

N° 11F0019M au catalogue — N° 362
ISSN 1205-9161
ISBN 978-0-660-22707-8

Direction des études analytiques : documents de recherche

Les effets du cancer sur l'emploi et les gains des survivants du cancer

par Sung-Hee Jeon

Date de diffusion : le 30 septembre 2014



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 sans frais suivants :

- | | |
|---|----------------|
| • Service de renseignements statistiques | 1-800-263-1136 |
| • Service national d'appareils de télécommunications pour les malentendants | 1-800-363-7629 |
| • Télécopieur | 1-877-287-4369 |

Programme des services de dépôt

- | | |
|-----------------------------|----------------|
| • Service de renseignements | 1-800-635-7943 |
| • Télécopieur | 1-800-565-7757 |

Comment accéder à ce produit

Le produit n° 11F0019M au catalogue est disponible gratuitement sous format électronique. Pour obtenir un exemplaire, il suffit de visiter notre site Web à www.statcan.gc.ca et de parcourir par « Ressource clé » > « Publications ».

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 « À propos de nous » > « Notre organisme » > « Offrir des services aux Canadiens ».

Publication autorisée par le ministre responsable de
Statistique Canada

© Ministre de l'Industrie, 2014

Tous droits réservés. L'utilisation de la présente
publication est assujettie aux modalités de l'entente de
licence ouverte de Statistique Canada (www.statcan.gc.ca/reference/licence-fra.htm).

This publication is also available in English.

Note de reconnaissance

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

Signes conventionnels

Les signes conventionnels suivants sont employés dans les publications de Statistique Canada :

- . indisponible pour toute période de référence
- .. indisponible pour une période de référence précise
- ... n'ayant pas lieu de figurer
- 0 zéro absolu ou valeur arrondie à zéro
- 0^s valeur arrondie à 0 (zéro) là où il y a une distinction importante entre le zéro absolu et la valeur arrondie
- ^p provisoire
- ^r révisé
- x confidentiel en vertu des dispositions de la *Loi sur la statistique*
- E à utiliser avec prudence
- F trop peu fiable pour être publié
- * valeur significativement différente de l'estimation pour la catégorie de référence ($p < 0,05$)

Les effets du cancer sur l'emploi et les gains des survivants du cancer

par

Sung-Hee Jeon
Division de l'analyse sociale et de la modélisation
Statistique Canada

11F0019M N° 362
ISSN 1205-9161
ISBN 978-0-660-22707-8

Septembre 2014

Direction des études analytiques Documents de recherche

La série de documents de recherche de la Direction des études analytiques permet de faire connaître, avant leur publication, les travaux de recherche effectués par le personnel de la Direction des études analytiques, les boursiers invités et les universitaires associés. Cette série a pour but de favoriser la discussion sur divers sujets, notamment le travail, la dynamique des entreprises, les pensions, l'agriculture, la mortalité, la langue, l'immigration, l'informatique statistique et la simulation. Le lecteur est invité à faire part aux auteurs de ses commentaires et suggestions.

Les documents de la série sont distribués aux établissements de recherche et aux bibliothèques spécialisées. On peut accéder gratuitement à ces documents à partir d'Internet, à l'adresse www.statcan.gc.ca.

Comité de révision des publications
Direction des études analytiques, Statistique Canada
Immeuble R.-H.-Coats, 24^e étage
Ottawa (Ontario) K1A 0T6

Table des matières

Résumé	5
Sommaire	6
1 Introduction	8
2 Stratégie empirique et modèle	10
2.1 Facteurs de confusion observables influant sur le cancer et les résultats sur le marché du travail.....	10
2.2 Facteurs de confusion non observables influant sur le cancer et les résultats sur le marché du travail.....	12
2.3 Spécification du modèle	12
3 Données	13
3.1 Échantillon de survivants du cancer	14
3.2 Échantillon de contrôle.....	15
3.3 Mesures des résultats sur le marché du travail	19
4 Résultats	20
4.1 Effet du cancer sur l'emploi	20
4.2 Effet du cancer sur les gains	21
4.3 Association de la méthode de la différence dans les différences avec l'appariement ..	22
4.4 Hétérogénéité de l'effet du cancer.....	23
5 Conclusion	29
6 Annexe	30
Bibliographie	35

Résumé

L'étude examine les effets du cancer sur la situation d'emploi et sur les gains annuels des survivants du cancer qui avaient un niveau élevé de participation à la vie active avant leur diagnostic. Le groupe de contrôle est composé de travailleurs semblables qui n'ont jamais reçu de diagnostic de cancer. L'étude est fondée sur un fichier de couplage de Statistique Canada qui regroupe des microdonnées du Recensement de 1991, du Registre canadien du cancer, des dossiers de décès et des dossiers fiscaux des particuliers. L'étude évalue les changements dans l'ampleur des effets du cancer au cours des trois premières années suivant l'année du diagnostic au moyen d'un grand échantillon de survivants du cancer qui ont reçu un diagnostic dans la plage d'âge de 25 à 61 ans. La stratégie empirique combine des modèles d'appariement et des modèles de régression pour traiter les différences observées et non observées entre l'échantillon de survivants du cancer et l'échantillon de contrôle, et pour améliorer l'inférence causale. Les résultats montrent les effets négatifs modérés du cancer sur la situation d'emploi et les gains annuels. Les constatations générales laissent entendre que, à long terme, le cancer est plus susceptible d'influer sur la situation d'emploi des survivants que sur leurs gains.

Mots clés : cancer, emploi, gains annuels

Classification du JEL : I1, J2, J3

Sommaire

L'étude examine les effets du cancer sur certains résultats sur le marché du travail des survivants du cancer au moyen de données de couplage nouvellement accessibles qui regroupent les microdonnées du Recensement du Canada de 1991 et les dossiers administratifs du Registre canadien du cancer, du registraire de l'état civil et des dossiers longitudinaux de l'impôt sur le revenu des particuliers. Ces données uniques servent à évaluer les effets du cancer sur les résultats sur le marché du travail des survivants du cancer canadiens pour tous les types de cancer. Les résultats sur le marché du travail des survivants du cancer sont comparés à ceux de l'échantillon de contrôle, tiré des mêmes fichiers de données composés de personnes possédant des caractéristiques semblables, mais qui n'ont jamais reçu de diagnostic de cancer. L'étude s'intéresse principalement à deux résultats sur le marché du travail — la situation d'emploi et les gains annuels — concernant les survivants du cancer qui ont reçu un diagnostic dans la plage d'âge de 25 à 61 ans et qui avaient un niveau élevé de participation à la vie active avant le diagnostic. L'étude mesure les changements dans l'ampleur des effets du cancer durant les trois années civiles ($T+1$, $T+2$ et $T+3$) après l'année du diagnostic (T). Elle examine aussi les différences entre les effets du cancer dans les sous-groupes ayant certaines caractéristiques observables et l'effet moyen du cancer pour l'ensemble de l'échantillon des personnes ayant reçu un diagnostic de cancer.

Les résultats de l'étude montrent que les survivants du cancer sont moins susceptibles de travailler comparativement aux personnes qui n'ont jamais reçu de diagnostic de cancer dans les trois périodes après le diagnostic. La probabilité de travailler dans la période $T+1$ est inférieure de 3,0 points de pourcentage, en moyenne, pour le groupe de survivants du cancer par rapport au groupe de contrôle. La différence des probabilités augmente et atteint 3,7 points de pourcentage dans la période $T+2$, et 4,8 points de pourcentage dans la période $T+3$. L'ampleur de l'effet négatif est plus grande pour les types de cancer qui ont un faible taux de survie que pour ceux qui ont un taux élevé de survie. La probabilité de travailler dans la période $T+1$ pour les survivants du cancer dont le taux de survie est faible est inférieure de 11,0 points de pourcentage à la probabilité de travailler dans le groupe de contrôle. Chez les personnes atteintes d'un type de cancer dont le taux de survie est élevé, cette probabilité est inférieure de 2,0 points de pourcentage à celle du groupe de contrôle.

Les gains annuels moyens des survivants du cancer sont aussi inférieurs à ceux des personnes de l'échantillon de contrôle. L'étude a révélé que les survivants du cancer gagnent en moyenne environ 12 % de moins dans la période $T+1$ que leurs homologues qui n'ont jamais reçu de diagnostic de cancer, lorsque nous tenons compte des transitions de l'emploi à une situation de non-emploi en incluant les personnes sans revenu dans l'échantillon. Lorsque nous examinons seulement les personnes qui ont gagné un revenu annuel, l'estimation de la perte de revenu moyenne des survivants du cancer dans la période $T+1$ est de l'ordre de 10 % à 11 %. Dans la deuxième et la troisième année après l'année du diagnostic, l'écart des gains entre les survivants du cancer et le groupe de contrôle se rétrécit, tout particulièrement lorsque les survivants du cancer continuent de travailler après avoir reçu leur diagnostic. Les effets du cancer sur les gains annuels sont plus grands et durent plus longtemps chez les survivants d'un cancer à faible taux de survie, ce qui est conforme aux résultats relatifs à la situation d'emploi.

Les constatations de l'étude laissent aussi entendre que, à long terme, l'effet négatif du cancer sur la situation d'emploi des survivants atteints du cancer à un plus jeune âge que l'âge moyen dans l'échantillon de survivants du cancer, soit 48 ans, est plus faible que l'effet moyen observé pour l'ensemble de l'échantillon de survivants du cancer. À l'inverse, l'ampleur de l'effet négatif sur la situation d'emploi est plus importante chez les survivants du cancer sans diplôme d'études secondaires que l'effet moyen observé dans l'échantillon. Toutefois, de façon plus

générale, les tendances des pertes de revenu après un diagnostic de cancer sont variables parmi les survivants du cancer ayant différentes caractéristiques.

Les résultats sur le marché du travail observés dans les données sont déterminés par l'offre et la demande dans le marché du travail. Toutefois, en raison du manque d'information dans les données, il est impossible d'examiner le rôle que joue chaque côté du marché du travail dans la réduction de la probabilité d'emploi des survivants du cancer par rapport à ceux qui n'ont jamais reçu de diagnostic de cancer.

1 Introduction

En 2012, on prévoyait environ 1 638 910 nouveaux cas de cancer diagnostiqués aux États-Unis, et le décès d'environ 577 190 Américains des suites du cancer cette même année (American Cancer Society, 2012). Au Canada, la Société canadienne du cancer (SCC) estimait qu'il y aurait environ 186 400 nouveaux cas de cancer diagnostiqués et 75 700 décès liés au cancer en 2012 (SCC, 2012). Les données de la SCC montrent que l'incidence du cancer dans la population canadienne s'est accrue de façon constante de 1983 à 2007 — en grande partie en raison de la croissance démographique et du vieillissement — même si les taux normalisés selon l'âge sont demeurés relativement stables pendant cette période. En revanche, les taux de mortalité attribuable au cancer parmi les hommes du Canada ont reculé lentement depuis la fin des années 1980 et, parmi les femmes, depuis le milieu des années 1990¹ (SCC, 2012). La plupart des pays membres de l'Organisation de coopération et de développement économiques (OCDE) ont connu un recul des taux de mortalité attribuable au cancer depuis 1980 (OCDE, 2007). Des diagnostics plus rapides, des améliorations apportées au traitement du cancer et de meilleurs soins de suivi ont considérablement augmenté le nombre de survivants du cancer. Le taux relatif de survie après cinq ans² pour tous les types de cancer au Canada a augmenté, passant de 56 % en 1992-1994 à 62 % en 2004-2006, selon les estimations de la SCC pour 2004-2006. De la même façon, le taux de survie après cinq ans aux États-Unis pour tous les types de cancer diagnostiqués dans la période de 2001 à 2007 était de 67 %, en hausse par rapport à 49 % en 1975-1977, et le nombre de survivants du cancer a atteint près de 12 millions en 2007 (American Cancer Society, 2012; Centers for Disease Control and Prevention, 2011).

La présente étude examine les effets du cancer sur les résultats sur le marché du travail des survivants du cancer. Elle utilise des données de couplage nouvellement accessibles qui combinent les microdonnées du Recensement du Canada de 1991 et les dossiers administratifs du Registre canadien du cancer et du registraire de l'état civil, et des dossiers longitudinaux de l'impôt sur le revenu des particuliers. Ces données uniques servent à évaluer les effets du cancer sur les résultats sur le marché du travail des survivants du cancer canadiens pour tous les types de cancer. Les résultats sur le marché du travail des survivants du cancer sont comparés à ceux de l'échantillon de contrôle, composé de personnes n'ayant jamais reçu de diagnostic de cancer et tiré des mêmes fichiers de données. L'étude s'intéresse particulièrement aux résultats des survivants du cancer en âge de travailler qui avaient un niveau élevé de participation à la vie active avant leur diagnostic et aux effets du cancer sur les résultats sur le marché du travail des survivants environ une à quatre années après le diagnostic de cancer. L'étude mesure les changements dans l'ampleur des effets négatifs sur le marché du travail, associés au cancer, pour les trois années suivant l'année civile du diagnostic. L'étude fait aussi l'examen des différences entre les effets du cancer dans les sous-groupes ayant certaines caractéristiques observables et l'effet moyen du cancer dans l'ensemble de l'échantillon des survivants du cancer.

Alors que le taux de survie au cancer augmente, les chercheurs tentent de plus en plus de comprendre l'effet du cancer sur les résultats économiques des survivants du cancer. Jusqu'à aujourd'hui, le cancer du sein a été le plus souvent traité dans ces publications, d'une part parce qu'il est le type de cancer le plus courant chez les personnes en âge de travailler et, d'autre part, parce que les personnes qui reçoivent un diagnostic de cancer du sein ont un taux de survie supérieur comparativement à celles qui ont d'autres types de cancer. Dans les études

-
1. Les taux de mortalité attribuable au cancer du sein chez les femmes ont reculé depuis le milieu des années 1980 (Société canadienne du cancer, 2012).
 2. Le taux relatif de survie après cinq ans correspond à la probabilité de survie pendant cinq ans après un diagnostic de cancer comparativement à la probabilité de survie pendant cinq ans dans l'ensemble de la population du même âge et du même sexe.

récentes, Bradley et coll. (2005) constatent que les personnes ayant survécu à un cancer du sein qui se trouvent dans la plage d'âge de 30 à 64 ans sont moins susceptibles de travailler peu de temps après le diagnostic (au cours des six prochains mois) comparativement aux autres femmes. Toutefois, lorsque Bradley et coll. examinent une plus longue période après le diagnostic, soit de 12 à 18 mois, ils constatent peu d'indications d'un effet négatif (Bradley et coll., 2007). Moran et coll. (2011), qui ont étudié les effets à long terme (de deux à six ans après le diagnostic) de tous les types de cancer sur l'emploi des travailleurs âgés de 28 à 54 ans, ont constaté que les survivants du cancer ont des taux d'emploi plus faibles que les autres adultes du même âge.

Une grande partie de la documentation sur les résultats économiques des survivants du cancer est assez récente, et les données présentées jusqu'ici sont variables, particulièrement en ce qui concerne les effets à long terme. La rareté des données démographiques qui contiennent le diagnostic de cancer et l'information sur le marché du travail a beaucoup nui aux chercheurs. Les enquêtes-ménages courantes sur la population contenant de l'information sur le marché du travail ne comprennent pas habituellement d'information sur le cancer : même si les survivants du cancer peuvent être déterminés dans ces enquêtes — soit généralement les travailleurs plus âgés — l'échantillon est très petit (Bradley et coll., 2002). D'autres sources de données peuvent contenir de l'information sur le marché du travail parce qu'elles sondent un échantillon de survivants du cancer repérables à l'aide des dossiers des patients atteints d'un cancer, auquel cas l'information sur le marché du travail pour l'échantillon de contrôle doit être obtenue auprès d'autres sources de données (Bradley et coll., 2005; Moran et coll., 2011). Toutefois, dans ces cas, les données sont recueillies dans une zone locale ciblée et pourraient ne pas être représentatives à l'échelon national. Le manque de données longitudinales concernant les résultats sur le marché du travail — particulièrement de données sur les gains — pour toute période prolongée avant et après un diagnostic de cancer constitue un autre problème courant. Les études antérieures qui examinaient les effets à long terme du cancer sur les résultats en matière d'emploi faisaient appel à des questions rétrospectives pour déterminer les antécédents sur le marché du travail avant le diagnostic (p. ex., Moran et coll., 2011); par conséquent, les résultats sont susceptibles de comporter d'importantes erreurs de mesure.

Cette étude examine deux résultats sur le marché du travail : la situation d'emploi (travailleur ou chômeur) et le total des gains annuels. L'un des atouts particuliers de cette étude est que ces variables des résultats sur le marché du travail sont fondées sur des données très précises au niveau de la personne, lesquelles proviennent des dossiers fiscaux du revenu annuel des particuliers pour les périodes avant et après le diagnostic de cancer. Cependant, à la différence de la plupart des études découlant de sondages, cette étude ne peut examiner explicitement l'effet du cancer sur l'offre de main-d'œuvre des survivants du cancer — leur activité sur le marché du travail ou leurs heures de travail — parce que ces variables, généralement accessibles dans les données d'enquête, ne sont pas accessibles dans les données administratives. En théorie, toutefois, un diagnostic de cancer peut influencer sur les décisions d'offre de main-d'œuvre d'une personne et, par conséquent, sur ses résultats sur le marché du travail d'au moins deux façons importantes. Premièrement, le cancer modifie la répartition du temps d'une personne : elle doit consacrer du temps à un traitement contre le cancer et à son rétablissement. Par conséquent, les survivants du cancer peuvent disposer de moins de temps à allouer au travail, et donc réduire leurs heures de travail ou simplement cesser de travailler. Il est toutefois probable que la situation entraîne la diminution temporaire de l'offre de main-d'œuvre d'une personne et, par conséquent, le recul des heures d'emploi ou des gains ne durerait pas longtemps. Deuxièmement, non seulement le cancer et le traitement du cancer réduisent-ils physiquement la quantité de temps dont disposent les survivants du cancer pour travailler durant la journée, mais comme pour tous les chocs au niveau de la santé, le cancer peut aussi réorienter les préférences d'une personne du travail aux loisirs. Dans ce cas, la réduction de l'offre de main-d'œuvre durerait plus longtemps et pourrait devenir permanente.

Cette situation entraînerait un recul constant de l'emploi ou des gains observés dans les données longitudinales³.

Selon les recommandations présentées dans Ho et coll. (2007), cette étude combine l'appariement et les modèles paramétriques pour améliorer les inférences causales des estimations fondées sur des données observationnelles. Les données sont prétraitées au moyen de la méthode d'appariement exacte avec groupement (Coarsened Exact Matching ou CEM) et de la repondération de l'échantillon de contrôle à l'aide des pondérations appariées de la méthode CEM pour que l'échantillon de contrôle soit le plus semblable possible à l'échantillon de survivants du cancer (Iacus et coll., 2012). La régression pondérée tient compte des effets confondants potentiels des caractéristiques observables avant le diagnostic de cancer, comme l'âge et les études, sur l'incidence du cancer et les résultats sur le marché du travail des survivants du cancer. De plus, l'étude tente de contrôler la corrélation potentielle parmi les caractéristiques inobservables, l'incidence du cancer et les résultats sur le marché du travail en combinant l'appariement à la méthode de la « différence dans les différences » (DDD) corrigée par régression (Heckman et coll., 1997, 1998).

Les résultats montrent que le taux d'emploi parmi tous les survivants du cancer était inférieur au taux d'emploi parmi les personnes n'ayant jamais reçu de diagnostic de cancer dans les trois périodes après le diagnostic. Toutefois, l'ampleur estimée de l'écart dans les taux d'emploi est relativement modérée (3 points de pourcentage dans la première année après le diagnostic de cancer). L'ampleur de l'effet négatif est plus importante pour les types de cancer ayant de faibles taux de survie que pour ceux qui ont des taux de survie élevés. Les gains annuels moyens des survivants du cancer sont aussi modérément inférieurs aux gains moyens de l'échantillon de contrôle (12,1 % dans la première année après le diagnostic de cancer). Les gains des survivants du cancer qui sont toujours sur le marché du travail dans la première année suivant l'année du diagnostic sont inférieurs à ceux de leurs homologues de l'échantillon de contrôle. Toutefois, ceux qui travaillaient après le diagnostic n'affichaient aucune différence statistiquement significative dans la deuxième et la troisième années pour ce qui est de la moyenne des gains annuels en comparaison avec les personnes n'ayant jamais reçu de diagnostic de cancer. Dans l'ensemble, les constatations laissent entendre qu'à long terme le cancer est plus susceptible d'influer sur la situation d'emploi que sur les gains des survivants.

Voici le plan du document : la section 2 décrit la stratégie empirique de détermination de l'incidence causale du cancer sur les résultats sur le marché du travail. La section 3 décrit toutes les sources des données et explique la règle de sélection pour les échantillons de survivants du cancer et de contrôle. La section 3 contient aussi des statistiques descriptives pour les échantillons de survivants du cancer et de contrôle, avant et après l'appariement. La section 4 expose les résultats. La conclusion se trouve à la section 5.

2 Stratégie empirique et modèle

2.1 Facteurs de confusion observables influant sur le cancer et les résultats sur le marché du travail

Un premier examen des données révèle que le cancer ne touche pas les personnes de l'échantillon de façon aléatoire. De plus, l'incidence du cancer est corrélée avec certaines caractéristiques observables : une indication est que les caractéristiques moyennes de

3. Les résultats observés sur le marché du travail sont déterminés par l'offre et la demande du marché du travail. Les données ne contiennent aucune information sur la réaction des employeurs au diagnostic de cancer des employés. La situation devient encore plus compliquée si le cancer ou le traitement contre le cancer est associé à une incapacité grave et prolongée limitant les fonctions ou le rendement au travail des survivants du cancer. Toutefois, les incapacités liées au cancer ne sont pas observées dans les données.

l'échantillon des survivants du cancer sont différentes de celles de l'échantillon de contrôle. Le fait de ne pas tenir compte de ces facteurs causerait un biais considérable dans les résultats. (Des statistiques sommaires détaillées des deux échantillons seront présentées dans la prochaine section.) La première préoccupation est que les facteurs qui augmentent les chances d'avoir le cancer ont aussi un impact direct sur les résultats sur le marché du travail. Le vieillissement, par exemple, augmente l'incidence du cancer et diminue en même temps la probabilité de travailler. Pour tenir compte de l'effet confondant des états préalables au cancer sur l'incidence du cancer et les variables des résultats, il est important de veiller à ce que les caractéristiques de l'échantillon de contrôle soient plus semblables à celles de l'échantillon de survivants du cancer.

L'étude applique la méthode CEM pour appairier l'échantillon de survivants du cancer à l'échantillon de contrôle en fonction des caractéristiques observables préalables à la période de cancer (Iacus et coll., 2012). Le CEM est un algorithme d'appariement exact multidimensionnel appliqué aux cellules produites par la division des variables continues en intervalles discrets ou par le regroupement des variables catégoriques en un moins grand nombre de catégories groupées⁴. L'algorithme CEM crée un ensemble de strates à partir des mêmes valeurs groupées de variables d'appariement; il réduit aussi les données appariées aux zones de soutien empirique communes en élaguant les observations non appariées des échantillons de traitement et de contrôle. Pour chaque strate, le CEM produit des pondérations $(n_t / n_c * N_c / N_t)$ ⁵ qui peuvent servir à repondérer les observations dans l'échantillon de contrôle apparié et à équilibrer les répartitions empiriques des variables d'appariement entre les échantillons de survivants du cancer et de contrôle.

Ho et coll. (2007) démontrent que le prétraitement des données brutes à l'aide de procédures d'appariement transforme les modèles paramétriques en un outil beaucoup plus fiable d'analyse empirique des effets causaux. Plus précisément, les estimations des effets causaux sont moins sensibles au choix de la spécification du modèle⁶. L'une des propriétés démontrées de la méthode CEM est qu'elle réduit le degré de dépendance du modèle (Iacus et coll., 2012)⁷. Pour faire une inférence causale, les pondérations d'appariement sont utilisées dans l'analyse de régression des résultats sur le marché du travail. De plus, la méthode de régression contrôle le reste du déséquilibre puisque, selon les groupements, il peut rester un certain déséquilibre dans les données appariées. Les estimations de l'effet du cancer dans le modèle mesurent les effets moyens du cancer sur les variables de résultats des survivants du cancer⁸.

-
4. La méthode CEM domine les autres méthodes d'appariement (p. ex., appariement par scores de propension) utilisées couramment pour réduire le déséquilibre lié aux premiers moments et aux moments supérieurs, aux non-linéarités, aux interactions et aux autres différences de répartition multidimensionnelle entre les groupes de traitement et de contrôle. Voir Iacus et coll. (2012, 2011) pour l'examen détaillé des propriétés de la méthode CEM et la comparaison avec d'autres méthodes d'appariement.
 5. Les pondérations attribuées à l'échantillon de contrôle apparié seront égales au ratio de la taille de l'échantillon de survivants du cancer (n_t) par rapport à la taille de l'échantillon de contrôle (n_c) dans chaque strate, multiplié par le ratio de la taille totale de l'échantillon de contrôle (N_c) par rapport à la taille totale de l'échantillon de survivants du cancer (N_t). Les pondérations de l'échantillon apparié des survivants du cancer sont des unités. Les pondérations des enregistrements non appariés correspondent à zéro.
 6. Ho et coll. (2007) recommandent aux chercheurs d'utiliser des modèles paramétriques pour rajuster l'échantillon apparié. Ils qualifient l'avantage d'une procédure en deux étapes comme étant doublement robuste dans le sens que, dans de faibles conditions, si le modèle d'appariement ou le modèle paramétrique est exact — mais pas nécessairement les deux — les estimations causales seront cohérentes.
 7. La dépendance du modèle se définit par la mesure dans laquelle la valeur prévue de la variable des résultats varie en fonction du modèle statistique pour un ensemble donné de variables explicatives (Ho et coll., 2007). L'une des principales raisons de l'appariement est d'éliminer la dépendance du modèle; toutefois, cette raison n'a jamais été mathématiquement démontrée pour les méthodes d'appariement antérieures utilisées couramment dans les diverses analyses. Pour obtenir plus de détails, consulter Iacus et coll. (2011).
 8. Dans les documents publiés sur l'évaluation, on parle de « l'effet moyen du traitement dans l'échantillon sur les personnes traitées ».

2.2 Facteurs de confusion non observables influant sur le cancer et les résultats sur le marché du travail

On s'inquiète également de la corrélation entre l'incidence du cancer et les caractéristiques individuelles non observables — qu'on appelle aussi « sélection en fonction de facteurs non observables » dans les documents relatifs à l'évaluation. Tout comme les différences observées, les différences non observées entre les échantillons de survivants du cancer et de contrôle peuvent aussi produire des résultats différents sur le marché du travail pour ces deux groupes. À titre d'exemple, le développement d'un cancer peut être corrélé à un mode de vie malsain non observé dans les données (p. ex., tabagisme ou mauvaise alimentation), ce qui, en retour, peut dépendre de caractéristiques individuelles précises non observées liées à la motivation personnelle. Les personnes ayant ces caractéristiques seraient moins susceptibles de développer un cancer (p. ex., en s'abstenant de fumer) et, en même temps, seraient plus susceptibles de gagner un revenu supérieur. Même si l'effet de confusion des caractéristiques non observées dans cet exemple n'est pas particulièrement fort, le défaut de contrôler la différence moyenne dans les caractéristiques non observées des groupes des survivants du cancer et de contrôle entraînerait la surestimation de l'effet négatif du cancer sur les résultats sur le marché du travail⁹.

Heckman et coll. (1997, 1998) ont proposé de combiner les démarches de DDD et d'appariement pour éliminer les différences non observables intemporelles entre les deux groupes. Toutefois, cette démarche nécessite des données recueillies avant et après le diagnostic, lesquelles font maintenant partie des données de couplages nouvellement accessibles utilisées dans cette étude. En adoptant la stratégie de Heckman et coll. (1997, 1998), la différence dans les gains des particuliers avant et après un diagnostic de cancer sert de variable dépendante pour laquelle on calcule une régression sur la variable du cancer à l'aide des pondérations d'appariement de la méthode CEM¹⁰.

2.3 Spécification du modèle

Dans le modèle statistique, T correspond à l'année du diagnostic de cancer dans l'échantillon de survivants du cancer et à l'année d'appariement dans l'échantillon de contrôle. Les années civiles avant et après l'année du diagnostic de cancer (ou les années d'appariement dans l'échantillon de contrôle) correspondent à $T - j$ et $T + j$, où $j = 1, \dots, J$. Les résultats sur le marché du travail pour les années $T + 1$, $T + 2$ et $T + 3$ sont modélisés en fonction d'un indicateur du cancer et de caractéristiques individuelles, et en fonction de caractéristiques d'emploi :

$$\text{Résultat sur le marché du travail} = f(\text{cancer}, \text{caractéristiques individuelles}, \text{caractéristiques d'emploi}) + \varepsilon.$$

Les caractéristiques individuelles sont notamment l'âge, le carré de l'âge, l'état matrimonial (en couple ou célibataire), un indicateur d'une incapacité de longue durée en 1991 et la province de résidence. Les variables utilisées pour l'appariement — le sexe, le niveau de scolarité, l'appartenance à une minorité visible — font aussi partie de la spécification du modèle¹¹. Les caractéristiques professionnelles sont des indicateurs d'un revenu non nul provenant d'un travail autonome et de l'affiliation syndicale (association professionnelle). Les variables

9. Moran et coll. (2011) expriment une préoccupation semblable dans leur étude des effets du cancer sur l'emploi.

10. En faisant la différence entre les gains individuels avant et après le diagnostic de cancer, il est aussi possible de tenir compte — au moins en partie — des effets fixes individuels non observés qui ne sont pas nécessairement corrélés au cancer (p. ex., l'éthique professionnelle d'une personne).

11. Cela ne change pas l'effet du cancer (le coefficient sur la variable du cancer) étant donné que grâce à la repondération, la variable du cancer est indépendante de la moyenne en fonction de ces variables d'appariement.

nominales du quintile pour le total des gains pour l'année $T-1$ et les variables nominales de l'année font aussi partie du modèle.

3 Données

Le Recensement de 1991 – Fichier de données longitudinales sur la main-d'œuvre (FDLMO) est un ensemble de données unique qui contient des données provenant de quatre sources : le Recensement de la population du Canada de 1991, la Base canadienne de données sur la mortalité (BCDM), la Base canadienne de données sur le cancer (BCDC) et le FDLMO.

La BCDM contient les enregistrements de décès des particuliers à partir de 1950. Les bureaux de l'état civil des provinces et des territoires fournissent ces enregistrements annuellement à Statistique Canada pour l'analyse à l'échelon national.

La BCDC est une banque de données qui combine deux sources de données sur le cancer : le Registre canadien du cancer (RCC) et le Système national de déclaration des cas de cancer (SNDCC). Le premier est une base de données sur les tumeurs axée sur la personne, qui comprend des renseignements cliniques et démographiques sur les résidents canadiens ayant un cancer depuis 1992 (Statistique Canada, 2008). La deuxième source est une base de données historique axée sur les tumeurs, qui contient les cas de cancer diagnostiqués datant d'aussi loin que 1969 (Carpenter et coll., 2008). Les dossiers personnels sur le cancer provenant du RCC servent à l'analyse; les renseignements historiques provenant du SNDCC servent à vérifier que les personnes qui figurent dans le RCC n'ont pas d'antécédents de cancer.

Le FDLMO représente 10 % de l'échantillon aléatoire des Canadiens qui ont produit une déclaration de revenus des particuliers (formulaire T1 général, *Déclaration de revenus et de prestations*) ou qui ont reçu un état de la rémunération (formulaire T4, *État de la rémunération payée*) de leur employeur chaque année à partir de 1983. Lorsque les personnes sont sélectionnées dans le FDLMO, elles sont suivies aussi longtemps qu'elles produisent une déclaration de revenus (formulaire T1 général) ou que leurs revenus sont déclarés à l'Agence du revenu du Canada (ARC) par leurs employeurs, peu importe leur situation d'emploi. La version actuelle du FDLMO contient des renseignements sur les salaires, les traitements et le revenu provenant d'un travail autonome, ainsi que des renseignements au niveau des entreprises¹². Les salaires et traitements proviennent des feuillets T4 émis par les employeurs. Le revenu provenant d'un travail autonome et les renseignements de base personnels (l'état matrimonial, la province de résidence, etc.) sont tirés des déclarations de revenus des particuliers (T1)¹³.

La Division de l'analyse de la santé de Statistique Canada a couplé au départ des renseignements personnels sélectionnés de la BCDM et de la BCDC aux dossiers individuels des personnes de 25 ans et plus dans le fichier du Recensement de 1991¹⁴. Ce premier couplage des données s'appelle la « cohorte du recensement canadien de 1991 : suivi de la mortalité et de l'incidence du cancer ». Les enregistrements de décès jusqu'en 2006 et les

12. Le revenu net provenant d'un travail autonome est le revenu provenant d'entreprises non constituées en société; le revenu des entreprises constituées en société est déclaré comme des salaires et traitements.

13. Les taux de déclaration au Canada sont très élevés principalement parce que les avantages fiscaux, comme les crédits d'impôt remboursables de la taxe sur les produits et services (TPS), encouragent les particuliers à produire une déclaration de revenus, même s'ils ont un faible revenu ou n'ont aucun revenu.

14. À cette étape, le couplage a été effectué à l'aide des méthodes probabilistes de couplage d'enregistrements de Statistique Canada. Consulter Wilkins et coll. (2008) pour obtenir des renseignements détaillés sur le couplage des données. La deuxième étape du couplage du FDLMO et de la cohorte du Recensement de 1991 a été effectuée au moyen d'un processus déterministe de couplage d'enregistrements à partir du numéro d'assurance sociale (NAS).

enregistrements de cancer jusqu'en 2003 ont été tirés des deux banques de données, soit la BCDM et la BCDC¹⁵. Par la suite, les enregistrements du FDLMO ont été couplés à la cohorte du recensement canadien de 1991 pour produire la composante du revenu indispensable.

L'échantillon de données du Recensement de 1991 – FDLMO contient 263 674 enregistrements individuels qui correspondent à environ 1,4 % de la population canadienne âgée de 25 ans et plus en 1991. Environ 58,8 % de la cohorte du Recensement de 1991 – FDLMO a été observée pendant les 28 années que couvre le FDLMO, de 1983 à 2010; la moyenne du nombre d'années de présence d'une personne dans l'échantillon est de 24,8. Les taux de déclaration de revenus étaient légèrement inférieurs dans les années 1980 comparativement aux décennies plus récentes; de 1990 à 2010, 66,9 % des 263 638 personnes sont observées pendant toute la période de 21 ans, et la moyenne du nombre d'années d'observation d'une personne dans l'échantillon est de 18,5.

3.1 Échantillon de survivants du cancer

Dans un premier temps, les personnes atteintes d'un cancer pour la première fois dans la période de 1992 à 2000 sont sélectionnées dans la cohorte du Recensement de 1991 – FDLMO. Puisque l'étude s'intéresse aux résultats sur le marché du travail des survivants du cancer, l'échantillon englobe seulement les personnes qui étaient âgées de moins de 62 ans lorsqu'elles ont reçu leur diagnostic de cancer. Les données du Recensement de 1991 – FDLMO contiennent seulement les sièges du cancer, mais ne contiennent aucun renseignement sur la gravité et le stade du cancer. Toutes les tumeurs sont d'abord regroupées selon les 26 sièges du cancer les plus courants, conformément au groupement du SEER (Surveillance, Epidemiology, and End Results)¹⁶ et aux codes de la Classification statistique internationale des maladies, 9^e édition (CIM-9) et de la Classification internationale des maladies : Oncologie, deuxième et troisième éditions (CIM-O-2/3) accessibles dans les données. Dans le cas des personnes qui avaient de multiples enregistrements de tumeur pour une année donnée, on a sélectionné l'enregistrement contenant une tumeur maligne et le siège primaire de la tumeur ayant le plus faible taux de survie selon les études précédentes sur les taux de survie au cancer (Ellison et coll., 2011). Ensuite, les 26 sites sont regroupés en trois catégories — taux de survie élevé, moyen et faible — selon le ratio de survie relatif pendant cinq ans (ci-après appelée la « catégorie de taux de survie au cancer ») pour chaque siège du cancer (Ellison et coll., 2011). Ces catégories de cancer servent d'approximation pour la gravité moyenne du cancer afin de déterminer l'impact de la gravité du cancer pour les résultats sur le marché du travail des survivants du cancer.¹⁷

Les données du Recensement de 1991 – FDLMO contiennent 5 185 personnes (2 120 hommes et 3 065 femmes) ayant reçu un diagnostic de cancer pour la première fois dans la période de 1992 à 2000 qui étaient âgées de 61 ans ou moins dans l'année du diagnostic¹⁸. Le tableau A.1 en annexe indique le nombre de cas pour chacun des 26 sièges du cancer. Parmi les 5 185 cas de cancer, environ 8 % (412) sont des tumeurs qui ne correspondent pas aux 26 sièges courants du cancer. Les sièges de cancer les plus courants sont le sein (21,5 %), les poumons et les bronches (10 %), le col utérin (7,9 %), la prostate (5,9 %) et le côlon (5,8 %). Lorsqu'on répartit les cancers selon la catégorie de taux de survie au cancer, 2 023 (39 %) cas appartiennent à la catégorie du taux de survie élevé, 1 659 (32 %) à la catégorie du taux moyen, et 1 091 (21 %) à la catégorie du taux faible. De plus, 2 020 enregistrements (39 % du

15. L'élaboration de données supplémentaires pour étendre les données de la cohorte du Recensement de 1991 aux années plus récentes est actuellement en cours. Des renseignements détaillés sur le Recensement de 1991 et les deux bases de données sont accessibles dans le site Web de Statistique Canada (www.statcan.gc.ca).

16. <http://seer.cancer.gov>.

17. Les différences dans les taux moyens de survie ne sont pas nécessairement équivalentes aux différences de morbidité pour tous les types de cancer différents. Toutefois, il n'y a pas de bon indice de la morbidité qui pourrait servir de mesure plus fiable de la gravité de tous les types de cancer.

18. Parmi ces personnes, 219 d'entre elles avaient reçu de multiples diagnostics de cancer entre 1993 et 2003.

total de l'échantillon) sont appariés aux enregistrements de décès provenant des données sur la mortalité dans la BCDM (la dernière année accessible étant 2006). Selon les catégories de taux de survie au cancer, 19,1 % des enregistrements de la catégorie du taux de survie élevé, 34,3 % des enregistrements de la catégorie du taux de survie moyen et 81,5 % de la catégorie du faible taux de survie sont appariés à des enregistrements de décès.

Pour effectuer l'analyse, 3 716 personnes qui ont survécu pendant au moins trois ans après avoir reçu un diagnostic de cancer sont conservées dans l'échantillon, et 77 personnes qui étaient atteintes de multiples nouveaux cancers au cours de ces trois années sont exclues de l'échantillon. Enfin, en fonction de la participation au marché du travail avant d'avoir reçu un diagnostic du cancer, l'échantillon des survivants du cancer englobe seulement les personnes qui travaillaient dans l'année où elles ont reçu un diagnostic (T) ainsi que dans les deux années précédentes ($T-1$ et $T-2$), puisque l'étude est centrée sur les effets du cancer sur le marché du travail parmi les personnes qui participaient très activement au marché du travail avant de recevoir leur diagnostic de cancer¹⁹. La sous-section 3.3 présentera de plus amples détails sur les mesures des résultats sur le marché du travail. En raison de cette restriction, 1 042 personnes sont encore soustraites de l'échantillon de survivants du cancer. En résumé, l'échantillon de survivants du cancer est composé de 2 597 survivants du cancer qui ont reçu un diagnostic de cancer pour la première fois dans la période de 1992 à 2000, et qui étaient âgés de 61 ans ou moins au moment du diagnostic. Ces personnes travaillaient au cours de l'année où elles ont reçu le diagnostic et les deux années précédentes; elles ont survécu au moins trois ans après le diagnostic et n'ont eu aucun autre diagnostic subséquent durant ces trois années.

Le tableau 1 illustre les types de cancer dans l'échantillon de survivants du cancer. Puisque l'échantillon de survivants du cancer contient uniquement les personnes qui ont reçu un diagnostic de cancer et qui ont survécu pendant plus de trois ans, il contient proportionnellement moins de personnes dans la catégorie du faible taux de survie par rapport à ce qui est illustré dans l'annexe A. La moitié de l'échantillon de survivants du cancer appartient à la catégorie de survie élevée. La deuxième colonne du tableau 2 dans la prochaine sous-section contient les caractéristiques générales de l'échantillon de survivants du cancer.

3.2 Échantillon de contrôle

La présente section décrit la sélection de l'échantillon de contrôle dont les résultats sur le marché du travail sont comparés à ceux de l'échantillon de survivants du cancer. Les critères généraux de sélection utilisés pour choisir l'échantillon de survivants du cancer sont d'abord appliqués : les personnes sont présélectionnées en fonction du diagnostic de cancer, de l'année, de l'âge, de l'enregistrement de décès et des antécédents professionnels provenant des mêmes données du Recensement de 1991 – FDLMO. Les personnes du groupe de 192 537 personnes qui n'ont pas d'enregistrement de cancer dans le RCC ou dans le SNDCC (enregistrements historiques de tumeurs) pour toutes les années accessibles sont choisies au départ dans les données. Comme dans l'échantillon de survivants du cancer, l'échantillon de contrôle est formé de personnes âgées de 61 ans ou moins pour chaque année de la période de 1992 à 2000. Ensuite, pour chaque année, les personnes qui ont vécu au moins trois ans suivant l'année T (ce critère exclut 1 053 personnes) et qui ont travaillé dans l'année T et les deux années précédentes ($T-1$ et $T-2$) (ce critère exclut encore 22 576 personnes) sont conservées dans l'échantillon. À cette étape, 168 908 personnes sont présélectionnées dans

19. Deux autres échantillons ont été conçus pour déterminer la solidité des résultats par rapport aux différentes restrictions applicables à la situation d'emploi d'une personne avant qu'elle reçoive un diagnostic de cancer. Un échantillon contient les personnes qui travaillaient dans les années T et $T-1$, et l'autre échantillon contient les personnes qui travaillaient dans les années T , $T-1$, $T-2$ et $T-3$. Toutes les constatations sont très robustes aux différentes restrictions d'échantillonnage. Consulter aussi la note de bas de page 25 pour en savoir davantage.

les données du Recensement de 1991 – FDLMO, et le nombre d'observations regroupées pour les neuf années de 1992 à 2000 atteint 1 228 551.

Tableau 1

Types (sièges) de cancer dans l'échantillon de survivants du cancer par catégorie de taux de survie

	Part de l'échantillon de survivants du cancer	
	Type	Survie
	pourcentage	
Taux élevé de survie	...	50,10
Thyroïde	3,93	...
Prostate	7,74	...
Testicules	1,85	...
Mélanome de la peau	6,93	...
Sein	25,68	...
Corps utérin	3,20	...
Lymphome hodgkinien	0,77	...
Taux moyen de survie	...	34,35
Leucémie lymphoïde chronique	0,85	...
Col utérin	10,63	...
Vessie (y compris in situ)	3,43	...
Rein et bassin du rein	2,89	...
Tissu mou	0,89	...
Larynx	0,85	...
Rectum	3,50	...
Côlon	4,74	...
Lymphome non hodgkinien	4,35	...
Oral (cavité buccale et pharynx)	2,23	...
Faible taux de survie	...	8,09
Ovaires	1,96	...
Myélome multiple	0,62	...
Leucémie (à l'exclusion de la leucémie lymphoïde chronique)	0,81	...
Estomac	0,35	...
Cerveau	1,19	...
Foie	x	...
Poumon et bronches	2,66	...
Œsophage	x	...
Pancréas	0,19	...
Autres	7,47	...
	nombre	
Nombre total de survivants du cancer	2597	...

... n'ayant pas lieu de figurer

x confidentiel en vertu des dispositions de la *Loi sur la statistique*

Note : Calculs de l'auteur.

Source : Statistique Canada, Recensement de 1991 – Fichier de données longitudinales sur la main-d'œuvre.

À l'étape suivante, la méthode CEM est appliquée pour appairer cet échantillon présélectionné avec l'échantillon de survivants du cancer afin d'obtenir un échantillon de contrôle apparié. Les variables d'appariement sont notamment l'âge dans l'année *T*, le sexe, le niveau de scolarité, l'appartenance à une minorité visible, l'année *T* et la province de résidence dans l'année *T*-1. Les âges sont regroupés en intervalle de cinq ans. L'appartenance à une minorité visible est répartie en trois catégories : ceux qui ne font pas partie des minorités, les Asiatiques et les autres minorités. Les Territoires du Nord-Ouest, le Nunavut et le Yukon sont regroupés en une seule catégorie (territoires du Nord). La méthode CEM limite automatiquement les échantillons

de survivants du cancer et de contrôle aux domaines de soutien empirique communs en élaguant les observations non appariées des deux échantillons. Dans le groupe de survivants du cancer, quatre personnes sur 2 597 ne peuvent être appariées à qui que ce soit dans le groupe de contrôle. Le groupe de contrôle apparié est composé de 624 835 observations (données annuelles totalisées pour 142 196 personnes)²⁰. Le tableau 2 contient les statistiques sommaires des variables d'appariement pour le groupe de survivants du cancer et le groupe de contrôle, avant et après l'appariement. Les statistiques sommaires des autres variables, qui ne sont pas utilisées dans la procédure d'appariement, mais qui sont utilisées pour l'analyse de régression, figurent dans le tableau A.2 en annexe. La dernière colonne du tableau 2 contient les différences dans les moyennes (proportions) des variables entre l'échantillon de survivants du cancer et l'échantillon de contrôle avant l'appariement et les tests t afin de déceler les différences entre ces deux échantillons²¹.

L'âge moyen dans l'échantillon de survivants du cancer est supérieur à celui dans l'échantillon de contrôle apparié au préalable. La répartition selon l'âge illustre que l'incidence du cancer a une corrélation positive avec l'âge. L'échantillon de survivants du cancer compte aussi proportionnellement un plus grand nombre de personnes qui n'ont pas de diplôme d'études secondaires que l'échantillon de contrôle. Les publications récentes contiennent des données probantes sur l'incidence causale du niveau de scolarité atteint sur la santé et un mode de vie sain (Jones et coll., 2011). L'échantillon de survivants du cancer renferme proportionnellement plus de femmes que l'échantillon de contrôle préapparié. Le cancer du sein, qui est le type de cancer le plus courant chez les femmes, est aussi le type de cancer le plus répandu dans l'échantillon de l'étude. Chez les hommes, le cancer de la prostate est le type de cancer le plus courant, mais il se développe habituellement plus tard dans la vie à un âge qui ne fait pas partie de l'échantillon d'enquête. Les données sont aussi conformes au fait que le cancer est généralement moins fréquent chez les Asiatiques que dans les populations non minoritaires.

Comme l'illustre la quatrième colonne, la répartition repondérée (après l'appariement) est pratiquement identique dans l'échantillon de survivants du cancer et dans l'échantillon de contrôle apparié. Les procédures d'appariement et de repondération nous permettent d'envisager le cancer comme un « traitement » assigné de façon aléatoire en fonction de caractéristiques observables (avant le cancer) et appariées de l'échantillon.

20. Comme les mêmes personnes peuvent paraître dans l'échantillon de contrôle dans plus d'une année d'appariement, les erreurs types comprises dans tous les modèles de régression sont groupées selon les identificateurs individuels.

21. En règle générale, les études antérieures qui ont recours aux méthodes d'appariement déclarent des différences moyennes normalisées avant et après dans les covariables, plutôt que des statistiques t (ou des résultats de l'essai t). Cette pratique sert à illustrer l'amélioration de l'équilibre des moyennes entre les deux échantillons parce que — à la différence des statistiques t — les différences moyennes normalisées ne dépendent pas de la taille de l'échantillon apparié ou non apparié (Imbens et Wooldridge, 2009). Toutefois, dans cette étude, il n'y a pas de différences moyennes restantes à la suite de la méthode CEM entre l'échantillon de survivants du cancer et l'échantillon de contrôle apparié (comme l'illustre le tableau 2). Le tableau 2 expose les résultats de l'essai t simplement pour illustrer que, pour les variables d'appariement, les différences moyennes entre les deux échantillons étaient statistiquement significatives avant l'application de la procédure d'appariement.

Tableau 2

Statistiques sommaires pour les échantillons de survivants du cancer et de contrôle avant et après l'appariement et l'élagage

	Échantillon de survivants du cancer	Échantillon de contrôle		Différence non appariée points de pourcentage
		Non apparié pourcentage	Apparié et répondu	
Âge (moyen) à la période <i>T</i>	48,12	43,05	47,97	5,07 ***
Groupe d'âge à la période <i>T</i>				
25 à 29	0,92	3,08	0,93	-2,16 ***
30 à 34	6,62	13,92	6,63	-7,30 ***
35 à 39	11,05	20,83	11,07	-9,78 ***
40 à 44	14,13	20,29	14,15	-6,16 ***
45 à 49	18,33	17,72	18,36	0,61
50 à 54	21,72	13,44	21,75	8,28 ***
55 à 59	20,64	8,48	20,63	12,16 ***
60 à 61	6,59	2,24	6,48	4,34 ***
Femme	62,84	46,02	62,86	16,82 ***
Plus haut niveau de scolarité atteint				
Sans diplôme d'études secondaires	25,45	22,21	25,49	3,25 ***
Études secondaires avec/sans certificat d'une école de métiers	39,24	40,86	39,26	-1,62 †
Études postsecondaires non universitaires	19,21	19,23	19,24	-0,02
Grade universitaire	16,10	17,70	16,01	-1,61 *
Minorité visible				
Groupe non minoritaire	93,96	92,03	94,06	1,93 ***
Asiatiques	3,97	5,26	3,93	-1,30 **
Autre	2,08	2,71	2,01	-0,63 *
Province/région à la période <i>T-1</i>				
Terre-Neuve-et-Labrador	1,54	1,82	1,58	-0,28
Île-du-Prince-Édouard	0,77	0,51	0,74	0,26 †
Nouvelle-Écosse	3,66	3,12	3,58	0,54
Nouveau-Brunswick	2,97	2,53	2,85	0,43
Québec	23,72	25,29	23,81	-1,57 †
Ontario	32,19	36,18	32,35	-3,99 ***
Manitoba	4,81	4,06	4,85	0,76 †
Saskatchewan	3,58	3,60	3,54	-0,02
Alberta	12,01	9,98	12,09	2,03 **
Colombie-Britannique	13,98	11,94	13,83	2,04 **
Territoires du Nord	0,77	0,93	0,77	-0,16
Valeur manquante	0,00	0,04	0,02	-0,04
Année de la période <i>T</i>				
1992	9,90	12,18	9,91	-2,29 **
1993	10,44	11,84	10,41	-1,40 *
1994	10,71	11,50	10,68	-0,79
1995	10,94	11,27	10,91	-0,34
1996	11,55	11,05	11,57	0,50
1997	11,13	10,84	11,15	0,29
1998	11,36	10,58	11,38	0,78
1999	12,17	10,43	12,15	1,73 **
2000	11,82	10,30	11,84	1,52 *
		nombre		
<i>n</i>	2 597	1 228 551	624 835	...

... n'ayant pas lieu de figurer

*** valeur significativement différente de l'estimation pour la catégorie de référence ($p < 0,001$)

** valeur significativement différente de l'estimation pour la catégorie de référence ($p < 0,01$)

* valeur significativement différente de l'estimation pour la catégorie de référence ($p < 0,05$)

† valeur significativement différente de l'estimation pour la catégorie de référence ($p < 0,10$)

Notes : La dernière colonne affiche les différences moyennes (proportion) entre l'échantillon de survivants du cancer et l'échantillon de contrôle avant l'appariement; des signes conventionnels indiquent l'importance des différences du test *t* entre les deux groupes. Les *T* sont les années d'appariement qui sont les années du diagnostic de cancer pour l'échantillon de survivants du cancer.

n : nombre d'observations. *T-1* : *T* moins 1. Calculs de l'auteur.

Source : Statistique Canada, Recensement de 1991 – Fichier de données longitudinales sur la main-d'œuvre.

Le tableau A.2 en annexe contient les statistiques sommaires pour les gains des deux années ($T-1$ et $T-2$) précédant le diagnostic de cancer dans l'échantillon de survivants du cancer, ainsi que les gains correspondant dans l'échantillon de contrôle apparié. En raison de l'appariement, les gains moyens et les répartitions moyennes dans l'échantillon de survivants du cancer dans les années précédant le diagnostic de cancer sont pratiquement identiques à ceux de l'échantillon de contrôle non apparié. Cette observation a deux répercussions : d'abord, le total des gains annuels peut être prévu de façon exacte dans les échantillons d'étude en fonction des variables utilisées pour l'appariement. Ensuite, s'il n'y avait pas eu de diagnostic de cancer, les gains moyens de l'échantillon de survivants du cancer auraient suivi la même tendance commune que celle de l'échantillon de contrôle apparié. Cette observation est particulièrement importante, et elle est conforme à l'hypothèse principale de la démarche DDD utilisée dans l'analyse empirique²².

3.3 Mesures des résultats sur le marché du travail

L'étude examine deux variables des résultats sur le marché du travail. La première est une variable binaire annuelle de la situation d'emploi; la seconde, une variable continue pour les gains annuels totaux (en dollars de 2010). Les gains annuels totaux d'une personne désignent la somme de tous les salaires et traitements touchés une année donnée, plus le revenu net provenant d'un travail autonome dans la même année. Les personnes qui avaient des gains non nuls — soit positifs ou négatifs — sont considérées comme ayant travaillé au cours de l'année. Dans les données, les fréquences des variables des résultats sont annuelles. Les données ne contiennent aucune information sur le nombre de mois ou de semaines qu'a travaillé chaque survivant du cancer avant ou après le diagnostic de cancer au cours de l'année du diagnostic. Pour s'assurer que les personnes n'avaient pas arrêté de travailler avant de recevoir leur diagnostic de cancer, l'échantillon de survivants du cancer englobe seulement les personnes qui ont des gains non nuls pour l'année du diagnostic de cancer (T), ainsi que pour les deux années précédentes ($T-1$ et $T-2$). L'étude examine l'effet du cancer sur les résultats sur le marché du travail au cours des années suivant le diagnostic en comparant les résultats sur le marché du travail de l'échantillon de survivants du cancer et de l'échantillon de contrôle pendant trois périodes consécutives : $T+1$, $T+2$ et $T+3$.

Il convient de souligner que T correspond à l'année civile, plutôt qu'au moment précis du diagnostic. Par conséquent, $T+j$ correspond aux années civiles suivant l'année du diagnostic, plutôt qu'aux intervalles de temps écoulés depuis le moment exact du diagnostic. À titre d'exemple, supposons qu'un survivant du cancer avait des gains nuls dans l'année $T+1$ (l'année suivant l'année du diagnostic). On sait que cette personne n'a pas travaillé pendant au moins 12 mois, mais cette période pourrait correspondre à une période aussi longue que 24 mois à partir du moment du diagnostic, si la personne a reçu son diagnostic au début de l'année T . Toutefois, si un survivant du cancer a des gains positifs pour l'année $T+1$, ces gains reflètent les gains totaux pour la période de 12 à 24 mois à partir du diagnostic du cancer. Par conséquent, la période de travail réelle à partir du jour du diagnostic du cancer jusqu'à la fin de l'année $T+1$ dépend du moment du diagnostic au cours de l'année T . Par conséquent, les données ne sont pas appropriées pour l'analyse des effets à court terme (p. ex., moins d'une année) d'un diagnostic de cancer sur les résultats sur le marché du travail; c'est pourquoi seuls les résultats à long terme sur le marché du travail sont examinés. Le tableau A.3 de l'annexe contient les moyennes non corrigées de l'emploi et des gains pour les années $T+1$, $T+2$ et $T+3$ selon l'état du cancer et la catégorie du taux de survie au cancer²³. Les différences moyennes

22. L'hypothèse sous-jacente pour l'interprétation causale est que, en l'absence de diagnostic de cancer, les gains de l'échantillon de survivants du cancer suivent la même tendance commune que ceux de l'échantillon de contrôle.

23. Bien que les taux de déclaration de revenus au Canada soient très élevés, les données sur les particuliers pour les années $T+1$, $T+2$ et $T+3$ peuvent être manquantes, et le volet déséquilibré. L'immigration et l'émigration, la déclaration tardive ou la non-admissibilité aux prestations sont des raisons possibles pour manquer l'année d'imposition.

non corrigées dans l'emploi et les gains entre l'échantillon de survivants du cancer et l'échantillon de contrôle apparié sont -4,36 % et -9 022 \$ pour l'année $T+1$.

4 Résultats

4.1 Effet du cancer sur l'emploi

Le volet de gauche du tableau 3 montre les estimations des effets du cancer sur la probabilité de travailler au cours des années $T+1$, $T+2$ et $T+3$. Puisque toutes les personnes dans l'échantillon travaillaient au cours des années T , $T-1$ et $T-2$, l'estimation de la probabilité de travailler pour chaque période correspond à l'estimation de la probabilité de la transition de l'emploi au non-emploi entre l'année T et $T+1$, T et $T+2$ et T et $T+3$, respectivement²⁴. Ici, l'effet du cancer sur l'emploi à la fin de chaque intervalle combine divers aspects de toutes les transitions sur le marché du travail qui ont eu lieu entre la période T et la période finale. Certaines personnes, par exemple, peuvent arrêter de travailler temporairement après un diagnostic de cancer pour suivre un traitement et se rétablir après le traitement, mais retourner au travail dans les périodes subséquentes. D'autres pourraient ne pas quitter le marché du travail immédiatement après le diagnostic de cancer, mais partiront dans les périodes subséquentes. Certaines personnes pourraient quitter le marché du travail de façon permanente après un diagnostic de cancer. Les différences entre les groupes de survivants et de contrôle dans la probabilité de travailler à la fin de chaque période reflètent l'effet cumulatif du cancer sur l'emploi à partir de l'année T jusqu'à cette période.

Le volet dans le coin supérieur gauche montre que la probabilité de travailler dans l'année $T+1$ est inférieure, en moyenne, de 3,0 points de pourcentage dans le groupe de survivants du cancer par rapport au groupe de contrôle. La différence des probabilités augmente et atteint 3,7 au cours de l'année $T+2$, et 4,8 points de pourcentage au cours de l'année $T+3$.

Le volet de droite du tableau 3 présente les effets du cancer sur la probabilité de travailler en fonction d'un emploi continu suivant le diagnostic de cancer jusqu'à l'année $T+1$, $T+2$ ou $T+3$. Cela équivaut à estimer la probabilité de quitter son emploi entre deux périodes consécutives (un taux de départ plus courant). Les résultats dans le volet en haut à droite laissent entendre que, même si le cancer n'interrompt pas la participation d'une personne au marché du travail peu après le diagnostic, il peut influencer sur son départ du marché du travail à plus long terme. Le taux de départ diminue au fil des années après le diagnostic de cancer, mais il est toujours supérieur de 1,3 point de pourcentage au taux de départ moyen des personnes qui n'ont jamais reçu de diagnostic de cancer dans l'année $T+3$, et la différence est statistiquement significative.

Dans le volet inférieur, à la place d'une variable binaire du cancer dans le modèle, trois catégories de taux de survie au cancer — taux élevé, taux moyen et taux faible — font partie du modèle d'estimation. Parmi ces trois catégories, l'effet négatif du cancer sur l'emploi est le plus important chez les survivants du cancer de la catégorie du faible taux de survie pour les trois périodes, comparativement aux personnes n'ayant jamais reçu de diagnostic de cancer. Dans le volet inférieur, la probabilité de travailler dans l'année $T+1$ pour les survivants du groupe du faible taux de survie est inférieure de 11 points de pourcentage à la probabilité de travailler dans le groupe de contrôle. Par ailleurs, le volet en bas à droite montre que, pour la catégorie du faible taux de survie, les taux de départ qui dépendent du travail au cours des trois années précédentes demeurent aussi élevés et statistiquement significatifs pour les trois périodes suivant le diagnostic. Toutefois, dans la catégorie du taux de survie élevé, pour ceux qui

24. L'interprétation des estimations sous forme de probabilité de départ nécessiterait simplement d'inverser les signes.

continuent de travailler après le diagnostic de cancer, le taux de départ estimé pour l'année $T+3$ ne diffère plus de façon significative de la moyenne observée dans le groupe de contrôle.

Tableau 3
Effet du cancer sur l'emploi

	Travaille dans chaque période : $\Pr(\text{travail}_{T+j}=1)$			Travaille continuellement jusqu'à chaque période : $\Pr(\text{travail}_{T+j}=1 \mid \text{travail}_{T+j-1}=1,$ $\text{travail}_{T+j-2}=1)$		
	$T+1$	$T+2$	$T+3$	$T+1$	$T+2$	$T+3$
Cancer						
Coefficient	-0,030 ***	-0,037 ***	-0,048 ***	-0,030 ***	-0,018 ***	-0,013 **
Erreur-type	0,005	0,006	0,007	0,005	0,005	0,005
n (échantillon pondéré)	588 593	586 698	584 994	588 593	567 922	547 901
Catégorie de taux de survie						
Élevé						
Coefficient	-0,021 **	-0,036 ***	-0,039 ***	-0,021 **	-0,025 ***	-0,008
Erreur-type	0,007	0,008	0,009	0,007	0,007	0,007
Moyen						
Coefficient	-0,034 ***	-0,025 *	-0,024 *	-0,034 ***	-0,001	-0,001
Erreur-type	0,009	0,010	0,011	0,009	0,007	0,007
Faible						
Coefficient	-0,107 ***	-0,129 ***	-0,206 ***	-0,107 ***	-0,080 **	-0,081 **
Erreur-type	0,025	0,028	0,031	0,025	0,025	0,027
n (échantillon pondéré)	588 411	586 514	584 815	588 411	567 746	547 733

*** valeur significativement différente de l'estimation pour la catégorie de référence ($p < 0,001$)

** valeur significativement différente de l'estimation pour la catégorie de référence ($p < 0,01$)

* valeur significativement différente de l'estimation pour la catégorie de référence ($p < 0,05$)

Notes : Les résultats sont les effets moyens partiels du modèle probit qui utilise des pondérations de la méthode Coarsened Exact Matching (CEM). Toutes les régressions sont conditionnelles à l'ensemble complet des variables de contrôle qui figure dans la liste de la sous-section 2.3. Dans le volet de gauche, les variables du travail autonome et de l'adhésion syndicale correspondent à la période T . Dans le volet de droite, les valeurs décalées des variables du travail autonome et de l'adhésion syndicale figurent à titre de contrôles. n : nombre d'observations; T : année du diagnostic du cancer pour l'échantillon de survivants du cancer. Calculs de l'auteur.

Source : Statistique Canada, Recensement de 1991 – Fichier de données longitudinales sur la main-d'œuvre.

4.2 Effet du cancer sur les gains

L'effet du cancer sur les gains annuels totaux est estimé au moyen d'une régression linéaire pondérée. Le volet de gauche du tableau 4 montre les estimations de l'effet du cancer pour chaque période sans condition relative au travail. Les gains annuels nuls font partie de l'analyse; par conséquent, les résultats présentés dans le volet de gauche comprennent les pertes de gain résultant des transitions d'une situation d'emploi avant un diagnostic de cancer au non-emploi après le diagnostic. Les estimations dans le volet de droite du tableau 4 sont conditionnelles au travail dans chaque période (gains nuls exclus). Ces résultats sont proches de l'effet du cancer dans la marge intensive des résultats sur le marché du travail et peuvent être comparés aux effets estimés dans la marge extensive figurant dans le tableau 3.

En moyenne, les survivants du cancer gagnent 5 079 \$, ou 12,1 %, de moins dans l'année $T+1$ que leurs homologues qui n'ont jamais reçu de diagnostic de cancer. Toutefois, l'effet négatif du cancer est moins grand dans l'année $T+2$, 9,7 %, et encore moins grand dans l'année $T+3$, 9,3 % (volet en haut à gauche dans le tableau 4). Lorsque l'effet du cancer est estimé en fonction du travail, les travailleurs qui ont reçu un diagnostic de cancer dans l'année T gagnent

4 675 \$ (10,6 %) de moins dans l'année $T+1$ que leurs homologues de l'échantillon de contrôle. L'effet négatif du cancer est moins important dans l'année $T+2$ (6,6 %) et dans l'année $T+3$ (5,4 %) que dans l'année $T+1$. En comparant ces résultats aux effets du cancer sur la probabilité de travailler illustrée dans le tableau 3, l'impact du cancer sur la situation d'emploi semble être plus persistant que son impact sur les gains annuels dans les périodes suivant le diagnostic.

Le volet inférieur du tableau 4 montre les résultats pour le modèle dans lequel la variable du cancer est répartie en trois catégories de taux de survie. Autant dans le volet de droite que dans celui de gauche, le recul des gains, par rapport à l'échantillon de contrôle, est supérieur pour les survivants de la catégorie du faible taux de survie comparativement à ceux de la catégorie du taux de survie élevé. Les effets du cancer sont plus importants et durent plus longtemps pour la catégorie du faible taux de survie que pour les autres catégories, ce qui est conforme aux résultats applicables à la situation d'emploi qui figurent dans le tableau 3. Avec le temps, les gains annuels moyens des survivants du cancer demeurant sur le marché du travail semblent s'apparenter aux gains annuels moyens des travailleurs ayant des caractéristiques semblables et n'ayant jamais reçu un diagnostic de cancer, particulièrement les travailleurs dont les taux de survie sont élevés.

Tableau 4
Effet du cancer sur les gains annuels (en dollars)

	Gains pour chaque période sans tenir compte des périodes de travail			Gains pour chaque période en tenant compte des périodes de travail		
	$T+1$	$T+2$	$T+3$	$T+1$	$T+2$	$T+3$
Cancer						
Coefficient	-5 079,37 ***	-3 980,47 ***	-3 756,72 ***	-4 674,72 ***	-2 940,20 **	-2 427,52 **
Erreur-type	902,42	878,38	810,23	953,02	939,40	880,57
n (échantillon pondéré)	596 374	595 210	594 014	577 036	564 400	552 832
Catégorie de taux de survie						
Élevé						
Coefficient	-4 420,21 **	-3 264,05 *	-3 018,97 *	-4 134,35 **	-1 953,90	-1 788,80
Erreur-type	1 520,97	1 483,54	1 237,04	1 600,95	1 606,21	1 350,72
Moyen						
Coefficient	-4 441,26 ***	-3 411,60 **	-2 943,02 *	-3 951,73 ***	-2 971,60 **	-2 335,15 †
Erreur-type	1 135,93	1 066,17	1 230,35	1 185,57	1 099,84	1 290,12
Faible						
Coefficient	-11 989,45 ***	-11 665,96 ***	-11 155,49 ***	-10 631,41 ***	-8 630,16 **	-5 254,77
Erreur-type	2 475,29	2 509,91	2 744,28	2 774,28	2 818,35	3 493,13
n (échantillon pondéré)	596 191	595 025	593 834	576 860	564 227	552 676

*** valeur significativement différente de l'estimation pour la catégorie de référence ($p < 0,001$)

** valeur significativement différente de l'estimation pour la catégorie de référence ($p < 0,01$)

* valeur significativement différente de l'estimation pour la catégorie de référence ($p < 0,05$)

† valeur significativement différente de l'estimation pour la catégorie de référence ($p < 0,10$)

Notes : Les résultats sont estimés pour un modèle de régression linéaire qui utilise des pondérations de la méthode Coarsened Exact Matching (CEM). Toutes les régressions sont conditionnelles à l'ensemble complet des variables de contrôle qui figure à la sous-section 2.3. n : nombre d'observations; T : année du diagnostic de cancer pour l'échantillon de survivants du cancer. Calculs de l'auteur.

Source : Statistique Canada, Recensement de 1991 – Fichier de données longitudinales sur la main-d'œuvre.

4.3 Association de la méthode de la différence dans les différences avec l'appariement

L'étude vérifie ensuite les différences entre les échantillons de survivants du cancer et de contrôle en faisant la différence entre les caractéristiques intemporelles non observées qui

peuvent être associées à la fois à l'incidence du cancer et aux gains. Les gains dans l'année $T-1$ sont soustraits des gains dans les années $T+1$, $T+2$ et $T+3$. Ces trois différences, avant et après le diagnostic, servent ensuite de variables dépendantes dans la régression pondérée de la différence dans les différences (DDD). Le volet supérieur du tableau 5 présente les résultats des régressions pondérées de la DDD pour les gains annuels. À titre de comparaison, le volet central du tableau 5 montre de nouveau les résultats de la régression linéaire pondérée déclarés au tableau 4 : le volet inférieur montre les résultats de la régression linéaire fondés sur l'échantillon non pondéré (avant l'appariement).

Le tableau 5 montre plusieurs tendances notables parmi les résultats de ces trois modèles²⁵. Premièrement, dans le volet supérieur, après avoir tenu compte, en outre, des différences non observables entre les échantillons de survivants du cancer et de contrôle, les effets négatifs du cancer perdent de l'importance comparativement à ceux du volet central, là où seules les différences observables entre les deux échantillons sont prises en compte. Ces différences dans les estimations témoignent des modestes différences intemporelles non observables entre l'échantillon de survivants du cancer (apparié) et l'échantillon de contrôle : cela mène à la surestimation de l'effet négatif du cancer dans le modèle présenté dans le tableau 4 (volet central du tableau 5). Deuxièmement, la comparaison des résultats du volet central à ceux du volet inférieur montre que, conformément à Ho et coll. (2007), le contrôle de la sélection des caractéristiques observables effectué par l'application de pondérations d'appariement améliore l'inférence causale, particulièrement dans l'année $T+1$ — une période relativement à court terme dans la présente étude. Toutefois, à plus long terme ($T+3$), la prise en compte de la sélection des caractéristiques observables dans le modèle semble moins importante. Troisièmement, la prise en compte des associations simultanées entre les caractéristiques individuelles intemporelles non observables et le cancer, d'une part, et ces caractéristiques et les résultats sur le marché du travail, d'autre part, semble plus importante pour estimer les effets à plus long terme du cancer sur les gains que pour estimer les effets à court terme.

4.4 Hétérogénéité de l'effet du cancer

Jusqu'à maintenant, l'analyse a été centrée sur l'effet général du cancer sur certains résultats sur le marché du travail. La question est maintenant de savoir si les effets du cancer sur les sous-groupes ayant certaines caractéristiques observables sont différents de l'effet moyen du cancer estimé pour l'ensemble de l'échantillon²⁶. Pour répondre à cette question, on estime des modèles d'interactions entre l'indicateur du cancer et les variables catégoriques d'âge, de niveau de scolarité et de gains préalables au diagnostic. Les interactions entre chaque ensemble de catégories et l'indicateur du cancer sont considérées séparément dans chaque régression; toutes les variables de contrôle des régressions précédentes sont aussi considérées. Tout d'abord, on calcule les effets moyens du cancer propres à un groupe; ensuite, les différences entre ces effets et l'effet moyen général (illustrées dans les tableaux 3 et 4) font l'objet d'un test de signification statistique. On utilise un modèle incluant les termes d'interaction plutôt que d'estimer des modèles distincts pour chaque catégorie afin d'obtenir la constante de l'incidence du cancer pour tous les sous-groupes en utilisant les pondérations CEM obtenues pour l'ensemble de l'échantillon. Parce que l'incidence pondérée du cancer est

25. Par ailleurs, les tableaux A.4-1 et A.4-2 en annexe présentent des résultats au moyen d'un échantillon de personnes qui travaillent dans les années T et $T-1$, et les tableaux A.5-1 et A.5-2 en annexe présentent les résultats au moyen d'un échantillon de personnes qui travaillent dans les années T , $T-1$, $T-2$ et $T-3$. Comme on pouvait s'y attendre, si l'échantillon est restreint aux personnes qui participent plus activement au marché du travail avant le diagnostic de cancer, l'effet négatif du cancer sur les gains est susceptible d'être plus important, mais les différences sont faibles. L'ensemble des résultats montre les mêmes tendances que dans le tableau 5. Il n'y a pas de différence dans l'effet négatif du cancer sur l'emploi peu importe les restrictions appliquées à la situation d'emploi des personnes avant le diagnostic de cancer.

26. L'étude ne compare pas les effets du cancer pour une sous-catégorie avec ceux de la catégorie de base puisque l'hétérogénéité des caractéristiques préalables au cancer entre l'échantillon de survivants du cancer et l'échantillon de contrôle a été délibérément supprimée en appliquant des pondérations d'appariement.

la même pour l'ensemble de l'échantillon et pour chaque sous-groupe (p. ex., les personnes âgées de 35 à 44 ans, les personnes ayant un grade universitaire, etc.), ces différences entre les effets estimés propres à un groupe et à l'ensemble de l'échantillon peuvent être interprétées comme reflétant les différences dans les caractéristiques (préalables au cancer) entre les sous-groupes et l'ensemble de l'échantillon plutôt qu'à titre de différences dans l'incidence du cancer.

Tableau 5
Effet du cancer sur les gains (en dollars) en tenant compte des caractéristiques individuelles non observables

	Sans tenir compte du travail dans la période T+j			En tenant compte du travail dans la période T+j		
	T+1	T+2	T+3	T+1	T+2	T+3
Régression pondérée de la DDD						
Cancer						
Coefficient	-4 967,85 ***	-3 858,98 ***	-3 386,03 ***	-4 269,39 ***	-2 428,57 ***	-1 763,90 **
Erreur-type	780,07	748,05	735,03	703,10	625,02	639,44
n (échantillon pondéré)	596 374	595 210	594 014	577 036	564 400	552 832
Régression linéaire pondérée (volet supérieur du tableau 4)						
Cancer						
Coefficient	-5 079,37 ***	-3 980,47 ***	-3 756,72 ***	-4 674,72 ***	-2 940,20 **	-2 427,52 **
Erreur-type	902,42	878,38	810,23	953,02	939,40	880,57
n (échantillon pondéré)	596 374	595 210	594 014	577 036	564 400	552 832
MCO de l'échantillon non pondéré						
Cancer						
Coefficient	-5 233,87 ***	-4 076,90 ***	-3 821,77 ***	-4 833,49 ***	-3 047,78 **	-2 472,27 **
Erreur-type	902,52	876,65	814,36	952,25	935,19	881,80
n (échantillon non pondéré)	1 166 397	1 163 918	1 161 372	1 129 985	1 107 613	1 088 126

*** valeur significativement différente de l'estimation pour la catégorie de référence (p<0,001)

** valeur significativement différente de l'estimation pour la catégorie de référence (p<0,01)

* valeur significativement différente de l'estimation pour la catégorie de référence (p<0,05)

Notes : Toutes les régressions sont conditionnelles à l'ensemble complet de variables de contrôle qui figure à la sous-section 2.3. DDD : différence dans les différences; n : nombre d'observations; MCO : moindres carrés ordinaires; T : année du diagnostic de cancer dans l'échantillon de survivants du cancer. Calculs de l'auteur.

Source : Statistique Canada, Recensement de 1991 – Fichier de données longitudinales sur la main-d'œuvre.

Les différences dans les effets du cancer (ΔEC) entre chaque sous-groupe ($[EC | x = 1]$) et l'effet général estimé pour l'ensemble de l'échantillon ($[EC]$) sont présentés dans les tableaux 6-1 et 6-2. Voici les hypothèses du test :

$$[\hat{pr}(\text{travail} | \text{cancer} = 1 \text{ et } x = 1) - \hat{pr}(\text{travail} | \text{cancer} = 0 \text{ et } x = 1)] - [\hat{pr}(\text{travail} | \text{cancer} = 1) - \hat{pr}(\text{travail} | \text{cancer} = 0)] = [EC_{\text{travail}} | x = 1] - [EC_{\text{travail}}] = 0 \quad (1)$$

$$[E(\text{gains} | \text{cancer} = 1 \text{ et } x = 1) - E(\text{gains} | \text{cancer} = 0 \text{ et } x = 1)] - [E(\text{gains} | \text{cancer} = 1) - E(\text{gains} | \text{cancer} = 0)] = [EC_{\text{gains}} | x = 1] - [EC_{\text{gains}}] = 0 \quad (2)$$

Le tableau 6-1 présente les ΔEC pour la probabilité de travailler dans la période $T+1$ et $T+3$, ainsi que les résultats de l'essai pour l'équation (1) pour chaque ensemble de caractéristiques de l'échantillon. Les effets du cancer sur la probabilité de travailler pour les groupes d'âge plus jeunes (de 25 à 34 ans et de 35 à 44 ans) sont supérieurs de 1 à 2 points de pourcentage dans l'année $T+1$ par rapport à l'effet du cancer estimé pour l'ensemble de l'échantillon (-0,030 dans le tableau 3), où la moyenne d'âge est de 48²⁷ : cette différence est statistiquement significative au niveau de 10 %. Les ΔEC pour les deux groupes d'âge les plus jeunes sont plus importantes et plus statistiquement significatives dans l'année $T+3$ que dans l'année $T+1$. Toutefois, les ΔEC pour les groupes d'âge plus avancés (de 45 à 54 ans et de 55 à 61 ans) ne sont pas statistiquement significatives pour l'une ou l'autre des périodes. Ce résultat laisse entendre que, pour les personnes atteintes d'un cancer lorsqu'elles sont plus jeunes, l'ampleur de l'effet négatif du cancer sur la probabilité de travailler est moins importante que l'effet moyen estimé pour l'ensemble de l'échantillon, particulièrement à long terme.

Les ΔEC sont aussi statistiquement significatives dans certaines catégories de niveau de scolarité. Dans l'année $T+1$, la seule ΔEC significative est observée chez les survivants du cancer ayant un grade universitaire pour qui l'effet du cancer sur la probabilité de travailler est supérieur de 2 points de pourcentage à l'effet général. Dans l'année $T+3$, la seule ΔEC significative s'observe pour les personnes sans diplôme d'études secondaires; elle est inférieure de 1,9 point de pourcentage à l'effet général du cancer sur la probabilité de travailler dans l'année $T+3$. En d'autres mots, dans l'année $T+3$, l'effet négatif du cancer sur la situation d'emploi chez les personnes sans diplôme d'études secondaires est plus important en fait d'ampleur (-0,048 - 0,019 = -0,067) que l'effet moyen général (-0,048)²⁸.

Chez les survivants du cancer dans le quintile inférieur (20 %) de la répartition des gains dans l'année $T-1$, la ΔEC dans l'année $T+1$ est de -0,042, ce qui signifie que l'effet négatif est supérieur de 4,2 points de pourcentage à l'effet général (-0,030). En revanche, pour ceux de la partie supérieure (60 %) de la répartition des gains dans l'année $T-1$, la ΔEC dans l'année $T+1$ (0,023) est positive et statistiquement significative; par conséquent, les effets négatifs du cancer sur la situation d'emploi pour ces groupes sont près de zéro dans l'année $T+1$ (-0,030 + 0,023 = -0,007). Dans l'année $T+3$, les ΔEC correspondantes sont beaucoup moins importantes et ne sont pas statistiquement significatives.

27. En d'autres mots, à la période $T+1$, la probabilité moyenne de travailler pour les personnes de 25 à 34 ans dans l'année du diagnostic de cancer (T) est inférieure de 1 point de pourcentage (-0,030 + 0,021 = -0,009) à la probabilité moyenne de travailler de leurs homologues dans l'échantillon de contrôle.

28. Dans une certaine mesure, cela s'explique peut-être parce que les personnes sans diplôme d'études secondaires sont plus susceptibles de recevoir un diagnostic de cancer ayant un faible taux de survie (ou un cancer plus grave) comme un cancer du poumon. En règle générale, il y a un lien négatif entre le niveau de scolarité et le tabagisme (Maralani, 2013).

Tableau 6-1
Hétérogénéité des effets du cancer sur la situation d'emploi, selon diverses
caractéristiques de l'échantillon

	Différences dans les effets du cancer (ΔEC)	
	points de pourcentage	erreur-type
$T+1^1$		
Âge dans l'année T		
25 à 34	0,021 †	0,013
35 à 44	0,013 †	0,007
45 à 54	-0,005	0,006
55 à 61	-0,012	0,010
Plus haut niveau de scolarité atteint		
Sans diplôme d'études secondaires	-0,010	0,008
Études secondaires, avec ou sans certificat professionnel	0,008	0,006
Études postsecondaires non universitaires	-0,015	0,011
Grade universitaire	0,020 †	0,012
Quintiles des gains dans l'année $T-1$		
Inférieur	-0,042 ***	0,012
Deuxième	-0,013	0,010
Troisième	0,023 **	0,009
Quatrième	0,023 *	0,009
Supérieur	0,023 **	0,009
$T+3^2$		
Âge dans l'année T		
25 à 34	0,036 *	0,016
35 à 44	0,020 *	0,010
45 à 54	-0,012	0,009
55 à 61	-0,012	0,013
Plus haut niveau de scolarité atteint		
Sans diplôme d'études secondaires	-0,019 †	0,011
Études secondaires, avec ou sans certificat professionnel	0,004	0,009
Études postsecondaires non universitaires	0,009	0,014
Grade universitaire	0,017	0,016
Quintiles des gains dans l'année $T-1$		
Inférieur	-0,013	0,014
Deuxième	-0,019	0,013
Troisième	0,010	0,013
Quatrième	0,011	0,014
Supérieur	0,017	0,013

*** valeur significativement différente de l'estimation pour la catégorie de référence ($p < 0,001$)

** valeur significativement différente de l'estimation pour la catégorie de référence ($p < 0,01$)

* valeur significativement différente de l'estimation pour la catégorie de référence ($p < 0,05$)

† valeur significativement différente de l'estimation pour la catégorie de référence ($p < 0,10$)

1. Effet moyen dans l'échantillon dans l'année $T+1$, -0,030.

2. Effet moyen dans l'échantillon dans l'année $T+3$, -0,048.

Notes : Les résultats sont calculés en fonction du modèle probit qui utilise des pondérations de la méthode Coarsened Exact Matching (CEM). Pour estimer les effets du cancer par tranche d'âge, les catégories d'âge sont incluses plutôt que l'âge et l'âge au carré. Autrement, toutes les régressions sont conditionnelles aux interactions entre le cancer et un ensemble de variables catégoriques et le jeu complet de variables de contrôle qui figure à la sous-section 2.3. T : année du diagnostic de cancer pour l'échantillon de survivants du cancer. $T-1$: T moins un. Calculs de l'auteur.

Source : Statistique Canada, Recensement de 1991 – Fichier de données longitudinales sur la main-d'œuvre.

Le tableau 6-2 présente les ΔEC pour les gains annuels et les résultats de l'essai de l'équation (2) dans l'année $T+1$ et $T+3$. Les pertes de gains subies après un diagnostic de cancer sont généralement petites pour les survivants du cancer compris dans le groupe d'âge des 25 à 34 ans ($-5\,079,37 + 4\,337,53 = -741,87$). Pour ce groupe d'âge, l'effet négatif du cancer sur la probabilité de travailler est plus petit que l'effet global de l'âge. Toutefois, les tendances des effets propres aux groupes sur les gains du cancer sont moins évidentes que celles observées pour la situation d'emploi montrées au tableau 6-1.

Tableau 6-2

Hétérogénéité des effets du cancer sur les gains (y compris les gains nuls), selon les caractéristiques de l'échantillon

	Différences dans les effets du cancer (ΔEC)	
	dollars	erreur-type
$T+1^1$		
Âge dans l'année T		
25 à 34	4 337,53 **	1 581,50
35 à 44	-982,43	1 098,64
45 à 54	-1 361,36	1 035,92
55 à 61	1 700,80	1 991,31
Plus haut niveau de scolarité atteint		
Sans diplôme d'études secondaires	400,24	1 165,61
Études secondaires, avec ou sans certificat professionnel	149,84	995,47
Études postsecondaires non universitaires	-903,13	1 417,18
Grade universitaire	78,07	3 881,43
Quintiles des gains dans l'année $T-1$		
Inférieur	2 802,57 **	1 064,16
Deuxième	666,51	1 024,41
Troisième	-43,56	1 082,79
Quatrième	-762,02	1 206,78
Supérieur	-3 674,27	3 833,93
$T+3^2$		
Âge dans l'année T		
25 à 34	5 020,94 **	1 740,39
35 à 44	-1 181,10	1 033,17
45 à 54	-720,37	964,59
55 à 61	752,23	1 639,10
Plus haut niveau de scolarité atteint		
Sans diplôme d'études secondaires	-597,08	1 022,11
Études secondaires, avec ou sans certificat professionnel	733,24	920,51
Études postsecondaires non universitaires	234,04	1 334,48
Grade universitaire	-1 130,42	3 287,29
Quintiles des gains dans l'année $T-1$		
Inférieur	2 081,94 *	977,97
Deuxième	1 322,43	954,97
Troisième	849,97	1 079,76
Quatrième	-226,33	1 284,24
Supérieur	-5 088,98	3 292,04

** valeur significativement différente de l'estimation pour la catégorie de référence ($p < 0,01$)

* valeur significativement différente de l'estimation pour la catégorie de référence ($p < 0,05$)

1. Effet moyen dans l'échantillon dans l'année $T+1$, -5 079,37.

2. Effet moyen dans l'échantillon dans l'année $T+3$, -3 756,72.

Notes : Les résultats sont calculés en fonction du modèle de régression linéaire qui utilise des pondérations de la méthode Coarsened Exact Matching (CEM). Pour estimer les effets du cancer par tranche d'âge, les catégories d'âge sont incluses plutôt que l'âge et l'âge au carré. Autrement, toutes les régressions sont conditionnelles aux interactions entre le cancer et un ensemble de variables catégoriques et l'ensemble complet de variables de contrôle qui figure à la sous-section 2.3. T : année du diagnostic de cancer pour l'échantillon de survivants du cancer. $T-1$: T moins un. Calculs de l'auteur.

Source : Statistique Canada, Recensement de 1991 – Fichier de données longitudinales sur la main-d'œuvre.

5 Conclusion

Des diagnostics plus rapides, des améliorations dans le traitement du cancer et de meilleurs soins de suivi ont considérablement augmenté le nombre de survivants du cancer au cours des deux dernières décennies. Alors que le taux de survie au cancer augmente, les chercheurs s'intéressent beaucoup à la vie dans différents domaines après le traitement du cancer. Pour leur part, les économistes s'intéressent particulièrement aux effets du cancer sur les choix et les résultats sur le marché du travail des survivants du cancer.

À l'aide de données de couplage uniques provenant du Recensement du Canada de 1991, des dossiers administratifs sur le cancer, des dossiers de mortalité et des dossiers longitudinaux de l'impôt sur le revenu, l'étude évalue les effets du cancer sur certains résultats sur le marché du travail des survivants du cancer en comparant leurs résultats à ceux du groupe de contrôle composé de personnes n'ayant jamais reçu de diagnostic de cancer. De manière générale, le cancer réduit la probabilité de travailler dans la première année suivant l'année du diagnostic de 3 points de pourcentage par rapport à l'échantillon de contrôle. La perte de gains moyenne des survivants du cancer dans la première année après le diagnostic est de l'ordre de 12 %, lorsque les transitions de l'emploi au non-emploi sont prises en compte en incluant dans l'échantillon les personnes dont les gains sont nuls. Lorsqu'on considère seulement les personnes dont la rémunération est différente de zéro, la perte moyenne estimée des gains des survivants du cancer est de l'ordre de 10 % à 11 %. Dans la deuxième et la troisième années après l'année du diagnostic, le cancer réduit la probabilité de travailler chez les survivants du cancer. Toutefois, l'écart des gains entre les survivants et les personnes du groupe de contrôle s'amenuise, particulièrement lorsque les survivants du cancer continuent de travailler après avoir reçu un diagnostic de cancer. En plus de tenir compte des différences dans les caractéristiques observables entre les personnes atteintes du cancer et celles qui n'ont jamais reçu de diagnostic de cancer, il semble important de prendre en compte les associations simultanées (1) entre les caractéristiques individuelles intemporelles non observables et le cancer, et (2) entre ces caractéristiques et les résultats sur le marché du travail pour estimer les effets à long terme du cancer sur les gains.

Les constatations de l'enquête laissent aussi entendre que les effets du cancer sur les résultats sur le marché du travail diffèrent pour diverses catégories de taux de survie, lesquelles servent d'approximation de la gravité du cancer dans cette étude. La probabilité de travailler et les gains annuels sont considérablement inférieurs chez les survivants de la catégorie du faible taux de survie que chez ceux de la catégorie du taux de survie élevé, pour toutes les périodes après le cancer. À long terme, l'effet négatif du cancer sur la situation d'emploi des personnes ayant reçu un diagnostic avant l'âge moyen dans l'échantillon est moins important que l'effet moyen observé dans l'ensemble de l'échantillon de survivants du cancer. À l'inverse, l'effet négatif du cancer sur la situation d'emploi des survivants du cancer sans diplôme d'études secondaires est plus important que l'effet moyen observé dans l'ensemble de l'échantillon.

Dans l'ensemble, les constatations laissent entendre qu'à long terme le cancer est plus susceptible d'influer sur la situation d'emploi des survivants (marge extensive de résultats sur le marché du travail) que sur leurs gains (marge intensive de résultats sur le marché du travail). L'amélioration de la compréhension générale du rôle que jouent l'offre et la demande dans la réduction de la probabilité d'emploi des survivants du cancer par rapport à ceux qui n'ont jamais reçu de diagnostic de cancer est une importante avenue de recherche.

6 Annexe

Tableau A.1

Sièges du cancer, catégorie de taux de survie et enregistrements des décès

	Siège du cancer		Catégorie de taux de survie	Enregistrement des décès constatés en 2006 dans chaque groupe
	Nombre de personnes	Part		
	nombre		pourcentage	
Taux élevé de survie	39,02	...
Thyroïde	138	2,66
Prostate	305	5,88
Testicules	56	1,08
Mélanome de la peau	241	4,65
Sein	1 117	21,54
Corps utérin	136	2,62
Lymphome hodgkinien	30	0,58
Décès	19,13
Taux moyen de survie	32,00	...
Leucémie lymphoïde chronique	32	0,62
Col utérin	409	7,89
Vessie (y compris in situ)	147	2,84
Rein et bassinnet du rein	145	2,80
Tissu mou	39	0,75
Larynx	51	0,98
Rectum	177	3,41
Côlon	298	5,75
Lymphome non hodgkinien	229	4,42
Oral (cavité buccale et pharynx)	132	2,55
Décès	34,30
Faible taux de survie	21,04	...
Ovaires	108	2,08
Myélome multiple	43	0,83
Leucémie (à l'exclusion de la leucémie lymphoïde chronique)	71	1,37
Estomac	74	1,43
Cerveau	108	2,08
Foie	35	0,68
Poumon et bronches	520	10,03
Œsophage	37	0,71
Pancréas	95	1,83
Décès	81,48
Valeur manquante	7,95	...
Autres	412	7,95
Décès	42,48
Total	5 185	100	100	...

... n'ayant pas lieu de figurer

Notes : Les données du Recensement de 1991 – Fichier de données longitudinales sur la main-d'œuvre contiennent au total 5 185 personnes qui ont reçu un premier diagnostic de cancer dans la période de 1992 à 2000, et qui étaient âgées de 61 ans ou moins dans l'année du diagnostic. Les sièges du cancer sont définis en fonction des définitions regroupées du programme de la surveillance, de l'épidémiologie et des résultats définitifs (SEER) pour la Classification internationale des maladies : Oncologie, deuxième et troisième éditions (CIM-O-2/3). Les chiffres ayant été arrondis, la somme peut ne pas correspondre aux totaux indiqués. Calculs de l'auteur.

Source : Statistique Canada, Recensement de 1991 – Fichier de données longitudinales sur la main-d'œuvre.

Tableau A.2
Statistiques sommaires pour les autres caractéristiques du groupe de survivants du cancer et du groupe de contrôle avant et après l'appariement et l'élagage

Variables ¹	Échantillon de	Échantillon de contrôle		Différence non appariée ²
	survivants du cancer	Non apparié	Apparié et repondéré	
			dollars	
Gains dans l'année <i>T</i> -1 (moyenne)	43 445	46 454	43 374	-3 009 *
		pourcentage		points de pourcentage
Quintiles des gains dans l'année <i>T</i>-1				
Inférieur	23,45	19,99	23,10	3,46 ***
Deuxième	21,79	20,00	21,64	1,80 *
Troisième	19,25	20,00	20,02	-0,75
Quatrième	17,60	20,01	17,85	-2,41 **
Supérieur	17,91	20,00	17,39	-2,10 **
			dollars	
Gains dans l'année <i>T</i> -2 (moyenne)	43 689	45 493	42 902	-1 804
		pourcentage		points de pourcentage
Quintiles des gains dans l'année <i>T</i>-2				
Inférieur	24,37	19,99	22,84	4,38 ***
Deuxième	20,83	20,00	21,32	0,83
Troisième	18,75	20,00	20,18	-1,25
Quatrième	17,10	20,01	17,92	-2,91 ***
Supérieur	18,95	20,00	17,75	-1,06
Emploi autonome dans l'année <i>T</i>				
Non	82,13	80,54	80,62	1,60 *
Oui	15,94	16,73	17,14	-0,79
Valeur manquante	1,93	2,74	2,25	-0,81 *
Union dans l'année <i>T</i>	41,16	40,87	40,82	0,29
État matrimonial dans l'année <i>T</i>				
Célibataire	23,57	22,92	22,37	0,64
En couple	73,55	72,96	74,18	0,59
Valeur manquante	2,89	4,12	3,45	-1,23 **
Incapacités ou handicaps de longue durée au Recensement de 1991				
Non	96,65	96,81	96,53	-0,16
Oui	2,43	2,25	2,56	0,17
Valeur manquante	0,92	0,94	0,91	-0,01
			nombre	
<i>n</i>	2 597	1 228 551	624 835	...

... n'ayant pas lieu de figurer

*** valeur significativement différente de l'estimation pour la catégorie de référence ($p < 0,001$)

** valeur significativement différente de l'estimation pour la catégorie de référence ($p < 0,01$)

* valeur significativement différente de l'estimation pour la catégorie de référence ($p < 0,05$)

1. Les variables ne sont pas utilisées pour l'appariement, mais elles font partie de l'analyse de régression.

2. Cette colonne affiche la différence des moyennes (proportion) entre le groupe de survivants du cancer et le groupe de contrôle avant l'appariement; des symboles normalisés indiquent les différences significatives du test-t entre ces deux groupes.

Note : Les gains dans l'année *T*-2 (*T* moins 2) ne font pas partie de l'analyse de régression. *T* : année du diagnostic du cancer pour l'échantillon de survivants du cancer; *n* : nombre d'observations. *T*-1 : *T* moins 1. Calculs de l'auteur.

Source : Statistique Canada, Recensement de 1991 – Fichier de données longitudinales sur la main-d'œuvre.

Tableau A.3**Moyenne non rajustée de l'emploi et des gains dans les périodes $T+1$, $T+2$ et $T+3$**

	$T+1$	$T+2$	$T+3$
Échantillon de survivants du cancer			
Tous les survivants du cancer			
Pourcentage de personnes qui travaillent	92,38	89,03	85,28
Moyenne des gains (en dollars)	37 124	37 517	36 577
<i>n</i>	2 573	2 561	2 540
Catégorie du taux de survie élevé			
Pourcentage de personnes qui travaillent	93,26	88,59	85,27
Moyenne des gains (en dollars)	38 149	38 412	37 266
<i>n</i>	1 291	1 288	1 276
Catégorie du taux de survie moyen			
Pourcentage de personnes qui travaillent	92,10	90,81	88,66
Moyenne des gains (en dollars)	37 424	37 891	37 377
<i>n</i>	886	881	873
Catégorie du faible taux de survie			
Pourcentage de personnes qui travaillent	84,47	79,70	69,80
Moyenne des gains (en dollars)	29 350	29 179	28 517
<i>n</i>	206	202	202
Valeur manquante			
Pourcentage de personnes qui travaillent	96,32	93,68	86,24
Moyenne des gains (en dollars)	37 182	38 579	36 837
<i>n</i>	190	190	189
Échantillon de contrôle			
Non apparié, sans cancer			
Pourcentage de personnes qui travaillent	96,85	95,19	93,77
Moyenne des gains (en dollars)	47 145	47 296	47 402
<i>n</i>	1 220 406	1 214 199	1 209 296
Apparié sans cancer			
Pourcentage de personnes qui travaillent	96,75	94,87	93,19
Moyenne des gains (en dollars)	46 146	45 933	45 691
<i>n</i>	621 050	617 958	615 693

Notes : Les calculs des gains moyens comprennent les gains nuls des personnes qui n'ont pas travaillé durant la période. *n* : nombre d'observations; *T* : année du diagnostic du cancer pour l'échantillon de survivants du cancer. Calculs de l'auteur.

Source : Statistique Canada, Recensement de 1991 – Fichier de données longitudinales sur la main-d'œuvre.

Tableau A.4-1

Effet du cancer sur l'emploi pour l'échantillon limité aux personnes travaillant dans les années T et $T-1$

	Travaille dans chaque période : $\Pr(\text{travail}_{T+j}=1)$			Travaille continuellement jusqu'à chaque période : $\Pr(\text{travail}_{T+j}=1 \mid \text{travail}_{T+j-1}=1, \text{travail}_{T+j-2}=1)$		
	$T+1$	$T+2$	$T+3$	$T+1$	$T+2$	$T+3$
Cancer						
Coefficient	-0,033 ***	-0,037 ***	-0,049 ***	-0,033 ***	-0,017 ***	-0,014 **
Erreur-type	0,005	0,006	0,007	0,005	0,005	0,005
n (échantillon pondéré)	612 760	610 639	608 809	612 760	589 413	567 726

*** valeur significativement différente de l'estimation pour la catégorie de référence ($p < 0,001$)

** valeur significativement différente de l'estimation pour la catégorie de référence ($p < 0,01$)

Notes : Ce tableau reproduit le volet supérieur du tableau 3 pour l'échantillon limité aux personnes qui travaillent dans les années T et $T-1$ (T moins 1). Les tailles des échantillons appariés dans l'année T sont 2 671 pour l'échantillon de survivants du cancer et 651 506 pour l'échantillon de contrôle n : nombre d'observations; T : année du diagnostic du cancer pour l'échantillon de survivants du cancer. Calculs de l'auteur.

Source : Statistique Canada, Recensement de 1991 – Fichier de données longitudinales sur la main-d'œuvre.

Tableau A.4-2

Effet du cancer sur les gains (en dollars) pour l'échantillon limité aux personnes travaillant dans les années T et $T-1$

	Sans tenir compte du travail dans l'année $T+j$			En tenant compte du travail dans l'année $T+j$		
	$T+1$	$T+2$	$T+3$	$T+1$	$T+2$	$T+3$
Régression pondérée de la DDD						
Cancer						
Coefficient	-4 820,26 ***	-3 768,02 ***	-3 297,79 ***	-4 132,46 ***	-2 389,55 ***	-1 710,38 **
Erreur-type	759,68	728,78	717,45	689,12	612,51	628,43
n (échantillon pondéré)	621 028	619 714	618 473	599 218	585 719	573 801
Régression linéaire pondérée						
Cancer						
Coefficient	-4 977,03 ***	-3 911,01 ***	-3 715,15 ***	-4 577,64 ***	-2 894,82 **	-2 364,64 **
Erreur-type	881,84	859,15	794,58	936,67	921,15	865,34
n (échantillon pondéré)	621 028	619 714	618 473	599 218	585 719	573 801
MCO de l'échantillon non pondéré						
Cancer						
Coefficient	-5 117,83 ***	-4 004,28 ***	-3 760,47 ***	-4 728,78 ***	-2 998,04 **	-2 396,81 **
Erreur-type	882,29	858,03	798,40	936,18	917,65	866,40
n (échantillon pondéré)	1 202 688	1 199 828	1 197 110	1 161 723	1 137 848	1 117 645

*** valeur significativement différente de l'estimation pour la catégorie de référence ($p < 0,001$)

** valeur significativement différente de l'estimation pour la catégorie de référence ($p < 0,01$)

Notes : Ce tableau reproduit le tableau 5 pour l'échantillon limité aux personnes qui travaillent dans les années T et $T-1$ (T moins 1). Les tailles des échantillons appariés dans l'année T sont 2 671 pour l'échantillon de survivants du cancer et 651 506 pour l'échantillon de comparaison. DDD : différence dans les différences; n : nombre d'observations; MCO : moindres carrés ordinaires; T : année du diagnostic du cancer pour l'échantillon de survivants du cancer. Calculs de l'auteur.

Source : Statistique Canada, Recensement de 1991 – Fichier de données longitudinales sur la main-d'œuvre.

Tableau A.5-1

Effet du cancer sur l'emploi pour l'échantillon limité aux personnes travaillant dans les années T , $T-1$, $T-2$ et $T-3$

	Travaille dans chaque période : Pr($\text{travail}_{T+j}=1$)			Travaille continuellement jusqu'à chaque période : Pr($\text{travail}_{T+j}=1 \mid \text{travail}_{T+j-1}=1,$ $\text{travail}_{T+j-2}=1$)		
	$T+1$	$T+2$	$T+3$	$T+1$	$T+2$	$T+3$
Cancer						
Coefficient	-0,031 ***	-0,038 ***	-0,049 ***	-0,031 ***	-0,019 ***	-0,014 **
Erreur-type	0,005	0,006	0,007	0,005	0,005	0,005
n (échantillon pondéré)	567 291	565 602	564 007	567 291	548 372	529 543

*** valeur significativement différente de l'estimation pour la catégorie de référence ($p < 0,001$)

** valeur significativement différente de l'estimation pour la catégorie de référence ($p < 0,01$)

Notes : Ce tableau reproduit le volet supérieur du tableau 3 pour l'échantillon limité aux personnes qui travaillent dans les années T , $T-1$ (T moins 1), $T-2$ (T moins 2) et $T-3$ (T moins 3). Les tailles des échantillons appariés dans l'année T sont 2 518 pour l'échantillon de survivants du cancer et 601 590 pour l'échantillon de contrôle n : nombre d'observations; T : année du diagnostic du cancer pour l'échantillon de survivants du cancer. Calculs de l'auteur.

Source : Statistique Canada, Recensement de 1991 – Fichier de données longitudinales sur la main-d'œuvre.

Tableau A.5-2

Effet du cancer sur les gains (en dollars) pour l'échantillon limité aux personnes travaillant dans les années T , $T-1$, $T-2$ et $T-3$

	Sans tenir compte du travail dans l'année			En tenant compte du travail dans l'année		
	$T+1$	$T+2$	$T+3$	$T+1$	$T+2$	$T+3$
Régression pondérée de la DDD						
Cancer						
Coefficient	-5 125,75 ***	-4 011,63 ***	-3 582,76 ***	-4 403,17 ***	-2 526,39 ***	-1 941,73 **
Erreur-type	801,48	767,53	749,09	721,18	640,12	647,78
n (échantillon pondéré)	574 695	573 679	572 570	556 933	544 931	533 775
Régression linéaire pondérée						
Cancer						
Coefficient	-5 314,87 ***	-4 201,35 ***	-4 021,46 ***	-4 931,66 ***	-3 147,42 **	-2 679,87 **
Erreur-type	923,75	899,01	824,45	975,52	962,15	896,13
n (échantillon pondéré)	574 695	573 679	572 570	556 933	544 931	533 775
MCO de l'échantillon non pondéré						
Cancer						
Coefficient	-5 449,73 ***	-4 333,37 ***	-4 120,28 ***	-5 062,16 ***	-3 283,89 ***	-2 756,85 **
Erreur-type	922,70	896,64	828,60	973,46	957,01	896,97
n (échantillon non pondéré)	1 134 030	1 131 814	1 129 527	1 100 455	1 079 254	1 060 522

*** valeur significativement différente de l'estimation pour la catégorie de référence ($p < 0,001$)

** valeur significativement différente de l'estimation pour la catégorie de référence ($p < 0,01$)

Notes : Ce tableau reproduit le tableau 5 pour l'échantillon limité aux personnes qui travaillent dans les années T , $T-1$ (T moins un), $T-2$ (T moins 2) et $T-3$ (T moins 3). Les tailles des échantillons appariés dans l'année T sont 2 518 pour l'échantillon de survivants du cancer et 601 590 pour l'échantillon de contrôle DDD : différence dans les différences; n : nombre d'observations; MCO : moindres carrés ordinaires; T : année du diagnostic de cancer pour l'échantillon de survivants du cancer. Calculs de l'auteur.

Source : Statistique Canada, Recensement de 1991 – Fichier de données longitudinales sur la main-d'œuvre.

Bibliographie

American Cancer Society. 2012. *Cancer Facts & Figures 2012*. Atlanta : American Cancer Society.

Bradley, C.J., H.L. Bednarek et D. Neumark. 2002. « Breast cancer survival, work and earnings ». *Journal of Health Economics* 21 (5) : 757 à 779.

Bradley, C.J., D. Neumark, H.L. Bednarek et M. Schenk. 2005. « Short-term effects of breast cancer on labour market attachment: results from a longitudinal study ». *Journal of Health Economics* 24 (1) : 137 à 160.

Bradley, C.J., D. Neumark, Z. Luo et M. Schenk. 2007. « Employment and cancer: findings from a longitudinal study of breast and prostate cancer survivors ». *Cancer Investigations* 25 (1) : 47 à 54.

Carpenter, M., M.E. Fair, C. Poliquin et P. Lalonde. 2008. *Base canadienne de données sur le cancer, 1969 à 1991 Historique et développement*. Section de la recherche sur l'hygiène du travail et de l'environnement, rapport n° 16. Division de la statistique de la santé. Ottawa : Statistique Canada.

Centers for Disease Control and Prevention. 2011. « Cancer survivors—United States, 2007 ». *Morbidity and Mortality Weekly Report* 2011 60 : 269 à 272.

Ellison, L.F., H. Bryant, G. Lockwood et L. Shack. 2011. « Analyses de la survie conditionnelle selon le siège du cancer ». *Rapports sur la santé* 22 (2) : 21 à 25. N° 82-003-XPE au catalogue de Statistique Canada.

Heckman, J., H. Ichimura, J. Smith et P. Todd. 1998. « Characterizing selection bias using experimental data ». *Econometrica* 66 (5) : 1017 à 1098.

Heckman, J., H. Ichimura et P. Todd. 1997. « Matching as an econometric evaluation estimator ». *Review of Economic Studies* 65 : 261 à 294.

Ho, D., K. Imai, G. King et E. Stuart. 2007. « Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference ». *Political Analysis* 15 (3) : 199 à 236.

Iacus, S.M., G. King et G. Porro. 2011. « Multivariate matching methods that are monotonic imbalance bounding ». *Journal of the American Statistical Association* 106 (493) : 345 à 361.

Iacus, S.M., G. King et G. Porro. 2012. « Causal inference without balance checking : Coarsened Exact Matching ». *Political Analysis* 20 (1) : 1 à 24.

Imbens, G.W., et J.M. Wooldridge. 2009. « Recent developments in the econometrics of program evaluation ». *Journal of Economic Literature* 47 (1) : 5 à 86.

Jones, A.M., N. Rice et P. Rosa Dias. 2011. « Long-term effects of school quality on health and lifestyle: evidence from comprehensive schooling reforms in England ». *Journal of Human Capital* 5 (3) : 342 à 376.

Maralani, V. 2013. « Educational inequalities in smoking: The role of initiation versus quitting ». *Social Science & Medicine* 84 : 129 à 137.

Moran, J.R., P.F. Short et C.S. Hollenbeak. 2011. « Long-term employment effects of surviving cancer ». *Journal of Health Economics* 30 (3) : 505 à 514

Organisation de coopération et de développement économiques (OCDE). 2007. *Coup d'œil sur la santé 2007 : indicateurs de l'OCDE*. Paris : publications de l'OCDE.

Société canadienne du cancer (SCC). 2012. *Statistiques canadiennes sur le cancer 2012*. Comité directeur des statistiques sur le cancer de la Société canadienne du cancer. Toronto : Société canadienne du cancer.

Statistique Canada. 2008. *Guide du système du Registre canadien du cancer, édition 2007*. Équipe de la refonte du Registre canadien du cancer. Produit n° 82-225-X au catalogue de Statistique Canada. Ottawa : Statistique Canada.

Wilkins, R., M. Tjepkema, C. Mustard et R. Choinière. 2008. « Étude canadienne de suivi de la mortalité selon le recensement, 1991 à 2001 ». *Rapports sur la santé* 19 (3) : 25 à 43. Produit n° 82-003-X au catalogue de Statistique Canada.