

2000s-55

**Le décrochage scolaire,
la performance scolaire
et le travail pendant les études :
un modèle avec groupe hétérogène**

*Marcel Dagenais, Claude Montmarquette,
Nathalie Viennot-Briot, Muriel Meunier*

Série Scientifique
Scientific Series



CIRANO
Centre interuniversitaire de recherche
en analyse des organisations

Montréal
Novembre 2000

CIRANO

Le CIRANO est un organisme sans but lucratif constitué en vertu de la Loi des compagnies du Québec. Le financement de son infrastructure et de ses activités de recherche provient des cotisations de ses organisations-membres, d'une subvention d'infrastructure du ministère de la Recherche, de la Science et de la Technologie, de même que des subventions et mandats obtenus par ses équipes de recherche.

CIRANO is a private non-profit organization incorporated under the Québec Companies Act. Its infrastructure and research activities are funded through fees paid by member organizations, an infrastructure grant from the Ministère de la Recherche, de la Science et de la Technologie, and grants and research mandates obtained by its research teams.

Les organisations-partenaires / The Partner Organizations

- École des Hautes Études Commerciales
- École Polytechnique
- Université Concordia
- Université de Montréal
- Université du Québec à Montréal
- Université Laval
- Université McGill
- MEQ
- MRST
- Alcan Aluminium Ltée
- AXA Canada
- Banque Nationale du Canada
- Banque Royale du Canada
- Bell Québec
- Bombardier
- Bourse de Montréal
- Développement des ressources humaines Canada (DRHC)
- Fédération des caisses populaires Desjardins de Montréal et de l'Ouest-du-Québec
- Hydro-Québec
- Imasco
- Industrie Canada
- Pratt & Whitney Canada Inc.
- Raymond Chabot Grant Thornton
- Ville de Montréal

© 2000 Marcel Dagenais, Claude Montmarquette, Nathalie Viennot-Briot et Muriel Meunier. Tous droits réservés. All rights reserved.

Reproduction partielle permise avec citation du document source, incluant la notice ©.

Short sections may be quoted without explicit permission, provided that full credit, including © notice, is given to the source.

Ce document est publié dans l'intention de rendre accessibles les résultats préliminaires de la recherche effectuée au CIRANO, afin de susciter des échanges et des suggestions. Les idées et les opinions émises sont sous l'unique responsabilité des auteurs, et ne représentent pas nécessairement les positions du CIRANO ou de ses partenaires.

This paper presents preliminary research carried out at CIRANO and aims at encouraging discussion and comment. The observations and viewpoints expressed are the sole responsibility of the authors. They do not necessarily represent positions of CIRANO or its partners.

Le décrochage scolaire, la performance scolaire et le travail pendant les études : un modèle avec groupe hétérogène*

Marcel Dagenais[†], Claude Montmarquette[‡], Nathalie Viennot-Briot[§],
Muriel Meunier[¶]

Résumé / Abstract

Le chômage chez les jeunes non qualifiés au Canada est particulièrement élevé et la mondialisation actuelle de l'économie ne fait qu'amplifier le problème. La première source de non-qualification des jeunes travailleurs est l'abandon des études secondaires. Dans cet article, nous prenons en compte l'effet du travail pendant les études sur la probabilité de quitter l'école au secondaire. Contrairement aux autres articles que nous avons réalisés sur le sujet, nous considérons ici la relation – décision de quitter l'école et travail pendant les études – d'une manière plus complexe à l'aide de *l'enquête auprès des sortants de 1991* de Statistique Canada. En effet, nous introduisons une réflexion sur le processus décisionnel qui conduit les étudiants à quitter l'école à l'aide de deux modèles. Dans un premier modèle, l'école est la principale préoccupation de l'individu tandis que dans le second, l'expérience de travail est avant tout privilégiée. Ne sachant pas a priori dans quel modèle un individu spécifique se trouvera, nous utilisons un modèle économétrique afin de déterminer la probabilité qu'un individu se retrouve dans le premier ou le second modèle. Les recherches identifient plusieurs déterminants du décrochage scolaire et différentes politiques qui pourraient être mise en œuvre par le gouvernement afin de réduire le décrochage scolaire.

Unemployment of the unqualified young workers in Canada has always been a preoccupation and is doomed to increase with the mondialisation of the economy. The first source of non qualification is to dropout of school. In this paper, we consider the effect of working while in school on the probability of dropping out at the secondary level. However, contrary to previous papers we have realised on this subject, we consider here the relation – decision to dropout and working while in school – in a more complex setting. In this paper, we entertain the possibility of two models to explain the decisions of the individuals

Adresse de l'auteur : Claude Montmarquette, CIRANO, 2020 rue University, 25^e étage, Montréal, Qc, Canada H3A 2A5
Tél. : (514) 985-4015 Fax : (514) 985-4039 courriel : montmarc@cirano.umontreal.ca
Cette recherche a été financée par la Direction générale de la recherche appliquée, Développement des Ressources Humaines Canada.

[†] Université de Montréal et CIRANO

[‡] Université de Montréal et CIRANO

[§] CIRANO

[¶] CIRANO

in our sample of Statistics Canada's School Leavers Survey. In a first model, school is the main preoccupation, while in the second model experience of work while in school is the leading decision. Not knowing a priori, which model a specific individual belongs to, we use a sophisticated econometric model to compute the probability that an individual with given characteristics is part of the first or the second model. The research identifies several determinants of dropping out and different policies which can be put in action by the government in order to reduce dropout.

Mots Clés : Abandon scolaire, travail pendant les études, performance scolaire, salaire minimum, conditions macroéconomiques sur le marché du travail

Keywords: School dropout, hours of work during study, grades, minimum wages, labour market conditions

Table des matières

1. L'analyse empirique de l'abandon et du travail pendant les études.....	1
1.1. Spécifications économétriques des comportements.....	1
1.1.1 Présentation du modèle Travail, Notes, Abandon (TNA).....	2
1.1.2 Présentation du modèle Notes, Travail, Abandon (NTA).....	5
1.1.3 Le modèle avec groupe hétérogène.....	7
1.2. Les résultats économétriques.....	8
2. Conclusion.....	22
Annexe A. Les données – définition et construction	24
Annexe B. Les spécifications économétriques des modèles TNA et NTA	28
Bibliographie.....	36

1. L'analyse empirique de l'abandon et du travail pendant les études

1.1 Spécifications économétriques des comportements

Dans la mesure où toute décision comporte un choix, on peut considérer que l'étudiant qui décide d'abandonner l'école a antérieurement fait une réflexion sur les autres alternatives qui s'offre à lui. On sait que l'individu ne poursuit ses études que si celles-ci lui procurent des avantages, dans la mesure où il choisit le niveau de scolarité qui selon lui, maximisera son bien-être. Bien que ces avantages puissent être liés au plaisir de l'étude, la théorie du capital humain¹ suggère surtout de considérer l'éducation comme un investissement. Cet investissement permet ainsi d'anticiper des flux de revenus et une consommation de biens et services plus élevés dans le futur. Mais l'investissement a un coût d'opportunité puisqu'il induit de devoir renoncer à consommer dans l'immédiat ou en d'autres termes de devoir renoncer à un travail permanent. En résumé, cela revient à considérer que l'individu fait un choix entre seulement deux alternatives, à savoir l'école ou le travail. Le seul compromis possible est celui du mariage études/travail à temps partiel. Et on peut alors envisager que les étudiants qui décident de travailler pendant leurs études cherchent en fait à réduire le coût d'opportunité lié aux études. Pourtant certains individus ne semblent pas être en mesure de gérer à fois leurs études et un travail.

En effet, la relation travail-abandon est plus complexe qu'elle n'y paraît et il est important de tenir compte de toutes les séquences dynamiques possibles de décisions. S'attaquer au problème de l'abandon scolaire induit implicitement de porter une attention particulière au processus décisionnel qui conduit les étudiants à quitter l'école. Il peut paraître simple de mettre en évidence la dynamique suivie par les individus concernant le travail pendant les études et la décision d'abandonner ou non l'école. Pourtant, un modèle théorique qui s'intéresse seulement aux variables affectant un ensemble d'individus, comme les conditions économiques de leur comté de résidence ou le salaire minimum, n'aborde le travail pendant les études que de façon incidente. Aussi, paraît-il pertinent de faire intervenir la performance scolaire dans le processus décisionnel comme une composante à part entière. En effet, un étudiant qui réussie parfaitement à allier travail et de bonnes notes aura peu de raisons d'abandonner l'école. Par contre, un étudiant qui à de mauvaises notes sera davantage tenté d'abandonner l'école. Reste à déterminer si ce sont le travail ou les mauvaises notes qui sont la cause de l'abandon scolaire.

On peut considérer que la décision de travailler précède toujours celle d'abandonner l'école². Dans ce cas, ce sont deux décisions séquentielles et il est important de noter qu'un tel raisonnement induit implicitement un effet négatif du travail extra-scolaire sur l'abandon. On peut alors raisonnablement supposer que cet effet se concrétise au niveau de la performance scolaire. En résumé, l'étudiant travaille, il délaisse son travail scolaire et a de mauvaises notes et décide alors d'abandonner l'école. Une lecture des données de l'échantillon laisse d'ailleurs penser que

¹ Becker, G. S. (1964) : *Human capital : A theoretical and empirical analysis with special reference to education*, Columbia University Press.

² Dagenais, M., C. Montmarquette, D. Parent et N. Viennot-Briot (1999) : *Travail pendant les études, performance scolaire et abandon* , Série Scientifique n° 99S-41, CIRANO.

ceux qui travaillent au-delà d'un certain nombre d'heures par semaine ont des résultats scolaires moins bons que ceux qui ne travaillent pas. Pourtant, le précédent rapport mettait en évidence que travailler moins de 10 heures par semaine est loin d'inciter au décrochage, et pour ceux d'entre eux qui obtiendront leur diplôme, l'expérience acquise parallèlement aux études pourra même favoriser leur insertion sur le marché de l'emploi. Un tel lien de causalité entre travail et abandon ne semble donc pas couler de source.

En effet, la réalité peut se révéler beaucoup plus complexe, voire totalement différente. L'individu pourrait également décider de trouver un emploi parce qu'il songe à abandonner l'école. C'est toujours une décision de type séquentiel mais dans la direction opposée au premier scénario. La modélisation d'un modèle qui prend en compte le deuxième processus de comportement s'avère fort utile, afin de valider ou non les intuitions concernant la séquence des décisions. Dans la mesure où l'on considère que le choix d'abandonner précède la participation au marché du travail, on suppose dans ce cas que ce sont les performances scolaires qui ont un impact négatif sur l'abandon. On établit alors que cet impact se concrétise dans la participation à un emploi sur le marché du travail. On notera toutefois qu'au niveau de l'échantillon, le manque d'informations sur la séquence des décisions ne permet pas de déterminer les probabilités séquentielles. Les résultats de l'estimation du modèle qui prend en considération ce dernier cas de figure ne sont pas réellement satisfaisantes. Néanmoins, cela permet de soulever une nouvelle réflexion, à savoir que le comportement des individus ne serait pas prédéterminé et permet d'introduire le concept d'hétérogénéité à l'intérieur du groupe.

1.1.1 Présentation du modèle Travail, Notes, Abandon (TNA)

Le modèle TNA considère le processus de décision suivant : l'individu prend la décision de travailler un certain nombre d'heures, le travail a des répercussions négatives, neutres ou positives sur ses résultats scolaires et l'étudiant décide d'abandonner ou non ses études secondaires. Le modèle approprié comprend alors : un probit ordonné sur les heures travaillées, un probit ordonné sur les notes (qui dépend des heures travaillées et d'autres variables), et un probit sur l'abandon (qui dépend des heures travaillées, des notes et d'autres variables).

Considérons l'équation (1) afin d'expliquer les déterminants du travail pendant les études :

$$T_i^* = x_i \beta + u_i \quad (1)$$

T_i^* mesure l'utilité du travail. En pratique, c'est une variable non observable ou variable latente. Seule la décision de travailler ou non est observée :

$$T_i = 0 \text{ si } T_i^* \leq 0;$$

$$T_i = 1 \text{ si } 0 < T_i^* \leq \varepsilon_1;$$

$$T_i = 2 \text{ si } \varepsilon_1 < T_i^* \leq \varepsilon_2;$$

$T_i = 3$ si $T_i^* > \varepsilon_2$.

Avec :

$T_i = 0$ si le nombre d'heures travaillées est égal à zéro;

$T_i = 1$ si le nombre d'heures travaillées est supérieur à zéro et inférieur ou égal à quinze;

$T_i = 2$ si le nombre d'heures travaillées est supérieur à quinze et inférieur ou égal à trente;

$T_i = 3$ si le nombre d'heures travaillées est supérieur à trente.

x_i un vecteur de variables exogènes, β est le vecteur des paramètres et u_i l'erreur résiduelle.

Considérons ensuite l'équation (2), afin d'expliquer les déterminants de la performance scolaire :

$$N_i^* = z_i \gamma + \sum_{j=1}^3 \lambda_j D_j + v_i \quad (2)$$

N_i^* mesure l'utilité de la performance scolaire de l'individu. C'est un variable latente car seule la performance scolaire³ est observée :

$$N_i = 0 \text{ si } N_i^* \leq 0 \Rightarrow v_i \leq -F \text{ avec } -F = -z_i \gamma - \sum_{j=1}^3 \lambda_j D_j ;$$

$$N_i = 1 \text{ si } 0 < N_i^* \leq \theta_1 \Rightarrow v_i \leq \theta_1 - F;$$

$$N_i = 2 \text{ si } N_i^* > \theta_1 \Rightarrow v_i > \theta_1 - F.$$

Avec :

$N_i = 0$ si la note obtenue par l'étudiant est inférieure à 60 %;

$N_i = 1$ si la note obtenue par l'étudiant est supérieure ou égale à 60 % et inférieure à 70 %;

$N_i = 2$ si la note obtenue par l'étudiant est supérieure ou égale à 70 %.

Dans ce modèle, on définit la performance scolaire en fonction du nombre d'heures travaillées, d'où :

$$D_j = 1 \text{ si } T = j$$

$$D_j = 0 \text{ sinon.}$$

C'est à dire :

$D_0 = 1$ si le nombre d'heures travaillées est égal à zéro;

$D_1 = 1$ si le nombre d'heures travaillées est supérieur à zéro et inférieur à quinze;

³ N équivaut plus spécifiquement, à la moyenne de l'individu au cours du dernier trimestre complet d'étude.

$D_2 = 1$ si le nombre d'heures travaillées est supérieur à quinze et inférieur ou égal à trente;

$D_3 = 1$ si le nombre d'heures travaillées est supérieur à trente;

z_i est un vecteur de variables exogènes. La note obtenue est également une fonction non linéaire du nombre d'heures travaillées. Nous utilisons une fonction par paliers. γ et λ représentent les paramètres de l'équation. v_i est l'erreur résiduelle.

L'équation de l'abandon (3) correspond à un probit simple. L'équation des notes ainsi que celle des heures travaillées correspondent à des probit ordonnés. On suppose que les erreurs résiduelles η_i , v_i et u_i suivent une loi normale trivariée. La forme précise de l'équation d'abandon est la suivante :

$$A_i^* = w_i \delta + \sum_{j=1}^3 \Pi_j D_j + \sum_{k=1}^2 \Phi_k M_k + \eta_i \quad (3)$$

A_i^* est une variable latente qui mesure l'utilité de l'abandon scolaire ou la propension au décrochage. Seule la décision ou non d'abandonner est observée :

$$A_i = 0 \text{ si } A_i^* \leq 0 \Rightarrow \eta_i \leq -G \text{ avec } -G = -w_i \delta - \sum_{j=1}^3 \Pi_j D_j - \sum_{k=1}^2 \Phi_k M_k ;$$

$$A_i = 1 \text{ si } A_i^* > 0 \Rightarrow \eta_i > -G.$$

L'abandon est fonction du nombre d'heures travaillées et de la performance scolaire de l'individu :

$$D_j = 1 \text{ si } T = j;$$

$$D_j = 0 \text{ sinon.}$$

Et :

$$M_k = 1 \text{ si } N = k;$$

$$M_k = 0 \text{ sinon.}$$

C'est à dire :

$$M_0 = 1 \text{ si la note obtenue par l'étudiant est inférieure à 60\%;}$$

$$M_1 = 1 \text{ si la note obtenue par l'étudiant est supérieure ou égale à 60\% et inférieure à 70\%;}$$

$$M_2 = 1 \text{ si la note obtenue par l'étudiant est supérieure à 70\%.}$$

La solution est obtenue par maximum de vraisemblance. Comme il y a deux possibilités pour l'abandon, trois pour la note obtenue et quatre pour les heures travaillées, il y a en tout 24 possibilités. Le logarithme de la fonction de vraisemblance que l'on trouvera à l'annexe B contient donc 24 termes, chacun des termes étant une intégrale différente de la loi normale en

trois dimensions. La solution numérique est obtenue par une méthode itérative de gradient. Pour générer la solution initiale, on traite chaque équation séparément à l'aide du modèle probit ou probit ordonné, selon le cas.

1.1.2 Présentation du modèle Notes, Travail, Abandon (NTA)

Le modèle NTA considère le processus de décision suivant : l'individu évalue ses notes, en fonction de ses résultats scolaires il décide de travailler ou non et cette dernière décision influence son choix d'abandonner ou non l'école au secondaire. Ce modèle comprend un probit ordonné sur les notes, un probit ordonné sur les heures travaillées (qui dépend des notes et d'autres variables), et un probit sur l'abandon (qui dépend des heures travaillées, des notes et d'autres variables).

Considérons l'équation (4) afin d'expliquer les déterminants de la performance scolaire :

$$N_i^* = y_i \alpha + \mu_i \quad (4)$$

N_i^* mesure l'utilité de la performance scolaire. C'est une variable latente et seule la performance scolaire est observée :

$$N_i = 0 \text{ si } N_i^* \leq 0;$$

$$N_i = 1 \text{ si } 0 < N_i^* \leq \Omega_1;$$

$$N_i = 2 \text{ si } N_i^* > \Omega_1.$$

Avec :

$N_i = 0$ si la note obtenue par l'étudiant est inférieure à 60 %;

$N_i = 1$ si la note obtenue par l'étudiant est supérieure ou égale à 60 % et inférieure à 70 %;

$N_i = 2$ si la note obtenue par l'étudiant est supérieure ou égale à 70 %.

y_i un vecteur de variables exogènes, α est le vecteur des paramètres et μ_i est l'erreur résiduelle.

Considérons ensuite l'équation (5), afin d'expliquer les déterminants du travail pendant les études :

$$T_i^* = t_i \tau + \sum_{j=1}^2 \lambda_j D_j + \vartheta_i \quad (5)$$

T_i^* mesure l'utilité du travail. En pratique, c'est une variable non observable ou variable latente. Seule la décision de travailler ou non est observée :

$$T_i = 0 \text{ si } T_i^* \leq 0 \Rightarrow \vartheta_i \leq -E \text{ avec } -E = -t_i \tau - \sum_{j=1}^2 \lambda_j D_j ;$$

$$T_i = 1 \text{ si } 0 < T_i^* \leq \zeta_1 \Rightarrow \vartheta_i \leq \zeta_1 - E;$$

$$T_i = 2 \text{ si } \zeta_1 < T_i^* \leq \zeta_2 \Rightarrow \vartheta_i \leq \zeta_2 - E;$$

$$T_i = 3 \text{ si } T_i^* > \zeta_2 \Rightarrow \vartheta_i > \zeta_2 - E.$$

Avec :

$T_i = 0$ si le nombre d'heures travaillées est égal à zéro;

$T_i = 1$ si le nombre d'heures travaillées est supérieur à zéro et inférieur ou égal à quinze;

$T_i = 2$ si le nombre d'heures travaillées est supérieur à quinze et inférieur ou égal à trente;

$T_i = 3$ si le nombre d'heures travaillées est supérieur à trente.

Dans ce modèle, on définit le nombre d'heures travaillées en fonction de la performance scolaire, d'où :

$$D_j = 1 \text{ si } N = j;$$

$$D_j = 0 \text{ sinon.}$$

C'est à dire :

$$D_0 = 0 \text{ si la note obtenue par l'étudiant est inférieure à 60\%;}$$

$$D_1 = 1 \text{ si la note obtenue par l'étudiant est supérieure ou égale à 60\% et inférieure à 70\%;}$$

$$D_2 = 1 \text{ si la note obtenue par l'étudiant est supérieure ou égale à 70\%.}$$

t_i est un vecteur de variables exogènes. ϑ_i est l'erreur résiduelle.

L'équation de l'abandon (6) correspond à un probit simple. L'équation des heures travaillées ainsi que celle des notes correspondent à des probit ordonnés. On suppose que les erreurs résiduelles ψ_i , ϑ_i et μ_i suivent une loi normale trivariée. La forme précise de l'équation d'abandon est la suivante :

$$A_i^* = p_i \xi + \sum_{j=1}^2 \Pi_j D_j + \sum_{k=1}^3 \Phi_k M_k + \psi_i \quad (6)$$

A_i^* est une variable latente qui mesure l'utilité de l'abandon scolaire ou la propension au décrochage. Seule la décision d'abandonner ou non est observée :

$$A_i = 0 \text{ si } A_i^* \leq 0 \Rightarrow \psi_i \leq -H \text{ avec } -H = -p_i \xi - \sum_{j=1}^2 \Pi_j D_j - \sum_{k=1}^3 \Phi_k M_k ;$$

$$A_i = 1 \text{ si } A_i^* > 0 \Rightarrow \psi_i > -H.$$

L'abandon est fonction du nombre d'heures travaillées et de la performance scolaire de l'individu :

$$D_j = 1 \text{ si } N = j;$$

$$D_j = 0 \text{ sinon.}$$

Et :

$$M_k = 1 \text{ si } T = k;$$

$$M_k = 0 \text{ sinon.}$$

C'est à dire :

$$M_0 = 1 \text{ si le nombre d'heures travaillées est égal à zéro;}$$

$$M_1 = 1 \text{ si le nombre d'heures travaillées est supérieur à zéro et inférieur ou égal à quinze;}$$

$$M_2 = 1 \text{ si le nombre d'heures travaillées est supérieur à trente;}$$

$$M_3 = 0 \text{ sinon.}$$

La solution est obtenue par maximum de vraisemblance. Comme il y a deux possibilités pour l'abandon, trois pour la note obtenue et quatre pour les heures travaillées, il y a en tout 24 possibilités. Le logarithme de la fonction de vraisemblance que l'on trouvera à l'annexe B contient donc 24 termes, chacun des termes étant une intégrale différente de la loi normale en trois dimensions. La solution numérique est obtenue par une méthode itérative de gradient. Pour générer la solution initiale, on traite chaque équation séparément à l'aide du modèle probit ou probit ordonné, selon le cas.

1.1.3 Le modèle avec groupe hétérogène

Il est vraisemblable, comme nous le notions en introduction, que certaines personnes se retrouvent dans le modèle TNA alors que d'autres feront partis du groupe NTA. Malheureusement les données et l'information disponible ne permettent pas d'identifier à priori qui se retrouvera dans un groupe plutôt que dans l'autre. Afin de résoudre cette difficulté, nous allons supposer que l'on connaît les déterminants de l'appartenance d'un individu à un groupe ou à un autre, et laisser aux données le soin de déterminer la probabilité que chaque individu dans l'échantillon appartienne à l'un ou l'autre groupe.

L'équation d'appartenance à TNA ou NTA s'écrit de la manière suivante :

$$c_i^* = s_i \lambda + m_i \tag{7}$$

c_i^* est une variable latente qui mesure l'utilité d'être dans le modèle TNA ou le modèle NTA. Ces utilités ne sont pas observables mais on suppose que leurs déterminants sont connus. Dans l'équation (7), ils sont définis dans s . L'erreur résiduelle m_i est supposée suivre une loi normale standardisée de sorte que la probabilité que l'individu i appartienne à TNA est égale à :

$$\int_{-\infty}^{-s_i \lambda} f(m_i).dm_i$$

Dans le cas contraire, l'individu appartient à NTA si :

$$\int_{-s_i \lambda}^{\infty} f(m_i).dm_i$$

La fonction de vraisemblance liant un modèle TNA et NTA et l'équation (7) est simplifiée en supposant que l'erreur résiduelle m_i est indépendante des erreurs résiduelles de TNA et NTA. On obtient :

$$L = \int_{-\infty}^{-s_i \lambda} f(m_i)TNA.dm_i + \int_{-s_i \lambda}^{\infty} f(m_i)NTA.dm_i$$

avec TNA et NTA définis en annexe.

1.2 Les résultats économétriques

Pour faciliter la lecture des résultats, nous présentons au tableau 1 un sommaire des variables retenues dans les estimations économétriques et quelques statistiques descriptives avec données pondérées et non pondérées. Les statistiques des variables endogènes que notre modèle veut expliquer sont d'abord présentées. Le suréchantillonnage des décrocheurs apparaît clairement dans l'enquête avec 12 % selon les données pondérées contre 24 % sans tenir compte des pondérations⁴. En moyenne, les élèves du secondaire travaillent 12 heures semaine alors que 36 % d'entre eux déclarent ne pas avoir travaillé au cours de leur dernière année au secondaire. 5 % des élèves ont des notes inférieures à 60, alors qu'une bonne majorité soit 70 %, présentent des notes supérieures à 70 %. Les variables explicatives ont été regroupées selon les caractéristiques personnelles des individus (sexe, a échoué au primaire ou non, notes et heures travaillées) leur milieu socio-économique (famille monoparentale, nombre de changements d'écoles, école privée ou publique, éducation des parents) et leur environnement institutionnel et macroéconomique (âge légal d'abandon, salaire minimum, taux de chômage et lieu de résidence). 52 % des participants à l'enquête sont de sexe masculin et 15 % ont échoué une ou plusieurs années au primaire. Pour 42 % des élèves de notre échantillon, un des deux parents a suivi des

⁴ Rappelons que la différence entre le 12 % de décrocheurs de notre échantillon et le 18 % de l'enquête sur les sortants tient au fait que nous intégrons dans notre modèle des variables renseignées comme déterminants de l'abandon et du travail pendant les études secondaires.

études collégiales ou universitaires. Le nombre de changements d'écoles concerne tous les changements depuis le primaire. Cette variable veut refléter la dimension d'accueil et le sentiment d'appartenance à un groupe. Ces situations sont jugées importantes par les spécialistes en éducation (Tinto [1993]⁵), comme déterminants de l'attrition.

L'intérêt de notre banque de données réside dans son caractère national et dans l'horizon temporel qu'elle couvre (5 ans, de 1987 à 1991). Elle permet d'évaluer par une expérience naturelle l'effet de variables d'environnement sur les décisions des individus concernant l'abandon et le travail pendant les études. Par exemple, 69 % des individus de notre échantillon vivaient dans une province où l'âge légal de l'abandon scolaire était de 16 ans, alors que pour les autres cet âge était de 15 ans. La même situation de variabilité dans les données s'observe pour les salaires minimums et les taux de chômage, comme l'indiquent les valeurs minimales et maximales rapportées pour ces variables au tableau 1. Les variables croisées veulent capter des influences plus particulières de la variable « taux de chômage » dans des contextes définis. Par exemple, la variable croisée « taux de chômage » et « famille avec deux parents » vérifie l'hypothèse que l'influence du taux de chômage sur le travail des étudiants varie selon la structure familiale. Rappelons que la construction de ces variables est expliquée dans l'annexe A et que nous avons des variables différentes de taux de chômage et de salaires minimums pour l'abandon et le travail pendant les études. Le travail concerne la dernière année de l'étudiant au secondaire alors que l'abandon ou la persévérance scolaire sont spécifiques à une période. Par exemple, pour les décrocheurs, le taux de chômage et le salaire minimum sont ceux en vigueur au moment de l'abandon alors que pour les diplômés ou les persévérants, il s'agit du taux de chômage et du salaire minimum moyen compilés sur la période possible d'abandon. Les variables croisées concernant le taux de chômage prennent en compte le fait que pour les individus qui n'habitent pas une région métropolitaine de recensement, le taux de chômage utilisé est celui de la province⁶. Enfin, nous retenons des variables indiquant la province où les études secondaires ont été effectuées.

⁵ Tinto, V. (1993) : *Leaving CEGEP : Rethinking the Causes and Cures of Student Attrition*, Chicago : University of Chicago Press.

⁶ Il est regrettable que les responsables de l'enquête de Statistique Canada n'aient pas jugé opportun de codifier les codes postaux. Cette initiative, tout en garantissant la confidentialité des données, aurait permis d'introduire d'autres variables d'environnement macro-économique et d'être plus précis sur celles actuellement retenues.

Tableau 1
Statistiques descriptives

	Moyenne non pondérée	Moyenne pondérée
<i>Caractéristiques Personnelles :</i>		
Femme	0,50	0,48
Homme	0,50	0,52
Echoue l'année primaire	0,15	0,15
N'abandonne pas les études au secondaire	0,76	0,88
Abandon	0,24	0,12
Note inférieure à 60	0,05	0,05
Note comprise entre 60 et 70	0,30	0,26
Note supérieure à 70	0,65	0,70
Aucune heure travaillée	0,41	0,36
Entre 1 et 15 heures travaillées	0,24	0,29
Entre 15 et 30 heures travaillées	0,29	0,30
Plus de 30 heures travaillées	0,06	0,05
<i>Caractéristiques Socio-économiques :</i>		
Famille monoparentale	0,23	0,20
Famille avec les 2 parents	0,77	0,80
Nombre de changements d'écoles	2,48	2,46
Ecole publique	0,94	0,90
Ecole privée	0,06	0,10
Aucun parent avec éducation post - secondaire	0,63	0,58
Parent(s) avec éducation post - secondaire	0,37	0,42
<i>Caractéristiques Institutionnelles :</i>		
Province où l'âge légal d'abandon est 15 ans	0,31	0,31
Province où l'âge légal d'abandon est 16 ans	0,69	0,69
Salaire minimum (abandon)	3,81	3,80
	[2,97;4,65]	[2,97;4,65]
Salaire minimum (travail)	3,78	3,84
	[2,97;4,43]	[2,97;4,43]
Taux de chômage (abandon)	10,09	8,96
	[2,7;23,6]	[2,7;23,6]
Taux de chômage (travail)	9,73	8,56
	[3,72;19,5]	[3,72;19,5]

Tableau 1 (suite)
Statistiques descriptives

	Moyenne non pondérée	Moyenne pondérée
<i>Caractéristiques Institutionnelles :</i>		
Croisée taux de chômage et résidant hors RMR (abandon)	6,45	4,41
Croisée taux de chômage et résidant hors RMR (travail)	6,10	4,03
Croisée taux de chômage et éducation des parents	3,51	3,59
Croisée taux de chômage et famille avec 2 parents	7,61	6,86
Résidant de la Colombie-Britannique	0,12	0,12
Terre-Neuve	0,09	0,03
Ile du Prince Edouard	0,07	0,01
Nouvelle-Ecosse	0,09	0,04
Nouveau-Brunswick	0,07	0,04
Québec	0,14	0,25
Ontario	0,14	0,29
Manitoba	0,09	0,05
Saskatchewan	0,08	0,05
Alberta	0,11	0,11

Note: Source: Enquête sur les sortants, 1991. Taille de l'échantillon: 5584.
Voir l'Annexe A pour une description complète des variables et de leur construction.

[;] : Indique les valeurs minimum et maximum observées.

Le tableau 2 donne les résultats de l'estimation du modèle avec groupe hétérogène. Ce modèle permet de calculer les probabilités que l'individu soit de type TNA ou de type NTA. Pour fins d'explications, ce calcul peut être décomposé en trois étapes (mais rappelons qu'en pratique, l'estimation est simultanée). D'abord, dans la mesure où le modèle a convergé, la valeur estimée des paramètres a permis de calculer la fonction de vraisemblance du modèle TNA. Ensuite, les trois variables explicatives retenues pour le choix d'appartenance au modèle TNA ou NTA ont servi à calculer la borne. Enfin, on intègre le fait d'appartenir à TNA et on obtient la probabilité que l'individu soit de type TNA. On procédera de la même manière afin de calculer la probabilité d'être de type NTA. Suite à l'estimation du modèle avec groupe hétérogène, on obtient dans notre cas que la probabilité moyenne d'être de type TNA est égale à 53 % alors que celle d'être de type NTA est de 47 %.

On trouve dans la colonne de gauche la liste des variables retenues. Avant de présenter les résultats des estimations économétriques, soulignons deux remarques de nature méthodologique. Premièrement, on doit noter que notre utilisation de fonctions en paliers peut offrir plus de flexibilité qu'une simple relation linéaire ou même quadratique, au prix cependant, d'avoir à fixer les seuils des notes et des heures travaillées de façon arbitraire. Nous avons exploré différents

seuils avant de retenir ceux qui sont présentés ici. Deuxièmement, on aurait pu vouloir introduire l'effet des notes et/ou de la décision d'abandonner ou non dans l'équation du travail et, de façon similaire, on aurait pu vouloir ajouter l'effet de la décision concernant l'abandon dans l'équation des notes. Ceci poserait cependant des problèmes techniques sérieux. D'une part, si on introduit, par exemple, à droite de l'équation du travail, les variables observées D_{4i} et D_{5i} , cela pose des problèmes de cohérence du modèle (Maddala, 1983⁷). D'autre part, si on introduit parmi les variables explicatives de chacune des équations, les variables latentes apparaissant comme variables dépendantes dans les autres équations, cela ne pose pas de problème majeur si ces variables latentes apparaissent sous forme linéaire. Cependant, la forme réduite du modèle devient beaucoup plus complexe et la situation se complique considérablement si l'effet des variables latentes est supposé non linéaire.

Notre procédure d'estimation admet que les variables non observables dans les trois équations peuvent être corrélées. On note, en particulier, une corrélation positive élevée entre les erreurs résiduelles de l'équation d'abandon et celles de l'équation du travail pour le modèle TNA. Pour illustrer la portée de ce résultat, imaginons qu'une des variables non observées soit la motivation de l'élève pour les études. En principe, l'élève motivé (non motivé) sera moins (plus) porté à abandonner l'école et moins (plus) disposé à travailler pendant ses études. En l'absence de variable définie dans le but d'évaluer l'influence de l'intérêt pour les études sur les choix de l'élève (décrochage, travail pendant les études), il faut ainsi se rabattre sur les résidus des deux équations. Comme il s'agit des mêmes individus, ces résidus sont donc corrélés.

Notons également que les estimateurs sont obtenus pour les données pondérées⁸ et que des corrections d'hétéroscédasticité pour tenir compte de l'hétérogénéité non observée ont été effectuées. Comme l'hétérogénéité non observée se répercute sur les résidus des équations, la correction que nous avons apportée à ce problème consiste à supposer que la variance des résidus est fonction de l'éducation des parents qui est observée. L'idée consiste à supposer qu'à l'intérieur du groupe des élèves de parents moins éduqués, on devrait observer plus de variabilité résiduelle au sein des décisions liées au travail, à la performance scolaire et à l'abandon qu'au sein du groupe des élèves de parents éduqués. La variance des résidus diminue ainsi avec la scolarité des parents. Rappelons que de ne pas tenir compte de l'hétéroscédasticité dans le cas de modèles non linéaires peut produire des estimateurs non convergents.

⁷ Maddala, G. (1983) : *Limited Dependent and Qualitative Variables in Econometrics*, New-York : Cambridge University Press.

⁸ La procédure retenue est la suivante. Soit N le nombre d'observations et Z_i les poids initiaux de la banque de données (discutés dans l'annexe B de l'Enquête sur les Sortants de 1991 : Guide des utilisateurs des microdonnées).

Définissons $W_i = Z_i / \left(\sum_i^N Z_i \right)$ La fonction de vraisemblance à maximiser est la sommation sur i de $L_i W_i$ où L_i est la contribution non pondérée de l'individu i au logarithme de la fonction de vraisemblance.

Tableau 2
Le modèle avec groupe hétérogène

	Modèle TNA			Modèle NTA		
	Travail	Notes	Abandon	Notes	Travail	Abandon
<i>Caractéristiques Personnelles:</i>						
Femme	Réf.	Réf.	Réf.	Réf.	Réf.	Réf.
Homme	0,3825 (10,078)	-0,3176 (-5,530)	0,4001 (7,190)	-0,3714 (-6,806)	-0,0004 (-0,007)	-0,0988 (-0,506)
Echoue l'année primaire	-	-0,6190 (-8,886)	-	-0,7859 (-11,592)	-	--
Note inférieure à 60	Réf.	Réf.	Réf.	Réf.	Réf.	Réf.
Note comprise entre 60 et 70	-	-	-0,9772 (-5,458))	-	-0,3721 (-2,580)	-1,2326 (-2,685)
Note supérieure à 70	-	-	-1,5873 (-5,659)	-	-0,5012 (-2,073)	-2,3910 (-2,808)
H1 : nombre d'heures travaillées = 0	Réf.	Réf.	Réf.	Réf.	Réf.	Réf.
H2 : 0 < nombre d'heures travaillées ≤ 15	-	0,1309 (1,143)	-1,4775 (-12,725)	-	-	0,7201 (3,405)
H3 : 15 < nombre d'heures travaillées ≤ 30	-	-0,1167 (-0,748)	-2,3069 (-19,211)	-	-	1,4936 (5,115)
H4 : nombre d'heures travaillées > 30	-	-0,0300 (-0,134)	-3,3063 (-23,777)	-	-	-0,5538 (0,000)
<i>Caractéristiques Socio-économiques:</i>						
Famille monoparentale	Réf.	Réf.	Réf.	Réf.	Réf.	Réf.
Famille avec les 2 parents	0,0633 (1,643)	0,2339 (3,777)	-	0,1254 (1,966)	-0,1436 (-1,803)	-
Nombre de changements d'écoles	-	0,0851 (2,977)	0,0596 (2,435)	-0,0470 (-1,658)	-	0,1871 (2,403)
Ecole publique	Réf.	Réf.	Réf.	Réf.	Réf.	Réf.
Ecole privée	-0,1919 (-3,224)	-0,1722 (-2,035)	-0,4225 (-4,511)	0,7660 (3,836)	-0,2506 (-1,978)	-1,2810 (-0,611)
Aucun parent avec éducation post - secondaire	Réf.	Réf.	Réf.	Réf.	Réf.	Réf.
Parents avec éducation post - secondaire	0,0458 (0,897)	0,1662 (2,303)	-0,2446 (-2,701)	0,1481 (1,907)	0,1320 (1,475)	0,2842 (1,250)

Tableau 2 (suite)
Le modèle avec groupe hétérogène

	Modèle TNA			Modèle NTA		
	Travail	Notes	Abandon	Notes	Travail	Abandon
<i>Variables Institutionnelles et macro-économiques :</i>						
Province où l'âge légal d'abandon est 15 ans	Réf.	Réf.	Réf.	Réf.	Réf.	Réf.
Province où l'âge légal d'abandon est 16 ans	-	-	-0,4447 (-5,475)	-	-	-1,5767 (-5,618)
Salaire minimum	0,4751 (12,176)	-	0,8120 (8,799)	-	-0,3089 (-7,290)	2,4137 (5,601)
Taux de chômage	-0,5706 (-9,887)	-	-0,9021 (-12,348)	-	-0,2168 (-2,389)	-1,0013 (-4,221)
Croisée taux de chômage et résidant hors RMR	0,0684 (2,631)	-	0,2612 (5,961)	-	-0,2245 (-5,190)	0,4460 (3,389)
Croisée taux de chômage et éducation des parents	0,1311 (3,220)	-	-	-	0,2403 (3,078)	-
Croisée taux de chômage et famille avec 2 parents	0,1867 (4,023)	-	-	-	-0,1932 (-2,237)	-
Résidant de la Colombie-Britannique	Réf.	Réf.	Réf.	Réf.	Réf.	Réf.
Terre-Neuve	-0,2177 (-1,217)	0,2189 (1,065)	-0,3110 (-1,305)	0,3586 (1,450)	-0,9048 (-0,845)	-1,8443 (-2,147)
Ile du Prince-Edouard	0,5245 (0,687)	0,4014 (0,598)	1,1392 (1,325)	0,3800 (0,475)	0,0001 (0,000)	1,9782 (0,416)
Nouvelle-Ecosse	0,0237 (0,152)	0,4552 (2,531)	-0,1075 (-0,567)	0,1670 (0,789)	-0,4337 (-1,723)	-0,0912 (-0,130)
Nouveau-Brunswick	-0,0068 (-0,045)	0,7147 (3,725)	-0,5219 (-2,168)	0,1085 (0,573)	0,0818 (0,310)	-1,5815 (-3,571)
Québec	-0,9169 (-10,460)	1,3284 (13,600)	-1,7840 (-8,714)	0,6669 (6,699)	0,2087 (1,936)	-4,0557 (-5,766)
Ontario	-0,7661 (-8,632)	0,6588 (7,882)	-1,8504 (-8,781)	0,2643 (2,726)	0,3236 (3,368)	-1,2395 (-2,892)
Manitoba	-0,7269 (-4,937)	0,7261 (3,482)	-1,1944 (-6,780)	0,1152 (0,547)	0,2524 (1,401)	-2,3659 (-3,739)
Saskatchewan	-0,7934 (-6,022)	0,4829 (3,112)	-2,2151 (-9,119)	0,1042 (0,584)	-0,2220 (-1,053)	-5,3063 (-4,926)
Alberta	0,3779 (4,300)	-0,0043 (-0,045)	-0,0822 (-0,630)	0,0260 (0,219)	-0,4037 (-3,715)	0,0677 (0,153)

Tableau 2 (suite)
Le Modèle avec groupe hétérogène

	Modèle TNA			Modèle NTA		
	Travail	Notes	Abandon	Notes	Travail	Abandon
<i>Autres:</i>						
Constante	1,0371 (10,443)	1,6131 (10,426)	3,4538 (10,465)	1,6780 (14,841)	0,1001 (0,454)	1,9381 (2,557)
σ (Education des parents)	-0,2129 (-5,354)	-0,1760 (-2,477)	-0,0635 (-1,141)	-0,0159 (-0,279)	-0,2143 (-2,719)	-0,9998 (-2,896)
ε_1		1,0280 (17,7657)			-	
ε_2		2,3037 (32,2316)			-	
θ_1		1,6125 (22,1473)			-	
Ω_1		-			1,3134 (24,6387)	
ζ_1		-			0,6918 (11,3143)	
ζ_2		-			9,2108 (3,6778.10 ⁻⁵)	
Constante			-0,0748 (-1,159)			
Ecole publique			Réf.			
Ecole privée			-0,2604 (-2,804)			
Aucun parent avec éducation post - secondaire			Réf.			
Parents avec éducation post - secondaire			0,0460 (0,761)			
<i>Vraisemblance</i>			-11203,57			

Note : Les t - statistiques apparaissent entre parenthèses.

Les coefficients de corrélations pour le modèle TNA sont :

ρ_{12} (travail, notes)	-0,0932	(-1,2761)
ρ_{13} (travail, abandon)	0,8929	(31,1098)
ρ_{23} (notes, abandon)	0,2133	(2,0433)

Les coefficients de corrélations pour le modèle NTA sont :

ρ_{12} (notes, travail)	0,1068	(1,0401)
ρ_{13} (travail, abandon)	-0,1180	(-0,4574)
ρ_{23} (notes, abandon)	-0,9096	(-8,8634)

L'étude des résultats de l'estimation du modèle avec groupe hétérogène permet de mettre en évidence l'effet "opposé" de certaines variables dans les différentes équations, selon que l'on se situe dans le cadre du modèle TNA ou du modèle NTA. Par effet opposé, on entend que la variable a un signe positif dans le modèle TNA alors que son signe est négatif dans le modèle NTA, ou inversement. Ces variables, auxquelles il semble primordial d'accorder un intérêt particulier, sont au nombre de quatre dans l'équation du travail. Il s'agit du "salaire minimum", des variables "croisée taux de chômage et résidant hors RMR" et "croisée taux de chômage et famille avec deux enfants", et de certaines "provinces" telles que le Québec, l'Ontario ou l'Alberta. Dans l'équation des notes, ces variables sont le "type d'école" (public ou privée) et le "nombre de changements d'écoles". Dans l'équation de l'abandon enfin, il s'agit de la variable concernant le "nombre d'heures travaillées".

Si l'on commence par l'équation du travail, il apparaît d'abord très intuitif d'obtenir un effet positif du salaire minimum sur le nombre d'heures travaillées dans le modèle TNA. Selon Card et Krueger⁹ (1995), plusieurs études récentes concluent d'ailleurs à un effet positif (ou nul) du salaire minimum sur le travail des jeunes. L'effet positif s'observe notamment dans le secteur de la restauration rapide (une situation de monopsonne local) où se trouvent vraisemblablement plusieurs étudiants du secondaire. L'explication de l'effet négatif de cette variable dans le modèle NTA est cependant un peu moins intuitive. On peut considérer que l'augmentation du salaire minimum entraîne un excès d'offre de main d'œuvre sur le marché du travail, dans la mesure où cela abaisse le salaire de réserve des individus qui ne travaillaient pas. Dans ce cas, on peut imaginer que la sélection à l'embauche se fera davantage sur des critères telles que les notes, même dans des emplois qui à priori n'exigeaient pas de telles caractéristiques. Or, ce sont pourtant ces emplois qui étaient majoritairement prisés par les "mauvais" élèves. On peut alors mettre en évidence que les individus qui avaient décidé de travailler à cause de leur mauvais résultats scolaires (cas du modèle NTA) risquent de moins travailler si le salaire minimum augmente. Un tel raisonnement pourrait ainsi expliquer l'effet négatif de la variable salaire minimum sur le travail dans le modèle NTA.

Ensuite, l'effet de la variable croisée taux de chômage et résidant hors RMR est positif dans le modèle TNA et négatif dans le modèle NTA. Pourtant, le taux de chômage a un effet négatif dans les deux modèles, et il semblerait donc que ce soit le fait d'être résidant ou non d'une RMR qui soit déterminant sur le signe de la variable croisée. Étant donné que ce signe diffère, on peut alors penser que le fait d'être résidant ou non d'une RMR ne représente pas le même effet selon le modèle considéré. On peut alors avancer que dans le modèle TNA, le fait d'être résidant ou non d'une RMR représente essentiellement la situation économique de la région. Par contre, comme le travail est fonction des notes dans le modèle NTA, le fait d'être résidant ou non d'une RMR illustre peut être davantage la situation scolaire de la région.

On note également un effet positif de la variable croisée taux de chômage et famille avec deux parents dans le modèle TNA. Cela signifie que l'élève qui vit avec deux parents ajuste moins ses heures travaillées si le taux de chômage augmente sur le marché du travail que l'élève d'une

⁹ Card, D. et Alan B. Krueger (1995) : *Myth and Measurement : The New Economics of the Minimum Wage*, Princeton : Princeton University Press.

famille monoparentale. Une explication possible de ce résultat est l'effet réseau de la famille biparentale qui protège l'emploi de l'élève lorsque le taux de chômage s'accroît. Dans le cas du modèle NTA, l'effet négatif laisse supposer que dans la mesure où le travail est fonction des notes, l'élève moyen d'une famille avec deux parents ajuste plus ses heures travaillées selon la condition d'emploi sur le marché du travail que l'élève d'une famille monoparentale. On peut penser que dans une telle famille, la valorisation des études sera mise en avant si le taux de chômage augmente sur le marché du travail.

Enfin, si l'on prend les exemples du Québec ou de l'Ontario, l'effet relatif aux provinces est négatif dans le modèle TNA et positif dans le modèle NTA. Dans la mesure où l'on s'intéresse à l'effet de la province directement sur le nombre d'heures travaillées (modèle TNA), on peut intuitivement penser que l'on se réfère davantage à la situation économique de la province. Dans notre cas, le fait d'habiter au Québec ou en Ontario réduit le nombre d'heures travaillées, ce qui refléterait une situation économique défavorable pour les jeunes comparativement aux autres provinces. Par contre, si l'on s'intéresse à l'effet de la province étant donné que le nombre d'heures travaillées a été déterminé en fonction des notes (modèle NTA), on peut penser que la province illustre plutôt la situation du système scolaire. Il semblerait que les individus travaillent plus lorsque leurs notes sont mauvaises, quand ils habitent dans la province du Québec ou de l'Ontario. On est alors en droit de se demander si les systèmes scolaires Québécois ou Ontarien sont vraiment efficaces contre le décrochage scolaire.

Par ailleurs, les déterminants du travail pendant les études montrent que le travail rémunéré est plus fréquent chez les garçons que chez les filles dans le modèle TNA, alors que cette variable n'est pas significative dans le modèle NTA. Le nombre d'heures de travail ne varie pas selon le niveau d'éducation des parents. Les élèves de l'école privée travaillent significativement moins d'heures que ceux fréquentant l'école publique dans les deux modèles. La variable est moins significative dans le modèle NTA, mais a un impact plus fort que dans le modèle TNA. La variable "taux de chômage" semble être l'élément déterminant pour expliquer les heures travaillées dans le modèle TNA, avec un coefficient estimé négatif élevé et statistiquement significatif. L'effet de cette variable est beaucoup moins important dans le modèle NTA. On pouvait s'attendre à ce que l'éducation des parents modifie l'effet du taux de chômage sur le travail pendant les études. De ce fait, les coefficients "croisée taux de chômage et éducation des parents" sont positifs et statistiquement significatifs dans les deux modèles. Pour finir, on notera que dans le modèle NTA, les variables des notes confirment le fait, de part leur signe négatif, que plus l'individu a de bonnes notes plus son nombre d'heures travaillées diminue. Ces variables sont significatives.

Si l'on poursuit avec l'équation des notes, on constate en premier lieu que l'effet du nombre de changements d'écoles sur les performances scolaires de l'individu semble intuitivement être négatif comme l'illustre le modèle NTA. En effet, on peut imaginer que l'élève aura à s'adapter chaque fois qu'il change d'école et que ses résultats scolaires en pâtiront. Dans le cas du modèle TNA, si l'on considère que le fait de changer d'école oblige l'individu à quitter son travail et à en trouver un autre, on peut alors avancer que l'effet sur les notes pourrait être positif puisque les notes sont fonction du travail.

Dans un deuxième temps, on constate que le type d'école a un effet négatif dans le modèle TNA et positif dans le modèle NTA. L'impact direct positif de l'école privé sur les notes peut être expliqué par le fait qu'il y a généralement un meilleur encadrement scolaire dans un environnement privé. Toutefois, pour les individus qui ont des mauvaises notes suite à leur travail, le fait d'être dans une école privée semble avoir un impact négatif.

Dans l'équation qui concerne la note globale obtenue par l'étudiant dans le dernier trimestre de fréquentation de l'école secondaire, on constate que les variables macroéconomiques comme le taux de chômage et le niveau du salaire minimum n'apparaissent pas car ces facteurs ne semblent pas avoir d'effet significatif. De plus, on constate qu'en moyenne, les hommes réussissent moins bien que les femmes, dans les deux modèles. Le fait d'appartenir à une famille où les deux parents sont présents a un effet positif sur le succès scolaire dans les deux modèles. Par contre, le fait d'avoir dû reprendre une année à l'école primaire pour cause d'échec s'avère un indice négatif important dans les deux modèles. Les coefficients des variables provinciales montrent des différences assez notables, ce qui pourrait à la limite, refléter simplement des divergences dans les échelles d'attribution des notes. On notera cependant que six provinces sont significatives et positives dans le modèle TNA (Nouvelle-Écosse, Nouveau-Brunswick, Québec, Ontario, Manitoba et Saskatchewan) alors que deux seulement le sont dans le modèle NTA (Québec, Ontario). En ce qui concerne l'effet direct des heures travaillées sur les notes que l'on retrouve uniquement dans le modèle TNA (cf. spécification du modèle), cette variable n'est pas significative.

Si l'on s'intéresse maintenant à l'équation d'abandon, on remarque que la variable du nombre d'heures travaillées est significative et négative dans le modèle TNA, alors qu'elle est positive et non significative au-delà de 30 heures dans le modèle NTA. Les résultats du modèle NTA sont très intuitifs car ils suggèrent que plus l'individu travaille, plus il a tendance à abandonner. Par contre, les résultats du modèle TNA sont plus surprenants puisque plus l'individu travaille, moins il a tendance à abandonner. Toutefois, il est difficile d'émettre des explications sans prendre en compte l'effet des heures travaillées sur les notes dans le modèle TNA ou l'effet des notes sur les heures travaillées dans le modèle NTA. Ceci sera fait ultérieurement à l'aide de l'étude des coefficients de corrélation et des probabilités conditionnelles.

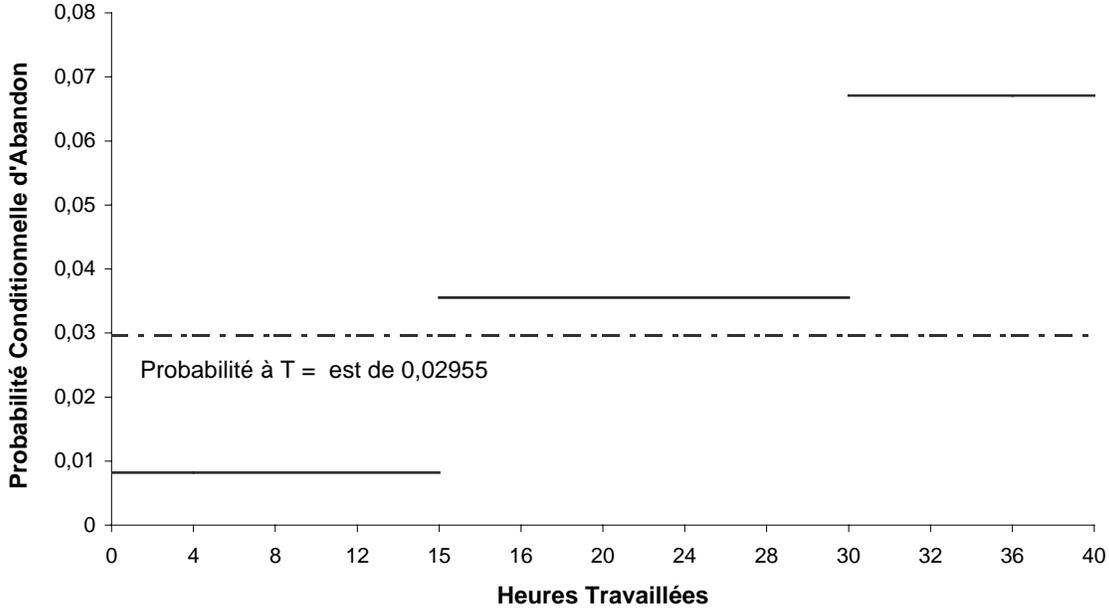
De plus, les résultats suggèrent que les hommes ont une plus grande propension à abandonner leurs études que les femmes dans le modèle TNA, alors que le sexe ne semble pas avoir d'effet dans le modèle NTA. Même chose en ce qui concerne le fait d'être dans une école privée. Les élèves qui ont subi plusieurs changements d'école depuis le secondaire ont également un taux d'attrition plus élevé dans les deux modèles, avec un effet beaucoup plus fort dans le modèle NTA. L'accroissement de l'âge légal d'abandon semble aussi avoir un effet dissuasif, effet plus important dans le cas du modèle TNA. On remarque qu'un taux de chômage faible incite davantage les élèves à décrocher. Par ailleurs, un salaire minimum élevé encourage aussi le décrochage scolaire avec un effet beaucoup plus élevé dans le modèle TNA que dans le modèle NTA. Ces résultats illustrent bien le concept économique de coût d'opportunité. Il est assez évident que dans ces conditions, la pire des situations pour l'abandon scolaire est un taux de chômage faible qui augmente à la fois les heures travaillées au cours des études et exerce une

pression pour hausser le salaire minimum. Un pareil scénario risque d'exercer un effet dévastateur sur la persévérance scolaire. Les comportements varient aussi de façon marquée entre les provinces.

Afin d'estimer les probabilités d'être dans le modèle TNA ou le modèle NTA, il a été nécessaire de faire un choix pertinent concernant les variables à retenir. Les variables choisies devaient l'être en fonction de leur pouvoir explicatif par rapport au fait que l'individu agit selon la catégorie TNA ou NTA. Seules des caractéristiques personnelles ou socio-économiques sont donc pertinentes dans la mesure où ce sont ces variables uniquement qui peuvent expliquer l'appartenance de l'individu au modèle TNA ou au modèle NTA. En effet, les variables institutionnelles et macroéconomiques ne peuvent être utilisées puisque l'on peut difficilement considérer qu'elles influenceront l'individu au préalable. Elles influenceront l'individu dans ses choix certes, mais une fois que l'on aura déterminé s'il appartient à la catégorie TNA ou NTA. Les variables retenues sont donc le type d'école et l'éducation des parents. On peut voir que la variable concernant le type d'école est significative et négative, c'est à dire que le fait d'être dans une école privée (comparativement au fait d'être dans une école publique) augmente la probabilité que l'individu soit de type TNA, autrement dit diminue la probabilité qu'il soit de type NTA. Par contre, la variable concernant l'éducation des parents n'est pas significative.

Si l'on s'intéresse désormais aux coefficients de corrélation, on remarquera rapidement que si le signe du coefficient est positif dans le cas du modèle TNA, il est alors négatif dans le cas du modèle NTA (ou inversement). Observons plus précisément les coefficients de corrélation entre les résidus de l'équation des heures travaillées et celle de l'abandon. Il est positif et très élevé (de l'ordre de 0,9) dans le modèle TNA et, comme il vient d'être mentionné, négatif et faible (de l'ordre de 0,12) dans le modèle NTA. On s'attendrait donc à ce que le travail augmente la probabilité d'abandonner dans le modèle TNA et diminue la probabilité d'abandonner dans le modèle NTA. Or, si l'on considère les coefficients des variables qui représentent les heures travaillées, on aboutit alors à une sorte de contradiction entre ce que suggèrent les coefficients de corrélation et les coefficients des variables. En effet, ces derniers impliquent que dans le cas du modèle TNA, plus le nombre d'heures travaillées augmente, plus la probabilité que l'individu abandonne est faible. Dans le cas du modèle NTA par contre, plus le nombre d'heures travaillées augmente, plus la probabilité que l'individu abandonne est forte. Afin d'analyser le phénomène de façon adéquate, tout laisse donc penser qu'il faut tenir compte de la forte corrélation entre les erreurs résiduelles des deux équations. Une évaluation correcte de la probabilité d'abandonner de l'étudiant moyen, conditionnelle au nombre d'heures travaillées et marginale par rapport aux notes scolaires, projette une vision différente. Cette évaluation qui donne l'effet total des heures travaillées (directement ou indirectement via leurs effets sur les notes et ces dernières sur l'abandon) sur cette probabilité d'abandon est reproduite à la figure 1.

**Figure 1. Modèle avec groupe hétérogène
Probabilité Conditionnelle d'Abandon (Données Pondérées)**



La probabilité conditionnelle d'abandon¹⁰ représente la probabilité que l'individu abandonne conditionnellement au nombre d'heure qu'il a choisi de travailler. On constate, en examinant la figure 1, que la propension à abandonner est plus faible pour l'étudiant qui travaille quelques heures (moins de 15 heures) que pour l'étudiant qui ne travaille pas du tout. Cependant, pour l'étudiant qui travaille plus de 15 heures, cette probabilité devient supérieure à celle de l'étudiant qui ne travaille pas et elle augmente encore pour l'étudiant travaillant plus de 30 heures. Le fait

¹⁰ La probabilité d'abandon conditionnelle a été calculée de la manière suivante :

$$P(A/T) = P(A, L_1 / T) + P(A, L_2 / T)$$

avec :

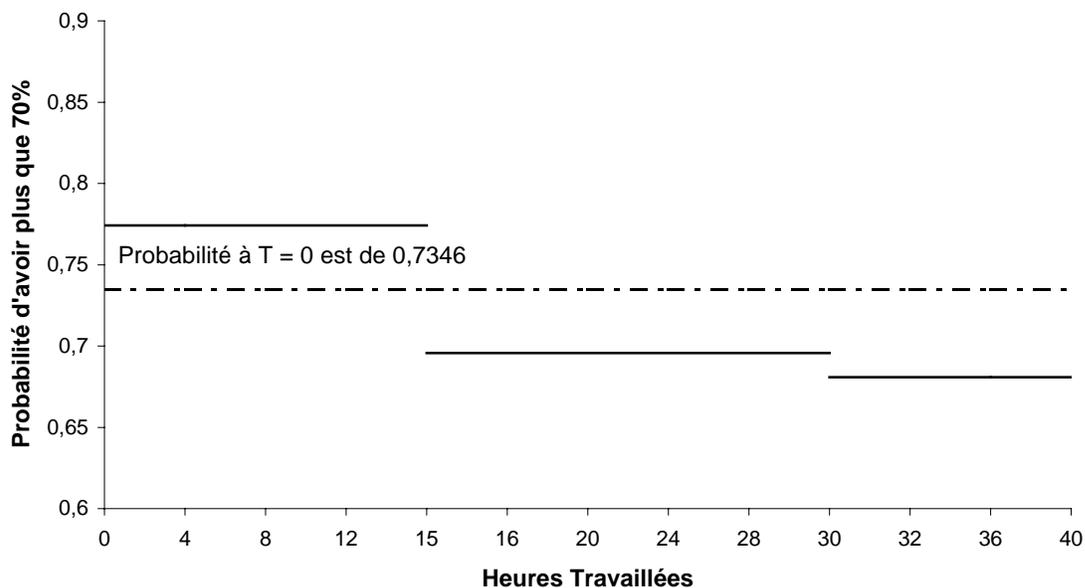
$$P(A, L_1 / T) = \frac{\sum_{K=0}^2 P(A, T, L_1, N_K)}{\sum_{i=0}^1 \sum_{j=1}^2 \sum_{k=0}^2 P(A_i, T, L_j, N_k)}$$

$$P(A, L_2 / T) = \frac{\sum_{K=0}^2 P(A, T, L_2, N_K)}{\sum_{i=0}^1 \sum_{j=1}^2 \sum_{k=0}^2 P(A_i, T, L_j, N_k)}$$

Notons que L_1 représente le modèle TNA et que L_2 représente le modèle NTA.

que l'étudiant qui travaille un nombre limité d'heures pas semaine ait tendance à moins abandonner doit peut-être s'interpréter à la lumière des résultats de l'étude faite aux Etats-Unis par Turner¹¹. Selon cette étude, l'élève en classe terminale du secondaire passe en moyenne 18 heures pas semaine à regarder la télévision et moins de 4 heures à faire ses devoirs. Le fait de travailler 20 heures par semaine réduit le temps d'études à la maison de 7,2 minutes et le temps consacré à la télévision de 3,6 heures. Il se peut donc que dans le contexte culturel nord-américain, le jeune qui est ambitieux et dynamique soit motivé à la fois pour bien réussir ses études et réduire le coût d'opportunité de celles-ci en travaillant à temps partiel.

**Figure 2. Modèle avec groupe hétérogène
Probabilité conditionnelle d'avoir des notes supérieures à 70
(Données Pondérées)**



Il faut noter que l'on observe le même genre de phénomène dans l'équation des notes, quand on évalue la probabilité d'avoir une note élevée ou faible, conditionnellement au nombre d'heures travaillées. On constate, en examinant la figure 2, que la propension à voir des notes supérieures à 70% est plus élevée pour l'étudiant qui travaille quelques heures (moins de 15 heures) que pour l'étudiant qui ne travaille pas du tout. Cependant, pour l'étudiant qui travaille plus de 15 heures, cette probabilité devient inférieure à celle de l'étudiant qui ne travaille pas et elle diminue encore pour l'étudiant travaillant plus de 30 heures pas semaine.

¹¹ Turner, Mark D. (1994) : *The effects of Part-Time Work on High School Students' Academic Achievement*, Mimeo, University of Maryland.

2. Conclusion

L'objectif de ce rapport était d'étudier les déterminants de l'abandon scolaire et du travail pendant les études en tenant compte du processus de décision. La caractéristique particulière de cette recherche a été de profiter de l'enquête rétrospective menée à l'échelle du pays par Statistique Canada en collaboration avec le Développement des ressources humaines Canada sur les sortants de l'école secondaire. La dimension nationale de l'enquête nous a permis de disposer d'une expérience naturelle pour comprendre le rôle de certaines variables macro-économiques et des institutions sur les déterminants de l'abandon aux études secondaires et du travail pendant les études. À cette variabilité entre provinces s'est ajoutée une variabilité temporelle puisque l'enquête portait sur des jeunes âgés de 18-20 ans en 1991 qui entre 1987 et 1991, correspondaient à l'un ou l'autre des trois statuts suivants : diplômé, persévérant ou sortant.

Ce travail a permis de mettre en évidence l'importance de tenir compte de l'hétérogénéité d'un groupe d'individus. Ce constat s'avère particulièrement intéressant dans la mesure où l'on peut se demander dans quelles proportions les différences de comportements ne seraient pas fonction du sexe de l'individu. En effet, les résultats suggèrent que les hommes ont une plus grande propension à abandonner leurs études que les femmes dans le modèle TNA, alors que le sexe ne semble pas avoir d'effet dans le modèle NTA. À cet égard, une version révisée de ce texte tiendra compte plus spécifiquement du sexe de l'individu. La prise en compte de l'hétérogénéité est d'autant plus pertinente que certaines conclusions sont modifiées selon le groupe pris en considération, sans toutefois remettre en question celles qui ont antérieurement été établis.

Un résultat important qui perdure dans cette recherche est le rôle du salaire minimum dans le phénomène du décrochage. Il est clair que le salaire minimum, s'il est élevé, contribue de façon significative à accroître l'abandon au secondaire. Chaque fois qu'ils procèdent à des hausses du salaire minimum, les gouvernements provinciaux créent donc des conditions propices à une recrudescence de l'abandon scolaire. Les premiers à subir les effets de ces hausses sont d'ailleurs les individus « à la marge de la décision » sortir/terminer les études secondaires. Néanmoins le fait de tenir compte de l'hétérogénéité du groupe permet de mettre en évidence que le salaire minimum n'a pas le même effet sur le travail selon le modèle considéré. Pour le modèle TNA, les coefficients confirment l'intuition qu'une hausse du salaire minimum augmente le travail pendant les études et augmente l'abandon. Par contre, pour le modèle NTA, les coefficients suggèrent qu'une hausse du salaire minimum réduit le travail pendant les études. L'explication est que la hausse du salaire minimum attire davantage de monde sur le marché du travail et évince donc les individus qui ont les moins bons résultats scolaires. On peut raisonnablement penser alors que si les élèves ont de mauvaises notes et n'arrivent pas à travailler pendant les études, ils vont certainement prendre la décision, aussi radicale soit elle, d'abandonner l'école.

Un autre résultat important concernait l'influence du nombre d'heures travaillées sur l'abandon. On avait obtenu dans le rapport précédent que pour plusieurs étudiants, travailler moins de 10 heures par semaine était loin d'inciter au décrochage, et pour ceux d'entre eux qui obtiendraient leur diplôme, l'expérience acquise parallèlement aux études pouvait même favoriser leur insertion sur le marché de l'emploi. Or, tenir compte de l'hétérogénéité du groupe a permis de mettre en relief que l'augmentation du nombre d'heures travaillées diminue l'abandon dans le

modèle TNA mais augmente l'abandon dans le modèle NTA. Cependant, il faut être très prudent pour l'interprétation d'un tel phénomène dans la mesure où l'abandon est fonction à la fois des heures travaillées mais aussi des notes et que selon le modèle considéré, les notes sont fonction des heures travaillées (modèle TNA) ou les heures travaillées fonction des notes (modèle NTA). Le fait de se référer aux probabilités conditionnelles permet d'avoir un aperçu plus global de l'impact des heures travaillées sur l'abandon puisqu'elles tiennent compte de la corrélation existant entre le travail, les notes et l'abandon. Les résultats obtenus à l'aide des probabilités conditionnelles restent conformes aux résultats antérieurs mais ne permettent pas d'expliquer la différence de signe selon le modèle considéré. Il semble donc intéressant de poursuivre les recherches à ce sujet.

Par ailleurs de nombreux résultats sont restés semblables. La baisse du taux de chômage peut inciter les étudiants à abandonner leurs études très tôt, même avant la fin du secondaire. Cette constatation doit donc inciter les autorités gouvernementales, à bien cibler leurs objectifs lorsqu'elles mettent en place des politiques visant à réduire le chômage chez les jeunes. Il faut éviter, en effet, que les politiques proposées n'incitent certains élèves à décrocher avant la fin du secondaire. Enfin, l'augmentation de l'âge légal d'abandon scolaire réduirait l'abandon scolaire.

Il reste toutefois à préciser que la prise en compte de ce modèle combiné a permis de soulever quelques problèmes méthodologiques. Un des problèmes rencontrés concerne les catégories établies pour les notes et les heures travaillées. Certaines catégories ne contiennent pas assez d'observations si bien que l'étude par sexe induira de les changer. Un autre problème concerne la difficulté d'interpréter les résultats dans la mesure où le travail, les notes et l'abandon sont corrélés. À ce sujet, la prise en compte des résidus corrélés aux autres résidus du modèle dans l'équation de sélection, est un prolongement de recherche actuellement en cours de réalisation.

Annexe A

Les données – définition et construction

Les données utilisées proviennent de l'Enquête auprès des Sortants (1991). Les données regroupent près de 6000 jeunes étudiants qui sont soit diplômé (ou persévéré) ou ont abandonné l'école secondaire entre 1987 et 1991.¹² Cette enquête posait des questions rétrospectives à propos des motifs de l'abandon scolaire chez les jeunes au secondaire de façon à identifier les facteurs socioéconomiques qui contribuent à l'abandon scolaire chez les jeunes. De plus, afin de mieux comprendre le rôle de la conjoncture économique sur l'abandon scolaire, un certain nombre de variables à caractère économique qui *a priori* n'étaient pas fournies par l'enquête, ont été construites à l'aide des données de Statistique Canada. Bien que la construction de certaines d'entre elles fut certes ardue, nul doute que la richesse de ce type de variables motivait grandement l'insertion de celles-ci dans l'analyse du choix de l'individu à l'égard de l'abandon scolaire. Les variables spécifiques retenues pour expliquer l'abandon scolaire sont :¹³

A.1 Spécification travail

Travail : Nombre d'heures travaillées au cours de la dernière année de fréquentation de l'école secondaire. 4 catégories: 0 heures, entre 1 et 15 heures, entre 15 et 30 heures et plus de 30 heures travaillées par semaine.¹⁴

Homme : Variable dichotomique qui indique si l'individu est un homme. 1 si l'individu est un homme, 0 si l'individu est une femme.

Famille avec deux parents : Variable dichotomique qui indique si l'étudiant habitait ou non avec ses deux parents au cours de ses études. 1 s'il vit avec ses deux parents, 0 sinon.

École privée : Variable dichotomique qui indique si l'individu a déjà fréquenté une école privée lors de son primaire ou secondaire. 1 si oui, 0 sinon.

Éducation des parents : Variable dichotomique qui indique le niveau d'éducation des parents. 1 si un des deux parents a suivi (complété ou non) un diplôme d'étude collégial ou universitaire, 0 autrement.

Salaire minimum : Variable construite à partir des données de Statistique Canada. Cette variable représente le salaire minimum réel en dollars constants de 1996 pertinent pour l'individu considéré, lors de sa dernière année d'étude. Si l'individu étudie ou a étudié dans une région métropolitaine de recensement (RMR), la salaire minimum est déflaté par l'indice des prix à la consommation de la région appropriée, tandis que si l'individu étudie ou a étudié en dehors d'une

¹² À noter que l'enquête originale comporte 9460 jeunes. Néanmoins, étant donné la non-disponibilité de l'information pour un grand nombre d'individus, l'étude empirique est effectuée à partir d'un échantillon plus restreint, plus précisément un échantillon de 5584 observations. Voir également la note 26 du texte.

¹³ À noter qu'une attention particulière sera accordée à la méthodologie de la construction des variables provenant des données de Statistique Canada.

¹⁴ Les données disponibles sont continues. Nous avons cependant préféré utiliser les catégories indiquées pour les raisons exposées plus haut.

région métropolitaine de recensement reconnue, le salaire minimum est déflaté par l'indice des prix à la consommation de la province appropriée.

Taux de chômage : Variable construite à partir des données de Statistique Canada. Cette variable représente le taux de chômage qui prévaut au sein de la RMR ou de la province appropriée pour l'individu considéré. À noter que le taux de chômage des 15 ans ou plus a été utilisé étant donné la non-disponibilité du taux de chômage des 15-24 ans au niveau des régions métropolitaines de recensement.

Croisée taux de chômage et région : Il s'agit d'une variable croisée qui est le produit du taux de chômage et d'une variable dichotomique qui distingue si l'individu étudié ou a étudié dans une région métropolitaine de recensement ou en région rurale.

Croisée taux de chômage et éducation des parents : Il s'agit d'une variable croisée qui est le produit du taux de chômage et de la variable dichotomique portant sur l'éducation des parents. Taux de chômage*1 si un des deux parents a suivi (complété ou non) un cours d'étude collégial ou universitaire, taux de chômage*0 si aucun des parents n'a suivi un cours d'étude collégiale.

Croisée taux de chômage et famille avec deux parents : Il s'agit d'une variable croisée qui est le produit du taux de chômage et de la variable dichotomique famille avec deux parents (famille bi-parentale ou non). Taux de chômage*1 si l'étudiant habitait avec ses deux parents au cours de ses études, taux de chômage*0 autrement.

Résidence : Variables dichotomiques qui indiquent dans quelle province les études secondaires ont été réalisées.

A.2 Spécification note

Notes : Moyenne au cours du dernier trimestre complet d'études primaires ou secondaires. Il y a 3 catégories : moins de 60%, entre 60 et 70% et plus de 70%.¹⁵

Homme : Voir définition dans la spécification travail.

Échoue : Variable dichotomique qui indique si l'individu a échoué une année d'étude à l'école primaire. 1 si l'individu a échoué, 0 sinon.

Note inférieure à 60 : Variable dichotomique qui indique la moyenne au cours du dernier trimestre complet d'études secondaires de l'étudiant. 1 s'il avait moins de 60%, 0 sinon.

Note comprise entre 60 et 70 : Variable dichotomique qui indique la moyenne au cours du dernier trimestre complet d'études secondaires de l'étudiant. 1 s'il avait entre 60 et 70%, 0 sinon.

Note supérieure à 70 : Variable dichotomique qui indique la moyenne au cours du dernier trimestre complet d'études secondaires de l'étudiant. 1 s'il avait plus de 70%, 0 sinon.

¹⁵ Il existe aussi une catégorie séparée pour 80 % et plus. Nous avons agrégé cette catégorie avec la catégorie : entre 70 % et 80 %, afin de réduire le nombre de catégories. Nos expériences préliminaires suggéraient que le comportement des individus dans la catégorie 80 % et plus n'était pas significativement différent de celui des individus dans la catégorie 70 % à 80 %.

Travail entre 1 et 15 heures : Variable dichotomique qui indique le nombre d'heures travaillées par semaine par l'individu. 1 s'il travaille entre 1 et 15 heures, 0 autrement.

Travail entre 15 et 30 heures : Variable dichotomique qui indique le nombre d'heures travaillées par semaine par l'individu. 1 s'il travaille entre 15 et 30 heures, 0 autrement.

Travail plus de 30 heures : Variable dichotomique qui indique le nombre d'heures travaillées par semaine par l'individu. 1 s'il travaille plus de 30 heures, 0 autrement.

Famille avec deux parents : Voir définition dans la spécification travail.

Nombre de changements d'écoles : Variable qui indique le nombre de changements d'école (passage du primaire au secondaire compris).

École privée : Voir définition dans la spécification travail.

Éducation des parents : Voir définition dans la spécification travail.

Résidence : Voir définition dans la spécification travail.

A.3 Spécification abandon

Abandon : Variable dichotomique qui indique si l'individu a abandonné l'école secondaire. 1 si l'individu a abandonné, 0 sinon.

Homme : Voir définition dans la spécification travail.

Note inférieure à 60 : Variable dichotomique qui indique la moyenne au cours du dernier trimestre complet d'études secondaires de l'étudiant. 1 s'il avait moins de 60%, 0 sinon.

Note comprise entre 60 et 70 : Variable dichotomique qui indique la moyenne au cours du dernier trimestre complet d'études secondaires de l'étudiant. 1 s'il avait entre 60 et 70%, 0 sinon.

Note supérieure à 70 : Variable dichotomique qui indique la moyenne au cours du dernier trimestre complet d'études secondaires de l'étudiant. 1 s'il avait plus de 70%, 0 sinon.

Travail entre 0 et 15 heures : Voir définition dans la spécification note.

Travail entre 15 et 30 heures : Voir définition dans la spécification note.

Travail plus de 30 heures : Voir définition dans la spécification note.

Famille avec deux parents : Voir définition dans la spécification travail.

Nombre de changements d'écoles : Voir définition dans la spécification note.

École privée : Voir définition dans la spécification travail.

Éducation des parents : Voir définition dans la spécification travail.

Âge légal : Variable dichotomique qui indique si, lorsque l'individu a abandonné l'école, l'âge légal d'abandon de sa province était de 15 ou 16 ans. 1 si 16 ans, 0 si 15 ans.

Salaire minimum : Variable construite à partir des données de Statistique Canada. Cette variable représente le salaire minimum réel au moment de l'abandon. Pour les diplômés/persévérants le salaire minimum réel en dollars constants de 1996 utilisé est le salaire moyen que l'individu aurait pu toucher au cours de la période possible d'abandon. Si l'individu étudie ou a étudié dans une région métropolitaine de recensement, le salaire minimum est déflaté par l'indice des prix à la consommation de la RMR appropriée, tandis que si l'individu étudie ou a étudié en dehors d'une région métropolitaine de recensement reconnue, le salaire minimum est déflaté par l'indice des prix à la consommation de la province appropriée.

Taux de chômage : Variable construite à partir des données de Statistique Canada. Cette variable représente le taux de chômage qui prévaut au sein de la RMR ou de la province appropriée lorsque l'individu a abandonné l'école. Pour les diplômés/persévérants le taux de chômage utilisé est le taux de chômage moyen au cours de la période possible d'abandon. Si l'individu étudie ou a étudié dans une région métropolitaine de recensement, le taux de chômage utilisé est celui de la RMR, tandis que si l'individu étudie ou a étudié en dehors d'une région métropolitaine de recensement reconnue, le taux de chômage utilisé est celui de la province appropriée. À noter que le taux de chômage des 15 ans ou plus a été utilisé étant donnée la non-disponibilité du taux de chômage des 15-24 ans au niveau des régions métropolitaines de recensement.

Croisée taux de chômage et région : Voir définition dans la spécification travail.

Croisée taux de chômage et éducation des parents : Voir définition dans la spécification travail.

Croisée taux de chômage et famille avec deux parents : Voir définition dans la spécification travail.

Résidence : Voir définition dans la spécification travail.

Annexe B

Les spécifications économétriques des modèles TNA et NTA

Les contributions individuelles à la fonction de vraisemblance du modèle TNA sont spécifiquement pour les décrocheurs, les persistants et les diplômés, selon leurs catégories vis-à-vis de leurs notes et le travail pendant les études :

$$P(T_i = 0, N_i = 0, A_i = 0) = \int_{-\infty}^{-x_i\beta} \int_{-\infty}^{-F-G} \int \phi_3(u_i, v_i, \eta_i, \rho_{u_i v_i}, \rho_{v_i \eta_i}, \rho_{u_i \eta_i}) d\eta_i dv_i du_i$$

$$P(T_i = 0, N_i = 0, A_i = 1) = \int_{-\infty}^{-x_i\beta} \int_{-\infty}^{-G} \int \phi_3(u_i, v_i, \eta_i, \rho_{u_i v_i}, \rho_{v_i \eta_i}, \rho_{u_i \eta_i}) d\eta_i dv_i du_i$$

$$P(T_i = 0, N_i = 1, A_i = 0) = \int_{-\infty}^{-x_i\beta} \int_{-F}^{\theta_1 - F} \int_{-\infty}^{-G} \phi_3(u_i, v_i, \eta_i, \rho_{u_i v_i}, \rho_{v_i \eta_i}, \rho_{u_i \eta_i}) d\eta_i dv_i du_i$$

$$P(T_i = 0, N_i = 1, A_i = 1) = \int_{-\infty}^{-x_i\beta} \int_{-F}^{\theta_1 - F} \int_{-G}^{\infty} \phi_3(u_i, v_i, \eta_i, \rho_{u_i v_i}, \rho_{v_i \eta_i}, \rho_{u_i \eta_i}) d\eta_i dv_i du_i$$

$$P(T_i = 0, N_i = 2, A_i = 0) = \int_{-\infty}^{-x_i\beta} \int_{\theta_1 - F}^{\infty} \int_{-\infty}^{-G} \phi_3(u_i, v_i, \eta_i, \rho_{u_i v_i}, \rho_{v_i \eta_i}, \rho_{u_i \eta_i}) d\eta_i dv_i du_i$$

$$P(T_i = 0, N_i = 2, A_i = 1) = \int_{-\infty}^{-x_i\beta} \int_{\theta_1 - F}^{\infty} \int_{-G}^{\infty} \phi_3(u_i, v_i, \eta_i, \rho_{u_i v_i}, \rho_{v_i \eta_i}, \rho_{u_i \eta_i}) d\eta_i dv_i du_i$$

$$P(T_i = 1, N_i = 0, A_i = 0) = \int_{-x_i\beta}^{\varepsilon_1 - x_i\beta} \int_{-\infty}^{-F-G} \int \phi_3(u_i, v_i, \eta_i, \rho_{u_i v_i}, \rho_{v_i \eta_i}, \rho_{u_i \eta_i}) d\eta_i dv_i du_i$$

$$P(T_i = 1, N_i = 0, A_i = 1) = \int_{-x_i\beta}^{\varepsilon_1 - x_i\beta} \int_{-\infty}^{-G} \int \phi_3(u_i, v_i, \eta_i, \rho_{u_i v_i}, \rho_{v_i \eta_i}, \rho_{u_i \eta_i}) d\eta_i dv_i du_i$$

$$P(T_i = 1, N_i = 1, A_i = 0) = \int_{-x_i\beta}^{\varepsilon_1 - x_i\beta} \int_{-F}^{\theta_1 - F} \int_{-\infty}^{-G} \phi_3(u_i, v_i, \eta_i, \rho_{u_i v_i}, \rho_{v_i \eta_i}, \rho_{u_i \eta_i}) d\eta_i dv_i du_i$$

$$P(T_i = 1, N_i = 1, A_i = 1) = \int_{-x_i\beta}^{\varepsilon_1 - x_i\beta} \int_{-F}^{\theta_1 - F} \int_{-G}^{\infty} \phi_3(u_i, v_i, \eta_i, \rho_{u_i v_i}, \rho_{v_i \eta_i}, \rho_{u_i \eta_i}) d\eta_i dv_i du_i$$

$$P(T_i = 1, N_i = 2, A_i = 0) = \int_{-x_i\beta}^{\varepsilon_1 - x_i\beta} \int_{\theta_1 - F}^{\infty} \int_{-\infty}^{-G} \phi_3(u_i, v_i, \eta_i, \rho_{u_i v_i}, \rho_{v_i \eta_i}, \rho_{u_i \eta_i}) d\eta_i dv_i du_i$$

$$P(T_i = 1, N_i = 2, A_i = 1) = \int_{-x_i\beta}^{\varepsilon_1 - x_i\beta} \int_{\theta_1 - F}^{\infty} \int_{-G}^{\infty} \phi_3(u_i, v_i, \eta_i, \rho_{u_i v_i}, \rho_{v_i \eta_i}, \rho_{u_i \eta_i}) d\eta_i dv_i du_i$$

$$P(T_i = 2, N_i = 0, A_i = 0) = \int_{\varepsilon_1 - x_i \beta}^{\varepsilon_2 - x_i \beta} \int_{-\infty}^{-F} \int_{-\infty}^{-G} \phi_3(u_i, v_i, \eta_i, \rho_{u_i v_i}, \rho_{v_i \eta_i}, \rho_{u_i \eta_i}) d\eta_i dv_i du_i$$

$$P(T_i = 2, N_i = 0, A_i = 1) = \int_{\varepsilon_1 - x_i \beta}^{\varepsilon_2 - x_i \beta} \int_{-\infty}^{-F} \int_{-G}^{\infty} \phi_3(u_i, v_i, \eta_i, \rho_{u_i v_i}, \rho_{v_i \eta_i}, \rho_{u_i \eta_i}) d\eta_i dv_i du_i$$

$$P(T_i = 2, N_i = 1, A_i = 0) = \int_{\varepsilon_1 - x_i \beta}^{\varepsilon_2 - x_i \beta} \int_{-F}^{-G} \int_{-\infty}^{\infty} \phi_3(u_i, v_i, \eta_i, \rho_{u_i v_i}, \rho_{v_i \eta_i}, \rho_{u_i \eta_i}) d\eta_i dv_i du_i$$

$$P(T_i = 2, N_i = 1, A_i = 1) = \int_{\varepsilon_1 - x_i \beta}^{\varepsilon_2 - x_i \beta} \int_{-F}^{-G} \int_{-G}^{\infty} \phi_3(u_i, v_i, \eta_i, \rho_{u_i v_i}, \rho_{v_i \eta_i}, \rho_{u_i \eta_i}) d\eta_i dv_i du_i$$

$$P(T_i = 2, N_i = 2, A_i = 0) = \int_{\varepsilon_1 - x_i \beta}^{\varepsilon_2 - x_i \beta} \int_{\theta_1 - F}^{-G} \int_{-\infty}^{\infty} \phi_3(u_i, v_i, \eta_i, \rho_{u_i v_i}, \rho_{v_i \eta_i}, \rho_{u_i \eta_i}) d\eta_i dv_i du_i$$

$$P(T_i = 2, N_i = 2, A_i = 1) = \int_{\varepsilon_1 - x_i \beta}^{\varepsilon_2 - x_i \beta} \int_{\theta_1 - F}^{-G} \int_{-G}^{\infty} \phi_3(u_i, v_i, \eta_i, \rho_{u_i v_i}, \rho_{v_i \eta_i}, \rho_{u_i \eta_i}) d\eta_i dv_i du_i$$

$$P(T_i = 3, N_i = 0, A_i = 0) = \int_{\varepsilon_2 - x_i \beta}^{\infty} \int_{-\infty}^{-F} \int_{-\infty}^{-G} \phi_3(u_i, v_i, \eta_i, \rho_{u_i v_i}, \rho_{v_i \eta_i}, \rho_{u_i \eta_i}) d\eta_i dv_i du_i$$

$$P(T_i = 3, N_i = 0, A_i = 1) = \int_{\varepsilon_2 - x_i \beta}^{\infty} \int_{-\infty}^{-F} \int_{-G}^{\infty} \phi_3(u_i, v_i, \eta_i, \rho_{u_i v_i}, \rho_{v_i \eta_i}, \rho_{u_i \eta_i}) d\eta_i dv_i du_i$$

$$P(T_i = 3, N_i = 1, A_i = 0) = \int_{\varepsilon_2 - x_i \beta}^{\infty} \int_{-F}^{-G} \int_{-\infty}^{\infty} \phi_3(u_i, v_i, \eta_i, \rho_{u_i v_i}, \rho_{v_i \eta_i}, \rho_{u_i \eta_i}) d\eta_i dv_i du_i$$

$$P(T_i = 3, N_i = 1, A_i = 1) = \int_{\varepsilon_2 - x_i \beta}^{\infty} \int_{-F}^{-G} \int_{-G}^{\infty} \phi_3(u_i, v_i, \eta_i, \rho_{u_i v_i}, \rho_{v_i \eta_i}, \rho_{u_i \eta_i}) d\eta_i dv_i du_i$$

$$P(T_i = 3, N_i = 2, A_i = 0) = \int_{\varepsilon_2 - x_i \beta}^{\infty} \int_{\theta_1 - F}^{-G} \int_{-\infty}^{\infty} \phi_3(u_i, v_i, \eta_i, \rho_{u_i v_i}, \rho_{v_i \eta_i}, \rho_{u_i \eta_i}) d\eta_i dv_i du_i$$

$$P(T_i = 3, N_i = 2, A_i = 1) = \int_{\varepsilon_2 - x_i \beta}^{\infty} \int_{\theta_1 - F}^{-G} \int_{-G}^{\infty} \phi_3(u_i, v_i, \eta_i, \rho_{u_i v_i}, \rho_{v_i \eta_i}, \rho_{u_i \eta_i}) d\eta_i dv_i du_i.$$

ϕ_3 correspond à la fonction de densité d'une normale standard trivariée et les $\rho_{u_i v_i}, \rho_{v_i \eta_i}, \rho_{u_i \eta_i}$ sont les coefficients de corrélation entre les termes d'erreurs.

Les contributions individuelles à la fonction de vraisemblance du modèle NTA sont spécifiquement pour les décrocheurs, les persistants et les diplômés selon leurs catégories vis-à-vis de leurs notes et le travail pendant les études :

$$P(N_i = 0, T_i = 0, A_i = 0) = \int_{-\infty}^{-y_i\alpha - E} \int_{-\infty}^{-H} \int_{-\infty}^{\infty} f_3(\mu_i, \vartheta_i, \psi_i, \rho_{\mu_i\vartheta_i}, \rho_{\vartheta_i\psi_i}, \rho_{\mu_i\psi_i}) d\psi_i d\vartheta_i d\mu_i.$$

$$P(N_i = 0, T_i = 0, A_i = 1) = \int_{-\infty}^{-y_i\alpha - E} \int_{-\infty}^{-H} \int_{-\infty}^{\infty} f_3(\mu_i, \vartheta_i, \psi_i, \rho_{\mu_i\vartheta_i}, \rho_{\vartheta_i\psi_i}, \rho_{\mu_i\psi_i}) d\psi_i d\vartheta_i d\mu_i$$

$$P(N_i = 0, T_i = 1, A_i = 0) = \int_{-\infty}^{-y_i\alpha} \int_{-E}^{\zeta_1 - E - H} \int_{-\infty}^{\infty} f_3(\mu_i, \vartheta_i, \psi_i, \rho_{\mu_i\vartheta_i}, \rho_{\vartheta_i\psi_i}, \rho_{\mu_i\psi_i}) d\psi_i d\vartheta_i d\mu_i$$

$$P(N_i = 0, T_i = 1, A_i = 1) = \int_{-\infty}^{-y_i\alpha} \int_{-E}^{\zeta_1 - E} \int_{-H}^{\infty} f_3(\mu_i, \vartheta_i, \psi_i, \rho_{\mu_i\vartheta_i}, \rho_{\vartheta_i\psi_i}, \rho_{\mu_i\psi_i}) d\psi_i d\vartheta_i d\mu_i$$

$$P(N_i = 0, T_i = 2, A_i = 0) = \int_{-\infty}^{-y_i\alpha} \int_{\zeta_1 - E}^{\zeta_2 - E - H} \int_{-\infty}^{\infty} f_3(\mu_i, \vartheta_i, \psi_i, \rho_{\mu_i\vartheta_i}, \rho_{\vartheta_i\psi_i}, \rho_{\mu_i\psi_i}) d\psi_i d\vartheta_i d\mu_i$$

$$P(N_i = 0, T_i = 2, A_i = 1) = \int_{-\infty}^{-y_i\alpha} \int_{\zeta_1 - E}^{\zeta_2 - E} \int_{-H}^{\infty} f_3(\mu_i, \vartheta_i, \psi_i, \rho_{\mu_i\vartheta_i}, \rho_{\vartheta_i\psi_i}, \rho_{\mu_i\psi_i}) d\psi_i d\vartheta_i d\mu_i$$

$$P(N_i = 0, T_i = 3, A_i = 0) = \int_{-\infty}^{-y_i\alpha} \int_{-\zeta_2 - E}^{\infty} \int_{-H}^{\infty} f_3(\mu_i, \vartheta_i, \psi_i, \rho_{\mu_i\vartheta_i}, \rho_{\vartheta_i\psi_i}, \rho_{\mu_i\psi_i}) d\psi_i d\vartheta_i d\mu_i$$

$$P(N_i = 0, T_i = 3, A_i = 1) = \int_{-\infty}^{-y_i\alpha} \int_{-\zeta_2 - E}^{\infty} \int_{-H}^{\infty} f_3(\mu_i, \vartheta_i, \psi_i, \rho_{\mu_i\vartheta_i}, \rho_{\vartheta_i\psi_i}, \rho_{\mu_i\psi_i}) d\psi_i d\vartheta_i d\mu_i$$

$$P(N_i = 1, T_i = 0, A_i = 0) = \int_{-y_i\alpha}^{\Omega_1 - y_i\alpha - E - H} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_3(\mu_i, \vartheta_i, \psi_i, \rho_{\mu_i\vartheta_i}, \rho_{\vartheta_i\psi_i}, \rho_{\mu_i\psi_i}) d\psi_i d\vartheta_i d\mu_i$$

$$P(N_i = 1, T_i = 0, A_i = 1) = \int_{-y_i\alpha}^{\Omega_1 - y_i\alpha - E} \int_{-\infty}^{-H} \int_{-\infty}^{\infty} f_3(\mu_i, \vartheta_i, \psi_i, \rho_{\mu_i\vartheta_i}, \rho_{\vartheta_i\psi_i}, \rho_{\mu_i\psi_i}) d\psi_i d\vartheta_i d\mu_i$$

$$P(N_i = 1, T_i = 1, A_i = 0) = \int_{-y_i\alpha}^{\Omega_1 - y_i\alpha} \int_{-E}^{\zeta_1 - E - H} \int_{-\infty}^{\infty} f_3(\mu_i, \vartheta_i, \psi_i, \rho_{\mu_i\vartheta_i}, \rho_{\vartheta_i\psi_i}, \rho_{\mu_i\psi_i}) d\psi_i d\vartheta_i d\mu_i$$

$$P(N_i = 1, T_i = 1, A_i = 1) = \int_{-y_i\alpha}^{\Omega_1 - y_i\alpha} \int_{-E}^{\zeta_1 - E} \int_{-H}^{\infty} f_3(\mu_i, \vartheta_i, \psi_i, \rho_{\mu_i\vartheta_i}, \rho_{\vartheta_i\psi_i}, \rho_{\mu_i\psi_i}) d\psi_i d\vartheta_i d\mu_i$$

$$P(N_i = 1, T_i = 2, A_i = 0) = \int_{-y_i\alpha}^{\Omega_1 - y_i\alpha} \int_{\zeta_1 - E}^{\zeta_2 - E - H} \int_{-\infty}^{\infty} f_3(\mu_i, \vartheta_i, \psi_i, \rho_{\mu_i\vartheta_i}, \rho_{\vartheta_i\psi_i}, \rho_{\mu_i\psi_i}) d\psi_i d\vartheta_i d\mu_i$$

$$\begin{aligned}
P(N_i = 1, T_i = 2, A_i = 1) &= \int_{-y_i\alpha}^{\Omega_1 - y_i\alpha} \int_{\zeta_1 - E}^{\zeta_2 - E} \int_{-\infty}^{\infty} f_3(\mu_i, \vartheta_i, \psi_i, \rho_{\mu_i\vartheta_i}, \rho_{\vartheta_i\psi_i}, \rho_{\mu_i\psi_i}) d\psi_i d\vartheta_i d\mu_i \\
P(N_i = 1, T_i = 3, A_i = 0) &= \int_{-y_i\alpha}^{\Omega_1 - y_i\alpha} \int_{\zeta_2 - E}^{\infty} \int_{-\infty}^{-H} f_3(\mu_i, \vartheta_i, \psi_i, \rho_{\mu_i\vartheta_i}, \rho_{\vartheta_i\psi_i}, \rho_{\mu_i\psi_i}) d\psi_i d\vartheta_i d\mu_i \\
P(N_i = 1, T_i = 3, A_i = 1) &= \int_{-y_i\alpha}^{\Omega_1 - y_i\alpha} \int_{\zeta_2 - E}^{\infty} \int_{-\infty}^{\infty} f_3(\mu_i, \vartheta_i, \psi_i, \rho_{\mu_i\vartheta_i}, \rho_{\vartheta_i\psi_i}, \rho_{\mu_i\psi_i}) d\psi_i d\vartheta_i d\mu_i \\
P(N_i = 2, T_i = 0, A_i = 0) &= \int_{\Omega_1 - y_i\alpha}^{\infty} \int_{-\infty}^{-E-H} \int_{-\infty}^{\infty} f_3(\mu_i, \vartheta_i, \psi_i, \rho_{\mu_i\vartheta_i}, \rho_{\vartheta_i\psi_i}, \rho_{\mu_i\psi_i}) d\psi_i d\vartheta_i d\mu_i \\
P(N_i = 2, T_i = 0, A_i = 1) &= \int_{\Omega_1 - y_i\alpha}^{\infty} \int_{-\infty}^{-E} \int_{-\infty}^{\infty} f_3(\mu_i, \vartheta_i, \psi_i, \rho_{\mu_i\vartheta_i}, \rho_{\vartheta_i\psi_i}, \rho_{\mu_i\psi_i}) d\psi_i d\vartheta_i d\mu_i \\
P(N_i = 2, T_i = 1, A_i = 0) &= \int_{\Omega_1 - y_i\alpha}^{\infty} \int_{-E}^{\zeta_1 - E - H} \int_{-\infty}^{\infty} f_3(\mu_i, \vartheta_i, \psi_i, \rho_{\mu_i\vartheta_i}, \rho_{\vartheta_i\psi_i}, \rho_{\mu_i\psi_i}) d\psi_i d\vartheta_i d\mu_i \\
P(N_i = 2, T_i = 1, A_i = 1) &= \int_{\Omega_1 - y_i\alpha}^{\infty} \int_{-E}^{\zeta_1 - E} \int_{-H}^{\infty} f_3(\mu_i, \vartheta_i, \psi_i, \rho_{\mu_i\vartheta_i}, \rho_{\vartheta_i\psi_i}, \rho_{\mu_i\psi_i}) d\psi_i d\vartheta_i d\mu_i \\
P(N_i = 2, T_i = 2, A_i = 0) &= \int_{\Omega_1 - y_i\alpha}^{\infty} \int_{\zeta_1 - E}^{\zeta_2 - E - H} \int_{-\infty}^{\infty} f_3(\mu_i, \vartheta_i, \psi_i, \rho_{\mu_i\vartheta_i}, \rho_{\vartheta_i\psi_i}, \rho_{\mu_i\psi_i}) d\psi_i d\vartheta_i d\mu_i \\
P(N_i = 2, T_i = 2, A_i = 1) &= \int_{\Omega_1 - y_i\alpha}^{\infty} \int_{\zeta_1 - E}^{\zeta_2 - E} \int_{-H}^{\infty} f_3(\mu_i, \vartheta_i, \psi_i, \rho_{\mu_i\vartheta_i}, \rho_{\vartheta_i\psi_i}, \rho_{\mu_i\psi_i}) d\psi_i d\vartheta_i d\mu_i \\
P(N_i = 2, T_i = 3, A_i = 0) &= \int_{\Omega_1 - y_i\alpha}^{\infty} \int_{\zeta_2 - E}^{\infty} \int_{-\infty}^{-H} f_3(\mu_i, \vartheta_i, \psi_i, \rho_{\mu_i\vartheta_i}, \rho_{\vartheta_i\psi_i}, \rho_{\mu_i\psi_i}) d\psi_i d\vartheta_i d\mu_i \\
P(N_i = 2, T_i = 3, A_i = 1) &= \int_{\Omega_1 - y_i\alpha}^{\infty} \int_{\zeta_2 - E}^{\infty} \int_{-H}^{\infty} f_3(\mu_i, \vartheta_i, \psi_i, \rho_{\mu_i\vartheta_i}, \rho_{\vartheta_i\psi_i}, \rho_{\mu_i\psi_i}) d\psi_i d\vartheta_i d\mu_i.
\end{aligned}$$

f_3 correspond à la fonction de densité d'une normale standard trivariée et les $\rho_{\mu_i\vartheta_i}, \rho_{\vartheta_i\psi_i}, \rho_{\mu_i\psi_i}$ sont les coefficients de corrélation entre les termes d'erreurs.

Le logarithme de la fonction de vraisemblance correspondante au modèle TNA s'écrit:

$$\begin{aligned}
TNA = & \sum_a^{n_1} \int_{-\infty}^{-x_i\beta-F-G} \int_{-\infty}^{-F-G} \int_{-\infty}^{-G} \phi_3(u_i, v_i, \eta_i, \rho_{u_i v_i}, \rho_{v_i \eta_i}, \rho_{u_i \eta_i}) d\eta_i dv_i du_i \\
& + \sum_b^{n_2} \int_{-\infty}^{-x_i\beta-F} \int_{-\infty}^{-G} \int_{-\infty}^{\infty} \phi_3(u_i, v_i, \eta_i, \rho_{u_i v_i}, \rho_{v_i \eta_i}, \rho_{u_i \eta_i}) d\eta_i dv_i du_i \\
& + \sum_c^{n_3} \int_{-\infty}^{-x_i\beta\theta_1-F-G} \int_{-F}^{-G} \int_{-\infty}^{\infty} \phi_3(u_i, v_i, \eta_i, \rho_{u_i v_i}, \rho_{v_i \eta_i}, \rho_{u_i \eta_i}) d\eta_i dv_i du_i \\
& + \sum_d^{n_4} \int_{-\infty}^{-x_i\beta\theta_1-F} \int_{-F}^{-G} \int_{-\infty}^{\infty} \phi_3(u_i, v_i, \eta_i, \rho_{u_i v_i}, \rho_{v_i \eta_i}, \rho_{u_i \eta_i}) d\eta_i dv_i du_i \\
& + \sum_e^{n_5} \int_{-\infty}^{-x_i\beta} \int_{\theta_1-F}^{-G} \int_{-\infty}^{\infty} \phi_3(u_i, v_i, \eta_i, \rho_{u_i v_i}, \rho_{v_i \eta_i}, \rho_{u_i \eta_i}) d\eta_i dv_i du_i \\
& + \sum_f^{n_6} \int_{-\infty}^{-x_i\beta} \int_{\theta_1-F-G}^{\infty} \int_{-\infty}^{\infty} \phi_3(u_i, v_i, \eta_i, \rho_{u_i v_i}, \rho_{v_i \eta_i}, \rho_{u_i \eta_i}) d\eta_i dv_i du_i \\
& + \sum_g^{n_7} \int_{-x_i\beta}^{\varepsilon_1-x_i\beta_1-F-G} \int_{-\infty}^{-G} \int_{-\infty}^{\infty} \phi_3(u_i, v_i, \eta_i, \rho_{u_i v_i}, \rho_{v_i \eta_i}, \rho_{u_i \eta_i}) d\eta_i dv_i du_i \\
& + \sum_h^{n_8} \int_{-x_i\beta}^{\varepsilon_1-x_i\beta_1-F} \int_{-\infty}^{-G} \int_{-\infty}^{\infty} \phi_3(u_i, v_i, \eta_i, \rho_{u_i v_i}, \rho_{v_i \eta_i}, \rho_{u_i \eta_i}) d\eta_i dv_i du_i \\
& + \sum_i^{n_9} \int_{-x_i\beta}^{\varepsilon_1-x_i\beta_1\theta_1-F-G} \int_{-F}^{-G} \int_{-\infty}^{\infty} \phi_3(u_i, v_i, \eta_i, \rho_{u_i v_i}, \rho_{v_i \eta_i}, \rho_{u_i \eta_i}) d\eta_i dv_i du_i \\
& + \sum_j^{n_{10}} \int_{-x_i\beta}^{\varepsilon_1-x_i\beta_1\theta_1-F} \int_{-F}^{-G} \int_{-\infty}^{\infty} \phi_3(u_i, v_i, \eta_i, \rho_{u_i v_i}, \rho_{v_i \eta_i}, \rho_{u_i \eta_i}) d\eta_i dv_i du_i \\
& + \sum_k^{n_{11}} \int_{-x_i\beta}^{\varepsilon_1-x_i\beta_1} \int_{\theta_1-F}^{-G} \int_{-\infty}^{\infty} \phi_3(u_i, v_i, \eta_i, \rho_{u_i v_i}, \rho_{v_i \eta_i}, \rho_{u_i \eta_i}) d\eta_i dv_i du_i \\
& + \sum_l^{n_{12}} \int_{-x_i\beta}^{\varepsilon_1-x_i\beta_1} \int_{\theta_1-F-G}^{\infty} \int_{-\infty}^{\infty} \phi_3(u_i, v_i, \eta_i, \rho_{u_i v_i}, \rho_{v_i \eta_i}, \rho_{u_i \eta_i}) d\eta_i dv_i du_i
\end{aligned}$$

$$\begin{aligned}
& + \sum_m^{n_{13}} \int_{\varepsilon_1 - x_i \beta}^{\varepsilon_2 - x_i \beta - F - G} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \phi_3(u_i, v_i, \eta_i, \rho_{u_i v_i}, \rho_{v_i \eta_i}, \rho_{u_i \eta_i}) d\eta_i dv_i du_i \\
& + \sum_n^{n_{14}} \int_{\varepsilon_1 - x_i \beta}^{\varepsilon_2 - x_i \beta - F} \int_{-\infty}^{\infty} \int_{-G}^{\infty} \phi_3(u_i, v_i, \eta_i, \rho_{u_i v_i}, \rho_{v_i \eta_i}, \rho_{u_i \eta_i}) d\eta_i dv_i du_i \\
& + \sum_o^{n_{15}} \int_{\varepsilon_1 - x_i \beta}^{\varepsilon_2 - x_i \beta \theta_1 - F - G} \int_{-F}^{\infty} \int_{-\infty}^{\infty} \phi_3(u_i, v_i, \eta_i, \rho_{u_i v_i}, \rho_{v_i \eta_i}, \rho_{u_i \eta_i}) d\eta_i dv_i du_i \\
& + \sum_p^{n_{16}} \int_{\varepsilon_1 - x_i \beta}^{\varepsilon_2 - x_i \beta \theta_1 - F} \int_{-F}^{\infty} \int_{-G}^{\infty} \phi_3(u_i, v_i, \eta_i, \rho_{u_i v_i}, \rho_{v_i \eta_i}, \rho_{u_i \eta_i}) d\eta_i dv_i du_i \\
& + \sum_q^{n_{17}} \int_{\varepsilon_1 - x_i \beta}^{\varepsilon_2 - x_i \beta} \int_{\theta_1 - F}^{\infty} \int_{-\infty}^{-G} \phi_3(u_i, v_i, \eta_i, \rho_{u_i v_i}, \rho_{v_i \eta_i}, \rho_{u_i \eta_i}) d\eta_i dv_i du_i \\
& + \sum_r^{n_{18}} \int_{\varepsilon_1 - x_i \beta}^{\varepsilon_2 - x_i \beta} \int_{\theta_1 - F - G}^{\infty} \int_{-\infty}^{\infty} \phi_3(u_i, v_i, \eta_i, \rho_{u_i v_i}, \rho_{v_i \eta_i}, \rho_{u_i \eta_i}) d\eta_i dv_i du_i \\
& + \sum_s^{n_{19}} \int_{\varepsilon_2 - x_i \beta}^{\infty} \int_{-\infty}^{-F - G} \int_{-\infty}^{\infty} \phi_3(u_i, v_i, \eta_i, \rho_{u_i v_i}, \rho_{v_i \eta_i}, \rho_{u_i \eta_i}) d\eta_i dv_i du_i \\
& + \sum_t^{n_{20}} \int_{\varepsilon_2 - x_i \beta}^{\infty} \int_{-\infty}^{-F} \int_{-\infty}^{\infty} \phi_3(u_i, v_i, \eta_i, \rho_{u_i v_i}, \rho_{v_i \eta_i}, \rho_{u_i \eta_i}) d\eta_i dv_i du_i \\
& + \sum_u^{n_{21}} \int_{\varepsilon_2 - x_i \beta}^{\infty} \int_{-F}^{\theta_1 - F - G} \int_{-\infty}^{\infty} \phi_3(u_i, v_i, \eta_i, \rho_{u_i v_i}, \rho_{v_i \eta_i}, \rho_{u_i \eta_i}) d\eta_i dv_i du_i \\
& + \sum_v^{n_{22}} \int_{\varepsilon_2 - x_i \beta}^{\infty} \int_{-F}^{\theta_1 - F} \int_{-G}^{\infty} \phi_3(u_i, v_i, \eta_i, \rho_{u_i v_i}, \rho_{v_i \eta_i}, \rho_{u_i \eta_i}) d\eta_i dv_i du_i \\
& + \sum_w^{n_{23}} \int_{\varepsilon_2 - x_i \beta}^{\infty} \int_{\theta_1 - F}^{\infty} \int_{-\infty}^{-G} \phi_3(u_i, v_i, \eta_i, \rho_{u_i v_i}, \rho_{v_i \eta_i}, \rho_{u_i \eta_i}) d\eta_i dv_i du_i \\
& + \sum_x^{n_{24}} \int_{\varepsilon_2 - x_i \beta}^{\infty} \int_{\theta_1 - F - G}^{\infty} \int_{-\infty}^{\infty} \phi_3(u_i, v_i, \eta_i, \rho_{u_i v_i}, \rho_{v_i \eta_i}, \rho_{u_i \eta_i}) d\eta_i dv_i du_i
\end{aligned}$$

Pour tenir compte de l'hétéroscédasticité, l'hypothèse usuelle de $\sigma = 1$ pour le probit (abandon) et les probits ordonnés (travail et notes) est remplacée par $\sigma_i = \exp(\varphi_1 m_i)$, $\sigma_i^* = \exp(\varphi_2 m_i)$ et $\sigma_i^{**} = \exp(\varphi_3 m_i)$ respectivement, où m_i correspond à l'éducation des parents et les φ sont des paramètres à estimer. Il faut diviser chacune des bornes des intégrales correspondantes au probit et aux probits ordonnés du logarithme de la fonction de vraisemblance, par ces expressions. Si à l'estimation, $\varphi_1 = \varphi_2 = \varphi_3 = 0$, on accepte l'hypothèse de l'homoscédasticité.

Le logarithme de la fonction de vraisemblance correspondante au modèle NTA s'écrit :

$$\begin{aligned}
\text{NTA} = & \sum_a^{n_1} \int_{-\infty}^{-y_i \alpha - E} \int_{-\infty}^{-H} f_3(\mu_i, \vartheta_i, \psi_i, \rho_{\mu_i \vartheta_i}, \rho_{\vartheta_i \psi_i}, \rho_{\mu_i \psi_i}) d\psi_i d\vartheta_i d\mu_i \\
& + \sum_b^{n_2} \int_{-\infty}^{-y_i \alpha - E} \int_{-\infty}^{-H} f_3(\mu_i, \vartheta_i, \psi_i, \rho_{\mu_i \vartheta_i}, \rho_{\vartheta_i \psi_i}, \rho_{\mu_i \psi_i}) d\psi_i d\vartheta_i d\mu_i \\
& + \sum_c^{n_3} \int_{-\infty}^{-y_i \alpha} \int_{-E}^{\zeta_1 - E - H} f_3(\mu_i, \vartheta_i, \psi_i, \rho_{\mu_i \vartheta_i}, \rho_{\vartheta_i \psi_i}, \rho_{\mu_i \psi_i}) d\psi_i d\vartheta_i d\mu_i \\
& + \sum_d^{n_4} \int_{-\infty}^{-y_i \alpha} \int_{-E}^{\zeta_1 - E} \int_{-H}^{\infty} f_3(\mu_i, \vartheta_i, \psi_i, \rho_{\mu_i \vartheta_i}, \rho_{\vartheta_i \psi_i}, \rho_{\mu_i \psi_i}) d\psi_i d\vartheta_i d\mu_i \\
& + \sum_e^{n_5} \int_{-\infty}^{-y_i \alpha} \int_{\zeta_1 - E}^{\zeta_2 - E - H} f_3(\mu_i, \vartheta_i, \psi_i, \rho_{\mu_i \vartheta_i}, \rho_{\vartheta_i \psi_i}, \rho_{\mu_i \psi_i}) d\psi_i d\vartheta_i d\mu_i \\
& + \sum_f^{n_6} \int_{-\infty}^{-y_i \alpha} \int_{\zeta_1 - E}^{\zeta_2 - E} \int_{-H}^{\infty} f_3(\mu_i, \vartheta_i, \psi_i, \rho_{\mu_i \vartheta_i}, \rho_{\vartheta_i \psi_i}, \rho_{\mu_i \psi_i}) d\psi_i d\vartheta_i d\mu_i \\
& + \sum_g^{n_7} \int_{-\infty}^{-y_i \alpha} \int_{-\zeta_2 - E}^{\infty} \int_{-\infty}^{-H} f_3(\mu_i, \vartheta_i, \psi_i, \rho_{\mu_i \vartheta_i}, \rho_{\vartheta_i \psi_i}, \rho_{\mu_i \psi_i}) d\psi_i d\vartheta_i d\mu_i \\
& + \sum_h^{n_8} \int_{-\infty}^{-y_i \alpha} \int_{-\zeta_2 - E}^{\infty} \int_{-H}^{\infty} f_3(\mu_i, \vartheta_i, \psi_i, \rho_{\mu_i \vartheta_i}, \rho_{\vartheta_i \psi_i}, \rho_{\mu_i \psi_i}) d\psi_i d\vartheta_i d\mu_i \\
& + \sum_i^{n_9} \int_{-y_i \alpha}^{\Omega_1 - y_i \alpha - E - H} \int_{-\infty}^{-\infty} f_3(\mu_i, \vartheta_i, \psi_i, \rho_{\mu_i \vartheta_i}, \rho_{\vartheta_i \psi_i}, \rho_{\mu_i \psi_i}) d\psi_i d\vartheta_i d\mu_i \\
& + \sum_j^{n_{10}} \int_{-y_i \alpha}^{\Omega_1 - y_i \alpha - E} \int_{-\infty}^{-H} f_3(\mu_i, \vartheta_i, \psi_i, \rho_{\mu_i \vartheta_i}, \rho_{\vartheta_i \psi_i}, \rho_{\mu_i \psi_i}) d\psi_i d\vartheta_i d\mu_i \\
& + \sum_k^{n_{11}} \int_{-y_i \alpha}^{\Omega_1 - y_i \alpha} \int_{-E}^{\zeta_1 - E - H} f_3(\mu_i, \vartheta_i, \psi_i, \rho_{\mu_i \vartheta_i}, \rho_{\vartheta_i \psi_i}, \rho_{\mu_i \psi_i}) d\psi_i d\vartheta_i d\mu_i \\
& + \sum_l^{n_{12}} \int_{-y_i \alpha}^{\Omega_1 - y_i \alpha} \int_{-E}^{\zeta_1 - E} \int_{-H}^{\infty} f_3(\mu_i, \vartheta_i, \psi_i, \rho_{\mu_i \vartheta_i}, \rho_{\vartheta_i \psi_i}, \rho_{\mu_i \psi_i}) d\psi_i d\vartheta_i d\mu_i
\end{aligned}$$

$$\begin{aligned}
& + \sum_m^{n_{13}} \int_{-y_i\alpha}^{\Omega_1-y_i\alpha} \int_{\zeta_1-E}^{\zeta_2-E-H} \int_{-\infty}^{\infty} f_3(\mu_i, \vartheta_i, \psi_i, \rho_{\mu_i\vartheta_i}, \rho_{\vartheta_i\psi_i}, \rho_{\mu_i\psi_i}) d\psi_i d\vartheta_i d\mu_i \\
& + \sum_n^{n_{14}} \int_{-y_i\alpha}^{\Omega_1-y_i\alpha} \int_{\zeta_1-E}^{\zeta_2-E} \int_{-\infty}^{\infty} f_3(\mu_i, \vartheta_i, \psi_i, \rho_{\mu_i\vartheta_i}, \rho_{\vartheta_i\psi_i}, \rho_{\mu_i\psi_i}) d\psi_i d\vartheta_i d\mu_i \\
& + \sum_o^{n_{15}} \int_{-y_i\alpha}^{\Omega_1-y_i\alpha} \int_{\zeta_2-E}^{\infty} \int_{-\infty}^{-H} f_3(\mu_i, \vartheta_i, \psi_i, \rho_{\mu_i\vartheta_i}, \rho_{\vartheta_i\psi_i}, \rho_{\mu_i\psi_i}) d\psi_i d\vartheta_i d\mu_i \\
& + \sum_p^{n_{16}} \int_{-y_i\alpha}^{\Omega_1-y_i\alpha} \int_{\zeta_2-E}^{\infty} \int_{-\infty}^{\infty} f_3(\mu_i, \vartheta_i, \psi_i, \rho_{\mu_i\vartheta_i}, \rho_{\vartheta_i\psi_i}, \rho_{\mu_i\psi_i}) d\psi_i d\vartheta_i d\mu_i \\
& + \sum_q^{n_{17}} \int_{\Omega_1-y_i\alpha}^{\infty} \int_{-\infty}^{-E-H} \int_{-\infty}^{\infty} f_3(\mu_i, \vartheta_i, \psi_i, \rho_{\mu_i\vartheta_i}, \rho_{\vartheta_i\psi_i}, \rho_{\mu_i\psi_i}) d\psi_i d\vartheta_i d\mu_i \\
& + \sum_r^{n_{18}} \int_{\Omega_1-y_i\alpha}^{\infty} \int_{-\infty}^{-E} \int_{-\infty}^{\infty} f_3(\mu_i, \vartheta_i, \psi_i, \rho_{\mu_i\vartheta_i}, \rho_{\vartheta_i\psi_i}, \rho_{\mu_i\psi_i}) d\psi_i d\vartheta_i d\mu_i \\
& + \sum_s^{n_{19}} \int_{\Omega_1-y_i\alpha}^{\infty} \int_{-E}^{\zeta_1-E-H} \int_{-\infty}^{\infty} f_3(\mu_i, \vartheta_i, \psi_i, \rho_{\mu_i\vartheta_i}, \rho_{\vartheta_i\psi_i}, \rho_{\mu_i\psi_i}) d\psi_i d\vartheta_i d\mu_i \\
& + \sum_t^{n_{20}} \int_{\Omega_1-y_i\alpha}^{\infty} \int_{-E}^{\zeta_1-E} \int_{-H}^{\infty} f_3(\mu_i, \vartheta_i, \psi_i, \rho_{\mu_i\vartheta_i}, \rho_{\vartheta_i\psi_i}, \rho_{\mu_i\psi_i}) d\psi_i d\vartheta_i d\mu_i \\
& + \sum_u^{n_{21}} \int_{\Omega_1-y_i\alpha}^{\infty} \int_{\zeta_1-E}^{\zeta_2-E-H} \int_{-\infty}^{\infty} f_3(\mu_i, \vartheta_i, \psi_i, \rho_{\mu_i\vartheta_i}, \rho_{\vartheta_i\psi_i}, \rho_{\mu_i\psi_i}) d\psi_i d\vartheta_i d\mu_i \\
& + \sum_v^{n_{22}} \int_{\Omega_1-y_i\alpha}^{\infty} \int_{\zeta_1-E}^{\zeta_2-E} \int_{-\infty}^{\infty} f_3(\mu_i, \vartheta_i, \psi_i, \rho_{\mu_i\vartheta_i}, \rho_{\vartheta_i\psi_i}, \rho_{\mu_i\psi_i}) d\psi_i d\vartheta_i d\mu_i \\
& + \sum_w^{n_{23}} \int_{\Omega_1-y_i\alpha}^{\infty} \int_{\zeta_2-E}^{\infty} \int_{-\infty}^{-H} f_3(\mu_i, \vartheta_i, \psi_i, \rho_{\mu_i\vartheta_i}, \rho_{\vartheta_i\psi_i}, \rho_{\mu_i\psi_i}) d\psi_i d\vartheta_i d\mu_i \\
& + \sum_x^{n_{24}} \int_{\Omega_1-y_i\alpha}^{\infty} \int_{\zeta_2-E}^{\infty} \int_{-\infty}^{\infty} f_3(\mu_i, \vartheta_i, \psi_i, \rho_{\mu_i\vartheta_i}, \rho_{\vartheta_i\psi_i}, \rho_{\mu_i\psi_i}) d\psi_i d\vartheta_i d\mu_i.
\end{aligned}$$

Le logarithme de la fonction de vraisemblance du modèle combiné s'écrit :

$$\log \left[\int_{-\infty}^{-s_i\tilde{\lambda}} f(m_i)TNA.dm_i + \int_{-s_i\tilde{\lambda}}^{\infty} f(m_i)NTA.dm_i \right]$$

Bibliographie

- Becker, G. S. (1964) : *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education*, Columbia University Press.
- Card, D. et Alan B. Krueger (1995) : *Myth and Measurement: The New Economics of the Minimum Wage*, Princeton : Princeton University Press.
- Dagenais, M., C. Montmarquette, D. Parent et N. Viennot-Briot (1999) : *Travail pendant les études, performance scolaire et abandon* , Série Scientifique n° 99s-41, CIRANO.
- Maddala, G. (1983) : *Limited Dependent and Qualitative Variables in Econometrics*, New-York : Cambridge University Press.
- Statistique Canada (1992) : *Enquête auprès des sortants – Guide à l'intention des utilisateurs de microdonnées*, Division de l'éducation de la Culture et du Tourisme.
- Statistique Canada (1995) : *Enquête du suivi auprès des sortants 1995*.
- Tinto, V. (1993) : *Leaving CEGEP: Rethinking the Causes and Cures of Student Attrition*, Chicago : University of Chicago Press.
- Turner, Mark D. (1994) : *The Effects of Part-Time Work on High School Students' Academic Achievement*, Mimeo, University of Maryland.

Liste des publications au CIRANO *

Cahiers CIRANO / *CIRANO Papers* (ISSN 1198-8169)

- 99c-1 Les Expos, l'OSM, les universités, les hôpitaux : Le coût d'un déficit de 400 000 emplois au Québec — Expos, Montréal Symphony Orchestra, Universities, Hospitals: The Cost of a 400,000-Job Shortfall in Québec / Marcel Boyer
- 96c-1 Peut-on créer des emplois en réglementant le temps de travail? / Robert Lacroix
- 95c-2 Anomalies de marché et sélection des titres au Canada / Richard Guay, Jean-François L'Her et Jean-Marc Suret
- 95c-1 La réglementation incitative / Marcel Boyer
- 94c-3 L'importance relative des gouvernements : causes, conséquences et organisations alternative / Claude Montmarquette
- 94c-2 Commercial Bankruptcy and Financial Reorganization in Canada / Jocelyn Martel
- 94c-1 Faire ou faire faire : La perspective de l'économie des organisations / Michel Patry

Série Scientifique / *Scientific Series* (ISSN 1198-8177)

- 2000s-54 Vertical R&D Spillovers, Cooperation, Market Structure, and Innovation / Gamal Atallah
- 2000s-53 Information Sharing and the Stability of Cooperation in Research Joint Ventures / Gamal Atallah
- 2000s-52 A Theory of Routines as Mindsavers / Bernard Sinclair-Desgagné et Antoine Soubeyran
- 2000s-51 Can Financial Intermediation Induce Economic Fluctuations? / Sanjay Banerji et Ngo Van Long
- 2000s-50 Information Technology Sophistication in Hospitals: A Field Study in Quebec / Guy Paré et Claude Sicotte
- 2000s-49 Does Comparable Worth Work in a Decentralized Labor Market? / Michael Baker et Nicole M. Fortin
- 2000s-48 Occupational Gender Composition and Wages in Canada: 1987-1988 / Michael Baker et Nicole M. Fortin
- 2000s-47 Économétrie, théorie des tests et philosophie des sciences / Jean-Marie Dufour
- 2000s-46 Factor Analysis and Independent Component Analysis in Presence of High Idiosyncratic Risks / Thierry Vessereau
- 2000s-45 Aspects non linéaires du marché des actions français / Thierry Vessereau
- 2000s-44 Étude du modèle d'évaluation par arbitrage sur le marché des actions suisses / Thierry Vessereau
- 2000s-43 Habit Formation with Recursive Preferences / Aylin Seckin

* Vous pouvez consulter la liste complète des publications du CIRANO et les publications elles-mêmes sur notre site Internet à l'adresse suivante :