

PERSPECTIVES DE L'IMT

Évaluer les résultats sur le marché du travail des étudiants postsecondaires et apprentis : une exploration des données disponibles

Principaux problèmes et constats

- La Plateforme de liens longitudinaux entre l'éducation et le marché du travail (PLLEMT) est un environnement de données unique qui nous permettra de mieux comprendre les transitions des Canadiens entre les programmes de formation et d'enseignement et le marché du travail.
- En ce moment, la PLLEMT offre un accès à trois ensembles de données administratives anonymisées, liés dans le temps : 1) le Système d'information sur les étudiants postsecondaires (SIEP); 2) le Système d'information sur les apprentis inscrits (SIAI); et 3) le Fichier des familles T1 des formulaires fiscaux (FFT1).
- Grâce à ses observations longitudinales, la PLLEMT offre une occasion extraordinaire d'analyser comment les individus s'en tirent en matière de revenu et de croissance du revenu, à la suite de leur participation à un programme universitaire, collégial ou d'apprentissage.
- À l'aide de données administratives, la PLLEMT offre une couverture complète des inscriptions à des programmes d'enseignement postsecondaire public et de formation d'apprenti au Canada, en plus d'un nombre important de liens supplémentaires potentiels. Les chercheurs seront ainsi capables de découvrir de nouveaux facteurs de succès des transitions vers le marché du travail.
- Les ensembles de données existants dans la PLLEMT ont toutefois certaines limites, comme le manque de données sur l'emploi occupé (p. ex. : la profession) et peu d'indicateurs sur la qualité du travail (p. ex. : aucune information sur le nombre d'heures travaillées, les avantages sociaux ou la culture du milieu de travail). De plus, pour l'ensemble du Canada, les données du SIEP sont seulement disponibles pour les années scolaires 2009-2010 à 2015-2016, et celles du SIAI, pour les années civiles 2008 à 2016, ce qui limite, du moins pour le moment, la durée de l'analyse longitudinale.

Introduction

Dans un récent billet, [Liens entre l'IMT et les microdonnées](#), nous avons exploré le potentiel qu'offrirait le couplage de différentes

données administratives. Dans cette optique, le [budget fédéral 2018](#) comprenait d'importants investissements pour la mise en œuvre de la Plateforme de liens longitudinaux entre

CiMT

l'éducation et le marché du travail (PLLEMT). La PLLEMT est une plateforme de couplage qui permet de combiner des fichiers de données distincts en une seule clé maîtresse anonymisée. En novembre 2018, la plateforme est devenue accessible dans les **Centres de données de recherche (CDR)** de Statistique Canada. Les chercheurs peuvent maintenant **accéder aux données de cette plateforme** en soumettant un **projet de recherche**, en passant une vérification des antécédents par la GRC et en devenant une « personne réputée être employée » de Statistique Canada. Une fois approuvés, les chercheurs peuvent accéder à la plateforme uniquement dans les environnements hautement sécurisés des CDR, situés dans certaines universités à travers le pays et au siège social de Statistique Canada à Ottawa.

La plateforme rassemble présentement trois principaux ensembles de données administratives que les chercheurs et analystes de données peuvent relier entre eux et selon le temps. Ces ensembles sont les suivants :

- **SIEP** : Le Système d'information sur les étudiants postsecondaires, qui comprend les dossiers des étudiants collégiaux et universitaires sur leurs programmes, diplômes et domaines d'études.
- **SIAI** : Le Système d'information sur les apprentis inscrits, qui comprend les dossiers des apprentis inscrits sur leurs programmes d'apprentissage, diplômes et métiers de formation.
- **FFT1** : Le Fichier des familles T1 des formulaires fiscaux, c'est-à-dire les formulaires fiscaux qui comprennent des renseignements sur les revenus d'emplois et d'autres sources, l'industrie d'emploi (**SCIAN à trois chiffres**, le Système de classification des industries de l'Amérique du Nord), ainsi que des indicateurs comme l'assurance-emploi et l'aide sociale.

Chaque ensemble de données contient aussi des variables sociodémographiques importantes, comme l'âge et le genre et, dans le cas du SIEP et du FFT1, des renseignements sur le statut des étudiants (p. ex. : citoyenneté canadienne ou visa étudiant; voir le **tableau 1**).

L'environnement de la PLLEMT permettra au CIMT et aux autres chercheurs d'explorer les transitions des étudiants et des apprentis vers le marché du travail. Dans cette publication, nous discutons plus amplement des possibilités et limites que présentent de telles analyses, mais nous présentons d'abord plus de précisions sur l'information disponible dans chacun des trois ensembles de données de la plateforme.

Coup d'œil sur les données

Les trois sources de données qui composent la PLLEMT réunissent un ensemble de renseignements complémentaires sur l'éducation, la formation et les revenus. Une liste des principales variables disponibles dans chaque ensemble est présentée dans le **tableau 1**. Veuillez noter que la liste n'est aucunement exhaustive, puisque la plateforme contient une grande variété d'informations.

Les données du Système d'information sur les étudiants postsecondaires (SIEP) comprennent des dossiers de plus de 300 établissements collégiaux et universitaires publics de chaque province et territoire, pour les années scolaires 2009-2010 à 2015-2016. Les données du SIEP contiennent des renseignements sur le statut de l'étudiant dans le programme (inscrit, diplômé), le statut d'étudiant international et le domaine d'étude pour chaque année scolaire.

Le Système d'information sur les apprentis inscrits (SIAI) contient des renseignements sur les travailleurs inscrits dans des programmes d'apprentis, y compris les formations en classe et en milieu de travail, ainsi sur les ouvriers

professionnels. Le SIAI comprend des informations sur les programmes d'apprentissage **Sceau rouge** ou non, et indique si la formation est obligatoire ou volontaire. Les programmes Sceau rouge sont des programmes d'apprentissage intégrés au travail durant généralement de deux à cinq ans, standardisés dans l'ensemble des provinces et territoires. Les données du SIAI comprennent des renseignements sur le statut d'apprenti (formation en cours, terminée, etc.), le métier et le groupe de base de la catégorie professionnelle (**la Classification nationale des professions 2016 ou CNP**) associée au métier. Elles sont disponibles pour les années civiles 2008 à 2016.

Le Fichier des familles T1 (FFT1) contient des informations sur les revenus individuels par catégorie, y compris les revenus d'emploi, d'aide sociale et d'assurance-emploi.

Au lancement initial de la PLLEMT, les données du FFT1 incluront les années 2004 à 2015 pour le SIEP (les données du FFT1 2016 seront ajoutées en mars 2019), et jusqu'à 2016 pour le SIAI. Lorsqu'un individu apparaît pour la première fois dans le SIEP ou le SIAI, ses données FFT1 deviennent disponibles à partir de cette année d'imposition jusqu'à 2015 (SIEP) ou 2016 (SIAI). Si les données du FFT1 d'un individu sont disponibles pour des années précédant la première apparition dans le SIEP ou le SIAI, la plateforme permet le couplage avec ces observations antérieures aussi.

Tableau 1 : Résumé des types de renseignements et variables de la PLLEMT

	SIEP	SIAI	FFT1
Période et statut	<ul style="list-style-type: none"> Ensemble des provinces et territoires : années scolaires 2009-2010 à 2015-2016 Maritimes : 2005-2006 à 2015-2016 Date d'entrée dans le programme Inscription à temps plein ou à temps partiel à l'automne Abandon ou diplôme à la fin de l'année de référence 	<ul style="list-style-type: none"> Années civiles 2008 à 2016 Dates d'inscription et de certification Inscription, annulation, transfert ou réussite 	<ul style="list-style-type: none"> Les années civiles couvertes varient par groupe : <ul style="list-style-type: none"> Pleine couverture SIEP : 2004 à 2015 Pleine couverture SIAI : 2004 à 2016
Variables principales	<ul style="list-style-type: none"> Domaine d'études Type de programme (ex. : premier, deuxième ou troisième cycle) Type de diplôme (ex. : certificat, baccalauréat ou maîtrise) 	<ul style="list-style-type: none"> Renseignements sur le métier (différents niveaux de détail, y compris la norme CNP 2016) Durée du programme (ex. : durée normale de complétion) Indicateurs sur les métiers Sceau rouge et les programmes obligatoires et volontaires 	<ul style="list-style-type: none"> Revenus d'emploi, y compris ceux du travail autonome Secteur d'emploi (SCIAN à 3 chiffres) Revenus d'entreprise Revenus non marchands provenant de l'assurance-emploi, du RPC, du RRQ ou de l'aide sociale (SIEP-FFT1 seulement) Crédits d'impôts (ex. : enfants à charge)

	SIEP	SIAI	FFT1
Données sociodémographiques	<ul style="list-style-type: none"> Statut d'étudiant au Canada (ex. : citoyen canadien, résident permanent, visa étudiant) Province ou territoire de résidence à l'admission Province ou territoire de l'établissement 	<ul style="list-style-type: none"> Province ou territoire du programme d'apprentissage Province ou territoire de résidence 	<ul style="list-style-type: none"> Date d'entrée ou de sortie des immigrants et émigrants pour cette année d'imposition (SIEP-FFT1 seulement) Revenu parental (SIEP-FFT1 seulement) Province ou territoire de résidence
	Âge et genre		

Opportunités pour évaluer les résultats

Afin de tirer profit de la structure longitudinale de la plateforme, les chercheurs peuvent adopter deux principales approches. La première est basée sur le nombre d'années depuis la fin du programme de formation ou d'enseignement (par la diplomation ou pour toute autre raison). L'autre approche est l'analyse par année, axée sur les résultats dans le marché du travail lors d'une année d'imposition donnée, pour différentes cohortes ayant pris part à des programmes de formation ou d'enseignement postsecondaire. Les deux approches permettent de regrouper les individus selon des caractéristiques démographiques, des types de programmes, ou encore des domaines ou métiers d'études.

Avec les deux approches, nous serons en mesure de suivre et d'évaluer l'avantage financier que procure la participation à des programmes d'enseignement et de formation, par exemple les revenus relatifs à différents domaines d'études ou types de diplômes. De la même manière, il sera possible de comparer les niveaux de revenus des diplômés à ceux des personnes n'ayant pas terminé leur programme. Enfin, plusieurs chercheurs seront intéressés par la possibilité de quantifier les trajectoires de revenus de plusieurs groupes sous-représentés (p. ex : les femmes, les immigrants, etc.) tout en contrôlant des

facteurs comme le domaine d'études et la réussite du programme.

Au-delà du niveau de revenu, les analyses des résultats sur le marché du travail pourraient se concentrer sur la volatilité des revenus. Dans ce cas, une analyse par année est probablement mieux adaptée étant donné que la volatilité des revenus est fortement influencée par les conditions macroéconomiques. La volatilité des revenus pourrait être représentée de plusieurs façons, y compris par des changements annuels dans le revenu, le recours à l'assurance-emploi ou à l'aide sociale, ou le nombre de T4 déclarés chaque année.

De toute évidence, il y a de nombreuses autres possibilités d'analyse à considérer. Bien qu'une liste exhaustive dépasse l'objectif de cette édition, voici d'autres pistes intéressantes :

- La distribution des anciens étudiants et apprentis inscrits vivant des périodes de faibles revenus.
- La mobilité sociale des individus par rapport à leur niveau de revenu familial (possible pour les données FFT1 liées au SIEP).

Nous espérons que les chercheurs de partout au pays fouilleront la plateforme en profondeur et qu'ils partageront largement leurs découvertes. Nos actions collectives, qu'elles soient indépendantes ou le fruit de collaborations,

donneront lieu à de nouvelles perspectives sur le marché du travail qui permettront d'informer le développement global de politiques fondées sur les données. Bien sûr, les résultats préliminaires ont déjà mis en évidence le potentiel de la plateforme à générer de l'information (voir l'**encadré 1**).

Limites et mises en garde

Bien que la PLLEMT offre de nombreuses occasions d'évaluer les résultats dans le marché du travail d'étudiants et d'apprentis inscrits, elle comporte de nombreuses limites que l'on doit garder en tête.

Information sur les professions

L'information sur l'emploi occupé que contient la PLLEMT est limitée. Même si les données sur les revenus sont solides, le FFT1 ne contient aucun renseignement sur la profession exercée. Les données du FFT1 comportent des informations sur l'industrie, sous forme de codes SCIAN à trois chiffres. Mais bien que ces codes offrent certaines indications sur le lien entre la formation et l'emploi, ils ne sont pas suffisants pour déterminer si une personne travaille dans un domaine lié à son programme de formation ou d'enseignement. Par exemple, à partir du SIEP, on peut observer des titulaires de diplômes en informatique (**CPE** : 11.0701) qui, selon le FFT1, travaillent maintenant dans le secteur bancaire (SCIAN3 : 522). Toutefois, sans code de la CNP, il est impossible de savoir si ces personnes travaillent dans le service de TI des banques, comme analystes dans un autre service ou comme cadre supérieur. Cette lacune pourra être comblée par l'ajout de nouvelles données à la plateforme, comme celles du recensement ou de l'Enquête nationale auprès des diplômés (ENG).

Qualité et quantité de travail

En plus du manque d'information sur les professions, la plateforme ne contient aucun renseignement sur la qualité ou la quantité de travail. Il n'y a par exemple aucune donnée sur les heures travaillées, la flexibilité d'horaire ou

Encadré 1 : Premiers résultats de la PLLEMT

Le 4 décembre 2018, Statistique Canada a publié les premiers résultats publics de couplages de données dans la PLLEMT. On rapporte que de 2010 à 2014, plus de 900 000 étudiants de moins de 35 ans ont terminé un programme d'études d'un établissement postsecondaire et sont entrés sur le marché du travail. Le revenu d'emploi médian deux ans après la fin des études était de 43 600 \$ pour les titulaires d'un grade de premier cycle (53 % de l'échantillon total) et de 39 100 \$ pour les titulaires d'un diplôme collégial (14 % de l'échantillon total). Une autre conclusion importante est l'écart de revenu entre les sexes pour les titulaires de diplômes collégial et universitaire. Cinq ans après l'obtention du diplôme, les femmes détenant un grade de premier cycle gagnaient 21,1 % moins (ou 13 300 \$/année) que les hommes ayant le même diplôme, et les femmes détenant un diplôme collégial, 29,8 % moins (ou 16 200 \$/année) que les hommes ayant le même niveau de scolarité.

même sur la volatilité des revenus au cours d'une année d'imposition. D'autres facteurs importants, comme la conciliation travail-vie personnelle ou le nombre d'emplois occupés simultanément, sont aussi absents des ensembles de données de la PLLEMT. En ce sens, l'ajout de l'Enquête sur la population active (EPA) pourrait aider et, au moins pour un sous-ensemble de la population, améliorer

la disponibilité des données sur la qualité et la quantité de travail. L'ajout des données du recensement ou de l'ENG à la plateforme, au moins pour une période de temps, pourrait aussi servir à déterminer, pour un plus grand échantillon, si un diplômé travaille à temps plein ou à temps partiel (et pourquoi), si l'emploi est permanent ou non, ou d'autres éléments relatifs à l'emploi.

Compétences des travailleurs

Généralement, les données administratives offrent peu d'information sur les compétences, qui doivent être autodéclarées ou mesurées par des tests. Une façon de contourner ce problème serait de se servir du niveau de compétence associé à chaque code CNP à quatre chiffres, s'ils devenaient disponibles grâce au recensement ou à l'EPA lors des prochaines mises à jour de la PLLEMT. Toutefois, comme mentionné dans le **numéro 3 de Perspectives de l'IMT**, les mesures directes d'un ensemble restreint de compétences sont de plus en plus possibles grâce à des tests en ligne, mais cela demeure une lacune dans le système d'information sur le marché du travail canadien.

Groupes de comparaison des revenus

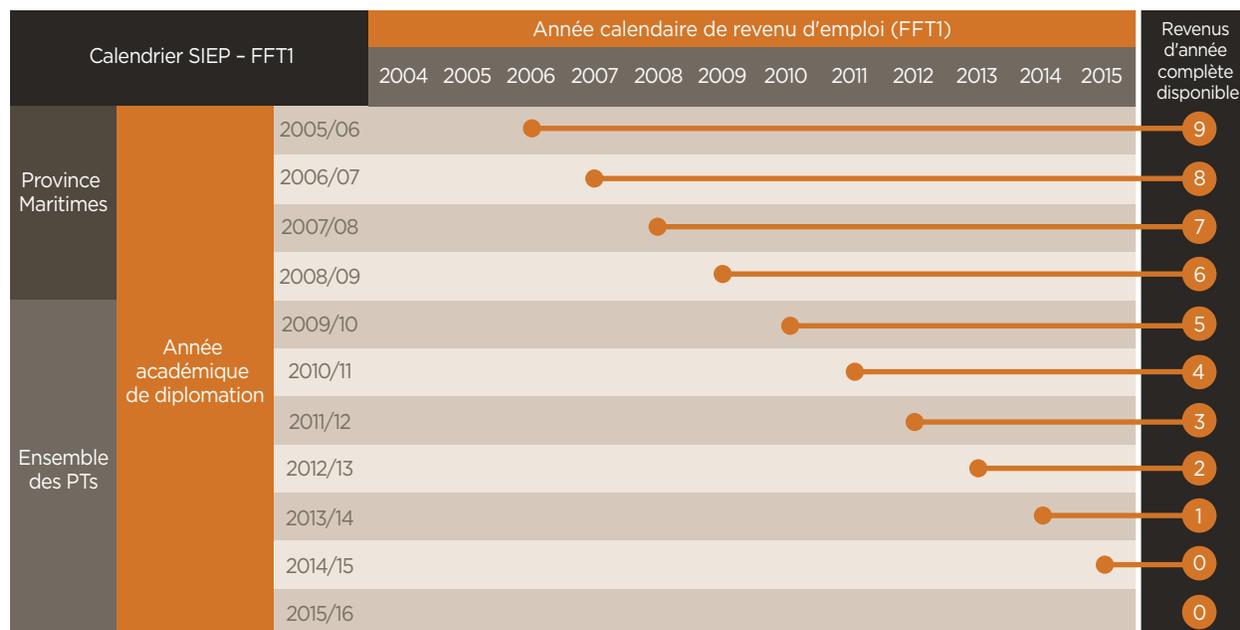
Bien que les niveaux et changements de revenus soient l'élément central de l'analyse possible sur la plateforme, une comparaison des groupes de revenus devrait être mise au point. Idéalement, les revenus des participants à des programmes de formation et d'enseignement devraient être comparés à ceux des individus n'ayant pas pris part à de tels programmes. Malheureusement, les données FFT1 de la plateforme sont seulement disponibles pour ceux qui sont dans le SIEP ou le SIAI. Au minimum, on pourrait comparer les revenus des diplômés avec ceux des personnes qui ont commencé des études mais ne les ont pas terminées.

Même si les chercheurs avaient un accès direct aux renseignements du FFT1 pour de tels groupes de comparaison, la façon d'aligner les cohortes de la PLLEMT avec celles qui n'y sont pas n'est pas claire. L'âge, le sexe et le domaine seraient les caractéristiques évidentes par lesquelles aligner le traitement des données et les groupes témoins (p. ex. : femme de 25 ans travaillant dans le secteur minier, dans l'univers SIEP/SIAI et hors de celui-ci). Toutefois, cette comparaison ne serait pas parfaite, puisque les revenus sont normalement mesurés selon le nombre d'années depuis l'entrée sur le marché du travail, ce qui est très différent pour les anciens étudiants, les apprentis, et ceux qui ont commencé à travailler directement à leur sortie du secondaire.

Couverture temporelle

Un problème lié à celui des groupes de comparaison de revenus est le fait que les analyses basées sur le nombre d'années depuis le diplôme regroupent des individus provenant de contextes macroéconomiques très différents. Considérons les données du SIEP présentées dans le **tableau 2**. Pour toutes les régions, la PLLEMT a de l'information sur une cohorte ayant cinq années complètes de revenus postdiplôme (diplômés 2009-2010); toutes les autres cohortes ont quatre années ou moins de renseignements sur les revenus. L'analyse des revenus pour les années $t+1$ depuis l'obtention du diplôme inclurait les diplômés de 2009-2010 à 2015-2016. Cela signifie que les revenus de l'année $t+1$ comptent ceux des individus qui ont obtenu leur diplôme dans un contexte économique postcrise financière ainsi que de ceux qui ont terminé leurs études lors du resserrement du marché du travail de 2015. Le contrôle de ces conditions macroéconomiques est une procédure importante qui n'est pas tout à fait simple.

Tableau 2. Calendrier SIEP-FFT1 des années d'obtention du diplôme (du SIEP) et des années de revenus d'emploi (du FFT1) démontrant les années d'analyses postdiplôme possibles pour chaque cohorte



Coûts de l'éducation

La dernière mise en garde quant aux analyses de revenus, et probablement la plus importante, concerne les coûts directs (frais de scolarité et autres) et indirects (manques à gagner) de l'éducation et des programmes de formation, dont les chercheurs devraient tenir compte. Les coûts de l'éducation ne font pas partie de la plateforme, et les coûts d'opportunité des revenus sont, bien sûr, inobservables. Les coûts d'hébergement sont aussi exclus de ces sources de données pour les étudiants qui n'habitent pas chez leurs parents pendant leurs études. Les chercheurs devront présenter des hypothèses réfléchies pour ajuster adéquatement et précisément leurs analyses selon ces coûts.

Couplages à venir

Certaines des mises en garde et limites soulignées précédemment pourraient être palliées par l'inclusion d'ensembles de données supplémentaires. Certains ensembles de données sont déjà envisagés pour intégration à la plateforme¹ :

- Données du Programme canadien de prêt aux étudiants (disponibilité dans les CDR en février 2019)
- Données du Programme canadien pour l'épargne-études (disponibilité prévue dans les CDR à l'automne 2020)
- Données des programmes de Prêt canadien aux apprentis et de subventions aux apprentis
- Données des prestataires de l'assurance-emploi, des relevés d'emploi et du Profil vectoriel de l'assurance-emploi
- Certains renseignements du recensement et de l'Enquête nationale auprès des ménages

¹ Selon les documents de gouvernance la Plateforme de liens longitudinaux entre l'éducation et le marché du travail, qui ne sont pas encore publics.

Bien que ces ensembles de données comportent aussi leurs propres limites et mises en garde, leur intégration à la plateforme permettrait de combler deux lacunes : 1) les coûts des programmes de formation et d'enseignement et 2) les informations propres à l'emploi. La première serait comblée par les données administratives sur les prêts étudiants, un élément important à considérer dans toute analyse des revenus nets et des rendements de l'éducation. La deuxième lacune serait, elle, comblée par le recensement de 2016 et l'Enquête nationale auprès des ménages (ENM) 2011. Ces deux sondages donnent toutefois lieu à de nouvelles limites. D'abord, le recensement et l'ENM couvrent respectivement 25 % et 33 % des ménages. Ensuite, l'information sur l'emploi contenue dans ces ensembles de données est disponible seulement chaque cinq ans.

Une autre source de données qui pourrait être intégrée à la PLLEMT est l'Enquête sur la population active (EPA), comme mentionné précédemment. Bien que sa couverture de la population soit beaucoup plus restreinte que celle du recensement ou de l'ENM, l'EPA offrirait des données plus à jour sur les catégories d'emploi. Notons aussi que l'intégration de l'EPA permettrait une mesure partielle des heures travaillées. De la même manière, les données de l'assurance-emploi ajouteraient des détails sur l'emploi pour ceux qui ont reçu des prestations d'assurance-emploi, tout en offrant de l'information sur la durée de la période de chômage.

La voie à suivre

La Plateforme de liens longitudinaux entre l'éducation et le marché du travail est une initiative riche et importante. Elle offrira des preuves empiriques et soutiendra la recherche et les découvertes, selon une diversité de perspectives, pour de nombreux intervenants et individus intéressés par les transitions des étudiants et apprentis inscrits vers le marché du travail. Dans cette optique, le **CIMT** prévoit de lancer une série d'analyses en profondeur des résultats sur le marché du travail, en partenariat avec des experts reconnus de l'**Education Policy Research Institute (EPRI)** et de Statistique Canada.

Il sera toutefois essentiel de contextualiser toute idée initiale tirée de la plateforme, au moyen des groupes de comparaison adéquats et de mises en garde claires sur l'analyse. Une telle contextualisation est incontournable afin d'assurer la pertinence des résultats pour les intervenants, les décideurs politiques, les étudiants et le grand public canadien. Par conséquent, nous attendons impatiemment l'intégration de nouveaux ensembles de données à la plateforme afin de combler les principales lacunes en matière d'information et d'élargir les possibilités de recherche.

Aux côtés de l'ensemble de la communauté de la recherche, le **CIMT** est impatient de plonger dans la plateforme, de partager ses découvertes, en soulignant adéquatement leurs limites, et d'adapter ces renseignements à la diversité des besoins des utilisateurs.

Remerciements

Le présent numéro de *Perspectives de l'IMT* a été préparé par Behnoush Amery, du CIMT. Nous aimerions remercier notre **Groupe consultatif pancanadien des intervenants** ainsi que notre **Comité d'experts en information sur le marché du travail** et **Statistique Canada** pour leurs commentaires et suggestions. Plus particulièrement, l'équipe souhaite souligner la précieuse rétroaction et contribution de Sylvie Gauthier, Christine Hinchley, Tamara Knighton, André Lebel (Statistique Canada), Ross Finnie, Michael Dubois, Masashi Miyairi (Education Policy Research Initiative) et Arthur Sweetman (Université McMaster).

Pour en savoir plus sur cette édition de *Perspectives de l'IMT* et les autres activités du CIMT, veuillez jeter un œil à notre **page de publications** ou encore contacter Behnoush Amery, économiste principale, à behnoush.amery@lmic-cimt.ca, ou Tony Bonen, directeur, recherche et analytique, à tony.bonen@lmic-cimt.ca.