qualité avec le registre des entreprises?

- Toutes les données permettant d'identifier de manière unique les entreprises doivent être recueillies.
 - Les noms et adresses seuls n'identifient pas nécessairement une entreprise de manière unique.

- Le numéro d'entreprise est la norme d'excellence pour coupler une entreprise au RE.
- Idéalement, sans numéro d'entreprise, plusieurs noms et adresses, y compris les codes postaux et les provinces ou territoires (p. ex. dénominations sociales, noms commerciaux et adresses), sont nécessaires pour coupler l'entreprise à la bonne entité du RE.
- D'autres variables utiles pour coupler les entreprises à la bonne entité du RE incluent les adresses courriel, les sites Web (URL), les numéros de téléphone et les géocodes, tels que les coordonnées de latitude et de longitude.

Fichier de microdonnées longitudinales des comptes nationaux.

La base de données la plus largement utilisée dans la BDCDEE pour l'analyse d'impact au sein de Statistique Canada est le Fichier de microdonnées longitudinales des comptes nationaux (FMLCN), qui comprend les déclarations de revenus des sociétés (T2) contenant de nombreux renseignements présents dans les états des résultats et les bilans des entreprises; les données de la paie (T4, PD7) liées au taux d'emploi et au revenu du travail; ainsi que les renseignements sur les ventes à partir desquelles la TPS et la TVH sont calculées. Notamment, le FMLCN fournit des renseignements au niveau de l'entreprise et ne couvre pas l'ensemble des entreprises non constituées en société sans employés. Par conséquent, si les clients de l'organisme comprennent des entreprises individuelles, les renseignements provenant du fichier des déclarations financières et de l'annexe 50 du fichier T2 peuvent être nécessaires pour leurs analyses d'études d'impact.

Base de données sur la diversité et les compétences

En s'appuyant sur les renseignements de la BDCDEE, la BDDC a été élaborée pour permettre aux chercheurs d'analyser le rendement des entreprises et d'évaluer les programmes de manière complète. La BDDC couvre la période de 2005 à 2021 (au moment de la rédaction du présent article) et peut fournir des renseignements démographiques sur les propriétaires d'entreprise et les employés selon les éléments suivants :

- sexe;
- âge;
- état matrimonial, statut d'immigrant et situation vis-à-vis de l'incapacité;
- peuples autochtones (propriétaires d'entreprise seulement);
- groupes racisés (propriétaires d'entreprise seulement);
- sexe des cadres supérieurs et des membres du conseil d'administration pour les sociétés cotées en bourse (projet de 2025-2026);
- renseignements sur les entreprises majoritairement détenues (si plus de 50 % de leurs actions appartiennent à des personnes qui s'identifient comme appartenant à un groupe particulier).

A2. Renseignements supplémentaires sur les modèles de régression

Étant donné que la méthode des doubles différences est une méthode fondamentale dans le cadre analytique de l'étude d'impact, la présente section fournit des précisions sur sa mise en œuvre et sa prolongation. L'approche d'étude d'événement généralise la méthode des doubles différences en permettant des effets de traitement dynamiques au fil du temps, tandis que

l'analyse de survie permet d'adapter la stratégie de la méthode des doubles différences aux contextes où le résultat d'intérêt est le taux de survie, le considérant comme la variable dépendante dans l'évaluation des effets de l'intervention.

La méthode des doubles différences est utilisée pour estimer l'effet moyen du traitement sur l'entreprise ayant bénéficié du programme. Les entreprises échantillonnées comprenaient des groupes de traitement et des groupes de comparaison appariés. Si le processus d'appariement est réussi, la seule caractéristique qui différera entre les groupes de traitement et les groupes de comparaison appariés sera le traitement.

Un estimateur de la méthode des doubles différences par la régression linéaire généralisée (équation 1) inclut des indicateurs binaires pour le traitement (T) et une période qui suit la participation au programme (P) afin de saisir les différences entre le groupe de traitement et le groupe témoin ainsi que les covariables X qui peuvent biaiser l'interprétation du paramètre d'interaction TP, qui représente l'effet moyen du traitement sur l'entreprise ayant bénéficié du programme.

$$Y = \beta_0 + \beta_1 T + \beta_2 P + \beta_3 TP + \beta_4 X + \varepsilon \quad (1)$$

Interprétation des coefficients : Les coefficients représentant l'effet du traitement sont interprétés en pourcentage de changement dans le résultat d'intérêt dans les groupes ayant bénéficié d'un programme par rapport aux groupes témoins. Toutefois, lorsque les variables ont été transformées à l'aide des formes logarithmiques ou d'une fonction sinus hyperbolique inverse (arsinh), les coefficients doivent être transformés pour dériver le changement en pourcentage en utilisant la formule suivante¹⁰ :

$$\% \Delta = (\exp(\beta) - 1) \times 100 \quad (2)$$

Ici, %Δ représente le coefficient β en tant que variation en pourcentage où β est la mesure de l'effet moyen du traitement sur l'entreprise ayant bénéficié du programme. Par exemple, si le coefficient pour le résultat relatif au nombre d'employés est égal à 0,2, la variation en pourcentage serait de 22,1 %, et non de 22 %. Quant aux coefficients à faibles valeurs, la différence entre la variation en pourcentage et le coefficient est faible, mais elle augmente de manière exponentielle en fonction de la valeur du coefficient.

A3. Comparaison des méthodes

Tableau 1

10. Bellemare et Wichman (2020) montrent que la méthode pour calculer le pourcentage de changement dans le résultat d'intérêt est la même, que les variables explicatives dans une régression des doubles différences standard soient transformées à l'aide d'une transformation « arsinh » ou d'une transformation logarithmique.

Caractéristique	Doubles différences	Étude d'événement	Analyse de survie
Question principale à laquelle elle répond	Quelle a été l'incidence globale d'un programme en comparant le changement d'un résultat pour le groupe de traitement au changement pour le groupe témoin?	Comment un résultat a-t-il évolué pendant les périodes juste avant et après l'introduction d'une politique ou d'un événement précis?	Comment un programme a-t-il influencé le moment d'un événement, par exemple, la durée nécessaire pour que quelque chose se produise (p. ex., la réussite ou l'échec d'une entreprise)?
Analogie simple	Vous comparez la croissance des ventes dans un magasin qui a lancé une nouvelle campagne publicitaire à celle d'un magasin similaire qui ne l'a pas fait. Vous examinez la différence dans la croissance de leurs ventes, et non seulement leurs chiffres de ventes	Vous assurez le suivi des bénéfices d'une entreprise pendant plusieurs années <i>avant, pendant et après</i> qu'elle commence à exporter. L'accent est mis sur le profil des variations lié au moment où l'exportation a commencé.	Vous mettez à l'essai deux groupes, les participants au programme et les non-participants, pour déterminer le groupe qui compte le plus d'entreprises actives. L'objectif est de déterminer si un groupe est actif de manière constante pendant plus d'années que l'autre.
Ce qu'elle compare	Le changement dans le résultat avant et après le programme pour le groupe de traitement par rapport au groupe témoin.	Le résultat sur plusieurs périodes avant et après un événement unique pour le groupe qui l'a vécu.	Le temps jusqu'à ce qu'un événement précis se produise pour le groupe de traitement par rapport au groupe témoin.
Principale hypothèse	Tendances parallèles : Vous devez supposer que si le programme n'avait jamais eu lieu, les deux groupes auraient continué à présenter des tendances similaires en matière de résultats.	Aucune anticipation : vous supposez que le résultat ne changerait pas déjà en prévision de l'événement, et qu'aucun autre événement majeur ne s'est produit exactement au même moment.	Censure indépendante : Vous devez supposer que les raisons pour lesquelles vous cessez d'observer un sujet (p. ex. fin de l'étude) ne sont pas liées à l'événement d'intérêt lui-même.
Exemple de cas d'utilisation	Une nouvelle subvention du gouvernement a-t-elle augmenté le revenu annuel moyen des entreprises qui l'ont reçue, comparativement à celles qui ne l'ont pas reçue?	L'annonce d'une nouvelle politique commerciale a-t-elle eu une incidence sur les niveaux d'investissement des entreprises dans les trimestres immédiatement avant et après son entrée en vigueur?	L'introduction d'un programme de mentorat pour les entreprises en démarrage a-t-elle augmenté leurs chances de survie?
Ce à quoi ressemble le résultat	Habituellement, un seul chiffre représente l'incidence estimée du programme (p. ex. « Le programme a permis de faire augmenter les revenus de 10 000 \$ en moyenne »).	Un graphique montrant l'incidence à chaque période entourant l'événement, vous permettant de voir si l'effet a été immédiat, s'est accru au fil du temps ou s'est estompé.	Un graphique selon une courbe de survie montrant le pourcentage d'entreprises toujours actives au fil du temps. Vous comparez la courbe du groupe traité avec celle du groupe témoin.

Bibliographie

Angrist, J. D. et Pischke, J.-S. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press.

Callaway, B. et Sant'Anna, P. H. C. (2021). *Difference-in-Differences with multiple time periods*. *Journal of Econometrics*, 225 (2), 200-230.

Khandker, S. R., Koolwal, G. B. et Samad, H. A. (2010). *Handbook on impact evaluation: Quantitative methods and practices*. Publications de la Banque mondiale.

Frenette, M., Chan, W. et Handler, T. (2025). Exploitation des possibilités d'intégration des données de Statistique Canada pour l'évaluation de programmes. *Rapports économiques et sociaux*, Statistique Canada, produit n° 36-28-0001 au catalogue.

Hamit-Hagggar, M. (2024). Évaluation de l'incidence économique du programme Service de croissance accélérée chez les entreprises participantes, 2017 à 2019. *Rapports sur les projets spéciaux sur les entreprises*, Statistique Canada, produit n° 18-001-X au catalogue.

Rosa, J., Nikzad, R., Demers, F. et Poghosyan, T. (2024). Évaluation de l'incidence des programmes fédéraux de soutien à la croissance et à l'innovation sur le rendement des entreprises au Canada à l'aide du modèle CDM. *Rapports sur les projets spéciaux sur les entreprises*, Statistique Canada, produit n° 18-001-X au catalogue.

Stuart, E. A. (2010). [Matching methods for causal inference: A review and a look forward](#). *Statistical Science*, 25(1), 1-21.