

Université de Montréal

Études en économie du travail

par

Daniel Parent

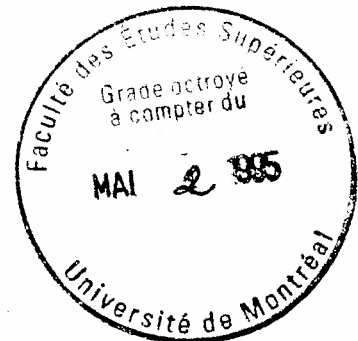
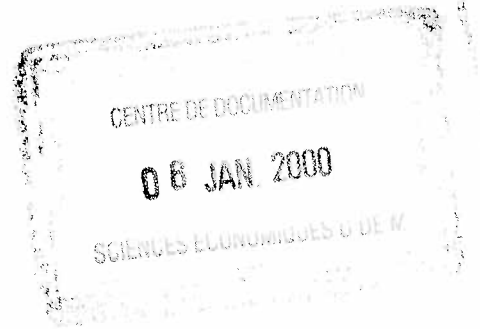
Département de sciences économiques

Faculté des arts et des sciences

**Thèse présentée à la faculté des études supérieures
en vue de l'obtention du grade de
Philosophiae Doctor (Ph.D.)
en sciences économiques**

Décembre 1994

©Daniel Parent





Université de Montréal

Faculté des études supérieures

Cette thèse intitulée:

Études en économie du travail

présentée par

Daniel Parent

a été évaluée par un jury composé des personnes suivantes:

Président du jury	Marcel Dagenais
Directeur de recherche	Bentley MacLeod
Codirecteur	Thomas Lemieux
Membre du jury	David Margolis
Examineur externe	Joseph Altonji

Thèse acceptée le: 10 avril 1995

Sommaire

Dans cette étude, je passe en revue de façon exhaustive les développements majeurs en commençant, comme il se doit, par la contribution fondamentale de Becker (JPE '62 et Woytinski Lecture '67). Après l'examen du modèle de Ben-Porath (JPE '67), qui s'avéra être le point de référence d'un pan entier de la littérature portant sur les choix de consommation et d'investissement dans une perspective de cycle vital, je porte ensuite mon attention sur les travaux de Mincer (JPE '58, '62 et NBER '74) qui, plus que quiconque, contribua à l'application pratique des concepts théoriques. Tout naturellement, l'accent est mis sur le développement de «sa» fonction de gains quoique nous soulignons aussi la compréhension précoce par Mincer des problèmes d'hétérogénéité individuelle. Nous enchaînons ensuite avec les développements plus récents tant du point de vue théorique (matching, modèle de paiements retardés de Lazear), que du point de vue empirique. À cet égard, nous examinons à la fois les contributions ayant fait usage des mesures directes d'accumulation du capital humain (e.g. nombre de semaines passées en formation) ainsi que celles n'ayant à leur disposition que les mesures indirectes habituelles (expérience, ancienneté).

En utilisant des données américaines du National Longitudinal Survey of Youth (NLSY), la deuxième étude s'attarde à examiner l'impact de la formation dispensée par l'employeur sur le profil salarial ainsi que sur la mobilité des jeunes travailleurs faisant leur entrée sur le marché du travail. En exploitant l'aspect longitudinal de l'échantillon de façon à tenir compte de l'hétérogénéité non observée, les résultats montrent (i) un impact économiquement et statistiquement significatif de la formation sur le salaire dans l'emploi courant, (ii) un impact substantiel sur le salaire de la formation acquise avec les employeurs précédents, (iii) une réduction d'environ 18% du salaire de départ pour les travailleurs en formation, et (iv) par un modèle de durée qui tient compte des épisodes multiples d'emploi (permettant alors l'utilisation de méthodes de type «effets fixes»), un degré substantiel de

spécificité du capital humain acquis par le biais de programmes de formation dispensés par l'employeur. La conclusion à tirer de ces résultats est que le capital humain acquis contient à la fois une composante générale rémunérée également par tous les employeurs ainsi qu'une composante spécifique qui réduit la mobilité tout en n'étant pas rémunérée.

Avec les données du NLSY ainsi que celles du «Panel Study of Income Dynamics» (PSID), la troisième étude cherche à déterminer s'il y a un rendement positif net lié à l'ancienneté dans la firme. Topel (JPE 1991) a montré avec un échantillon du PSID l'existence d'un rendement substantiel (25% en 10 ans). Toutefois, du moment que l'on inclut l'expérience dans l'industrie courante dans l'équation de salaire (en plus de l'ancienneté dans la firme ainsi que l'expérience totale de travail), l'effet d'ancienneté disparaît presque complètement, que l'on estime par simples moindres carrés généralisés ou par la méthode des variables instrumentales (IV-GLS), et ce, avec les deux échantillons différents. À noter également que ce résultat est robuste au degré d'agrégation des classes d'industries.

Par le biais de la méthode des moments (réf.: Chamberlain (1984), Gallant et Jorgenson (1979)) et avec des données non balancées du NLSY, le dernier volet de la thèse cherche à tester les prédictions engendrées par la théorie du matching ainsi que la théorie du capital humain quant à la structure de covariance des résidus d'une équation de salaire typique à la Mincer. Le processus de sélection qu'implique le matching dans sa version «pure search good» fait en sorte que l'on devrait observer une diminution de la contribution du terme reflétant la variance dans la qualité du match lorsque l'on suit les travailleurs à mesure qu'ils acquièrent de l'ancienneté dans leur emploi. Les résultats corroborent cette prédiction surtout pour les travailleurs plus scolarisés. Par contre, dans la version «pure experience good» de la théorie, la prédiction à l'effet que la variance devrait s'accroître au tout début de la relation d'emploi s'avère non validée par les résultats, sauf peut-être pour les travailleurs moins scolarisés. Par ailleurs, la théorie du capital humain prédit que l'on devrait observer une corrélation négative entre la pente (le rendement lié à l'ancienneté) et l'ordonnée (ou salaire) à l'origine dans un emploi puisque les travailleurs sont supposés payer pour la

formation. Ce résultat aussi est corroboré, surtout pour les travailleurs ayant au moins un diplôme d'école secondaire. Il importe de noter que la prédiction de la théorie du capital humain quant à la corrélation négative entre la pente et l'ordonnée à l'origine dans un emploi est exclusive à cette théorie. Hause ('80) avait déjà examiné une variante de cette prédiction, soit qu'il devrait y avoir une corrélation négative entre le rendement lié à l'expérience totale et l'ordonnée à l'origine pour un même individu. Or, cette prédiction peut également être engendrée par la théorie du matching. En conséquence, il s'avère nécessaire de généraliser cette prédiction au profil ancienneté-salaire afin de pouvoir discriminer l'apport distinct de la théorie du capital humain.

TABLE DES MATIÈRES

Sommaire	i
Table des matières	iv
Liste des tableaux	vii
Liste des figures	ix
Remerciements	x
Introduction générale	1
Première étude	
Survol des contributions théoriques et empiriques liées au capital humain	
1. Introduction	6
2. Théorie du capital humain	
2.1) Préambule	7
2.2) Capital humain général et capital humain spécifique	8
2.3) Technologie de production du capital humain et profil d'investissement dans le temps	13
2.4) L'équation de salaire de Mincer	23
3. Premiers résultats empiriques	
3.1) Remarques préliminaires	32
3.2) Résultats de Mincer	32
3.3) Capital spécifique à la firme	35
3.4) Théorie du matching en bref	37
4. Développements récents de la théorie du capital humain	39
5. Résultats récents	
5.1) Remarques préliminaires	42
5.2) Le modèle de base et les résultats d'estimation	42
6. Mesures directes de l'acquisition de capital humain	
6.1) Remarques préliminaires	45
7. Conclusion	52
Tableaux	
Figures	

Deuxième étude**Wages and Mobility: the Impact of Employer-Provided Training**

I. Introduction	71
II. Theoretical Framework	73
III. The Data	76
IV. Results	
IV.1 Earnings Equation Estimates	80
IV.2 Specification of the Hazard Model	84
IV.3 The Impact of Training on the Starting Wage	86
V. Conclusion	88
Tableaux	
Figures	

Troisième étude**Industry-Specific Capital and the Wage Profile: Evidence from the NLSY and the PSID**

I. Introduction	102
II. The Data	103
III. Results	
III.1 Basic Model	104
III.2 Earnings Equation Estimates	106
III.3 Comparison with Data from the PSID	108
IV Conclusion	109
Tableaux	

Quatrième étude**Matching, Human Capital, and the Covariance Structure of Earnings**

I. Introduction	120
II. The Theory of Matching and its Empirical Implications	123
II.1 The Theory in a Nutshell	124
III. The Theory of Human Capital and the Covariance Structure of Earnings	134
IV. Empirical Implementation	
IV.1 The Data	136
IV.2 Log-Earnings Equation	137
IV.3 Econometric Models of Matching	138

IV.4 Econometric Model of Human Capital	139
IV.5 Estimation Methodology	142
V. Results	144
V.1 Matching Models	144
V.2 Human Capital Models	146
V.3 Matching with Control for Human Capital Accumulation	148
VI. Conclusion	148
Tableaux	148
Figures	
Appendix	164
Conclusion générale	169
Bibliographie générale	171

Liste des tableaux

Première étude

1. Premiers résultats de Mincer	54
2. Résultats de Mincer et Jovanovic (1981)	54
3. Résultats comparés d'Altonji et Shakotko (1987), Abraham et Farber (1987) et Topel (1991)	55
4. Répertoire des banques de données contenant des mesures explicites du capital humain	56
5. Effet de la formation sur le salaire et la productivité	59
6. Effet de la formation préalable à l'emploi sur le salaire et la productivité	61
7. Effet de la formation sur le salaire de départ	63

Deuxième étude

1. Mean Sample Characteristics (Weighted)-NLSY	90
2. Number of Job-Worker Observations with Training by Industrial Aggregates	91
3. Summary Statistics-Training (Completed) vs No Training	92
4. Training Programs Undertaken between 1988 and 1991	93
5. Earnings Functions Estimates: The Impact of Training	94
6. Hazard Model Results	95
7. OLS Regression Estimates of the Effect of Ongoing Training on Starting Wages	96

Troisième étude

1. Mean Sample Characteristics (Weighted)-NLSY	110
2. Mean Sample Characteristics-PSID	110
3. Earnings Functions Estimates-NLSY: Industry vs Tenure Est	111
4. Earnings Functions Estimates-NLSY: Industry vs Tenure Est	112
5. Industry Effect (1 digit) vs Tenure Effect by Occupations NLSY	113
3. Earnings Functions Estimates-PSID: Industry vs Tenure Est	115
4. Earnings Functions Estimates-PSID: Industry vs Tenure Est	116
5. Industry Effect (1 digit) vs Tenure Effect by Occupations PSID	117

Quatrième étude

1. Mean Sample Characteristics (Weighted)-NLSY	150
2. Estimation of Covariance Structure Implied by the Theory of Matching	151
3. Evolution of the Variance of the Job-Match Component from Job 1 to Job 5 and Beyond	152
4. Evolution of the Variance of the Job-Match Component Within Jobs	153
5. Covariance Structure of Residuals as Function of Tenure, Experience and Unobserved Components (Entire Sample)	154
6. Evolution of Parameters of Quadratic Functions from Job to Job	155
7. Covariance Structure of Residuals as Function of Tenure, Experience and Unobserved Components (by Education Level)	156
8. Evolution of Parameters of Quadratic Functions from Job to Job (by Education Level)	157
9. Evolution of the Variance of the Job-Match Component from Job 1 to Job 5 and Beyond With Quadratic Function in Experience	158
10. Evolution of the Variance of the Job-Match Component Within Jobs with Quadratic Function in Experience	159

Liste des figures**Première étude**

1. Capital humain général	64
2. Capital humain spécifique à la firme	65
3. Offre et demande de capital humain	66
4. Stock d'équilibre de capital humain	67
5. Profil d'investissement et de revenu	68
6. Point de renversement	69

Deuxième étude

1. Timing of First Training Program (OJT)	97
2. Timing of First Training Program (OFT)	98
3. Timing of First Training Program (APP)	99
4. Tenure-Training Relationship	100

Quatrième étude

1. Within-Job Evolution of Variance (Total Sample)	160
2. Within-Job Evolution of Variance (Workers with Less than a High School Diploma)	161
3. Within-Job Evolution of Variance (Workers with a High School Diploma)	162
4. Within-Job Evolution of Variance (Workers with Some College Education)	163

REMERCIEMENTS

La rédaction de cette thèse aurait sans doute été encore plus ardue n'eût été de l'appui et de la collaboration de plusieurs personnes. Comme il se doit, je tiens à remercier en premier lieu mes deux superviseurs, Thomas Lemieux et Bentley MacLeod, qui ont toujours été d'une disponibilité et d'une patience exemplaires. Il me paraît évident qu'un des événements majeurs ayant contribué au développement de cette thèse fut l'arrivée de Thomas au département en 1992. Par sa (déjà) vaste expérience et sa non moins vaste compétence en économie du travail, il a su orienter avec doigté le néophyte que j'étais dans le domaine empirique. Je ne me souviens pas d'une seule fois où je n'ai pas appris quelque chose de nouveau lorsque j'allais l'«achaler» dans son bureau. Pour un chercheur et un pédagogue, je ne puis trouver de meilleur compliment. Quant à Bentley, il importe de souligner qu'il est à l'origine de cette thèse par sa suggestion du capital humain comme thème central. Aussi, tout comme dans le cas de Thomas, la porte de son bureau m'a toujours été ouverte (même lorsqu'il bûchait sur un de ses programmes en C++, ce qui n'est pas peu dire). Son énergie et son enthousiasme tout à fait contagieux se sont avérés une grande source d'inspiration tout au long des quatre dernières années.

En second lieu, je remercie Marcel Dagenais pour son apport essentiel à ma formation en économétrie. J'ai eu la bonne idée, voilà deux ans, de demander à suivre un cours de lecture dirigée en économétrie avec Monsieur Dagenais. Rarement un cour privé dispensé par une université publique n'aura été aussi formateur! De façon générale, je remercie également tout le département de sciences économiques qui m'a accueilli dans son programme de doctorat en 1990. Nul doute que la formation acquise durant ma scolarité s'est avérée précieuse pour l'élaboration de mon projet de recherche. De même, l'ambiance agréable créée par les étudiants que j'ai eu la chance de côtoyer a contribué à ce que les moments plus difficiles soient plus aisément supportables.

D'un point de vue plus personnel, je tiens à souligner la patience indéfectible de ma

compagne, Nathalie. Bien que je suis persuadé qu'elle est bien aise de voir l'aboutissement de ce travail de moine, j'ai rarement (je mentirais si je disais jamais!) senti autre chose qu'un appui aussi discret que ferme. Enfin, je me dois de remercier mon oncle Pascal Parent qui, au moment où j'en avais le plus besoin, m'a prodigué conseils et encouragements lorsque j'ai décidé d'effectuer un retour aux études. Sans lui, cette thèse n'aurait tout simplement pas vu le jour.

Introduction générale

S'il est un domaine où la cohabitation entre la théorie et la pratique s'est avérée fructueuse, c'est bien en économie du travail. Ce constat est aussi valable pour le domaine plus restreint touchant au capital humain. Dans la thèse qui suit, nous nous pencherons de nouveau sur certaines prédictions de la théorie du capital humain, plus particulièrement en ce qui concerne l'impact d'acquérir du capital humain, à la fois sur le salaire et sur la mobilité des travailleurs. En même temps, nous chercherons à déterminer dans quelle mesure le capital humain acquis par les travailleurs semble être de nature spécifique à la firme. Puisque la première étude a pour but explicite de faire le survol de la littérature, cette introduction générale portera surtout sur les questions principales abordées dans les trois autres études de la thèse.

Dans la deuxième étude, en utilisant des données américaines du "National Longitudinal Survey of Youth" (NLSY), nous examinons l'impact des programmes de formation dispensés par les employeurs sur le profil salarial ainsi que sur la mobilité des jeunes travailleurs faisant leur entrée sur le marché du travail. De façon plus spécifique, les questions que nous examinons sont les suivantes: a) dans quelle mesure les travailleurs formés bénéficient-ils en termes de salaire? b) les habiletés acquises par le biais des programmes formels de formation sont-elles transférables d'un employeur à un autre? c) si oui, les travailleurs paient-ils pour la formation acquise par le biais de salaires de départ plus faibles (les données montrent que les employeurs défraient la grande majorité des coûts directs)? d) quel est le degré de spécificité par rapport à la firme du capital humain acquis?

Naturellement, le processus par lequel les travailleurs sont choisis pour suivre des programmes de formation par leur employeurs respectifs est vraisemblablement endogène, tout comme le processus par lequel les travailleurs et les firmes entament les relations d'emploi. Par conséquent, estimer par moindres carrés ordinaires une équation qui fait dépendre le salaire du nombre de semaines de formation complétée conduira à des estimés biaisés des coefficients. De façon à contourner ce problème, nous exploitons l'aspect

longitudinal de la banque de données en construisant un estimateur par variables instrumentales selon la méthodologie proposée par Altonji et Shakotko (1987). Toutes les variables indépendantes (endogènes) spécifiques à la relation d'emploi, soit la formation accumulée dans l'emploi courant ainsi que l'ancienneté dans la firme, sont instrumentées en prenant les déviations par rapport à la moyenne durant la relation. Quant aux variables (endogènes) spécifiques à l'individu, elles sont instrumentées en prenant les déviations par rapport à la moyenne individuelle pour toute la période où le travailleur est présent dans l'échantillon. Ces instruments sont, par construction, non corrélés avec les composantes d'erreur non observées, ce qui n'est pas le cas des variables originales. De plus, puisque l'on suit les mêmes individus dans le temps, les résidus de régression seront corrélés dans le temps. Pour apporter un correctif à ce problème, nous utiliserons alors la méthode des moindres carrés généralisés.

Cette méthode d'estimation est également utilisée dans la troisième étude où l'on cherche à déterminer dans quelle mesure l'effet positif de l'ancienneté sur le salaire est dû à une erreur de spécification plutôt qu'au partage du rendement découlant des investissements spécifiques à la firme. Faute d'avoir une mesure directe du capital humain acquis (ce qui n'est pas le cas pour les données du NLSY utilisées dans la deuxième étude), les économistes ont de façon routinière utilisé l'ancienneté dans la firme comme variable représentant l'accumulation de capital spécifique à la firme alors que l'expérience totale du travailleur sur le marché était vue comme représentant l'accumulation de capital général. Bien que le débat ne soit pas clos, le dernier article majeur (et publié) ayant analysé cette question, soit celui de Topel (1991), faisait état d'un rendement substantiel lié à l'ancienneté. Selon Topel, la prime d'ancienneté est de l'ordre de 25% pour quelqu'un qui passe dix ans dans la même firme. Assez curieusement, aucune attention n'a été portée quant à la pertinence de décomposer la carrière d'un individu en deux morceaux, soit l'ancienneté et l'expérience totale. S'il s'avérait que le capital acquis par le travailleur soit de nature spécifique à l'industrie plutôt qu'à la firme, alors les résultats de Topel seraient erronés. Pour attaquer ce problème, nous ajoutons une autre variable dans l'équation de salaire, soit l'ancienneté dans l'industrie. Cette variable

est facile à construire puisque nous connaissons le code de classification industrielle du travailleur dans chacun de ses emplois. Nous pouvons donc déterminer s'il change de groupe industriel ou non lorsqu'il change d'emploi. Tout comme dans la deuxième étude, les problèmes d'endogénéité des régresseurs sont traités en faisant usage de la méthode des variables instrumentales telle qu'utilisée par Altonji et Shakotko (1987). Quant aux données, elles proviennent du NLSY (1979-1991) ainsi que du "Panel Study of Income Dynamics" (PSID, 1981-1987). En somme, alors que dans la deuxième étude nous disposons de mesures directes d'acquisition du capital humain, soit le nombre total de semaines passées en formation, dans l'étude numéro 3, nous faisons usage des mesures usuelles indirectes.

La dernière étude de cette thèse s'attaque au problème de tenter de discerner les prédictions de la théorie du capital humain de celles de la théorie du matching. Pour ce faire, nous nous concentrons sur les moments d'ordre deux (variance et covariance) des résidus provenant d'une régression standard liant le log du salaire à un vecteur de variables indépendantes, soit le genre de régression effectuée dans les études précédentes. Comme les deux explications concurrentes que constituent le matching et le capital humain donnent le même genre de prédiction quant aux premiers moments (l'espérance du salaire conditionnellement à un vecteur de variables explicatives), il nous faut examiner les prédictions de ces deux théories en ce qui concerne la structure de covariance des résidus si l'on espère pouvoir apporter des réponses plus satisfaisantes. À cette fin, nous utiliserons encore une fois les données du NLSY.

La première partie de l'étude s'attarde à dériver les prédictions de la théorie du matching quant au comportement de la variance des salaires. Ensuite, nous faisons de même pour la théorie du capital humain, généralisant en cela la contribution de Hause (1980). La méthodologie économétrique permettant de tester les différentes théories est ensuite décrite avant de passer aux résultats proprement dits. Il importe de noter que l'un des aspects méthodologiques de notre démarche, soit l'utilisation d'un modèle à coefficients aléatoires, s'inscrit en lignée directe avec la contribution essentielle de Mincer (1974).

Donc, on pourrait résumer la démarche de toute la thèse en affirmant que les deuxième et troisième études cherchent à analyser de façon standard tous les aspects systématiques de la problématique d'acquisition du capital humain alors que la dernière étude s'attaque au côté stochastique et non observable. De plus, au passage, nous analysons les implications d'une théorie concurrente à la théorie du capital humain afin de dégager de façon aussi claire que possible son apport distinct au processus de détermination des salaires.

HS 10

Première étude

**Survol des contributions théoriques et empiriques liées au capital
humain**

HS 1A

1. INTRODUCTION

Depuis la contribution originale de Mincer (1974), les travaux portant sur l'évaluation empirique de la théorie du capital humain par l'entremise d'une fonction liant le logarithme du revenu ou du salaire réel à l'accumulation de capital humain¹ se sont accumulés à un rythme régulier. Bien que les conclusions qu'on puisse tirer varient d'une étude à l'autre, cette divergence d'opinion découle invariablement d'une différence d'approche quant aux problèmes économétriques liés à ce genre d'exercice plutôt qu'à une remise en question de l'exercice lui-même. Ce constat ne vise qu'à illustrer la pertinence de la démarche mise de l'avant par Mincer ainsi que sa robustesse face aux développements théoriques des vingt dernières années en ce qui a trait à la relation d'emploi.

Cet article se veut tout d'abord une synthèse de ces travaux, mais pour que cette synthèse soit davantage qu'une simple énumération de résultats, il faudra bien sûr lui tisser une toile de fond de nature théorique afin qu'apparaisse clairement la cohérence entre les développements empiriques et les fondements théoriques sur lesquels ils reposent. Dans le même ordre d'idée, la nature des résultats étant pour une part non négligeable tributaire des méthodes économétriques employées, il faudra également tenir compte de cet aspect de la problématique. Toutefois, le but de ce survol étant de couvrir les contributions qui nous semblent essentielles, nous ne pourrons faire le tour de tout ce qui s'est publié sur le sujet depuis plus de trois décennies.

Le plan de l'exposé est le suivant. Après avoir exposé les premiers développements théoriques dans la première partie de la section 2, nous terminerons cette section en montrant la démarche suivie par Mincer pour en arriver à une formulation empirique du modèle. La section 3, quant à elle, sera consacrée aux premiers résultats empiriques. Les développements théoriques récents seront abordés dans la section 4, après quoi nous passerons aux articles

¹ Faute d'une traduction adéquate pour ce que Mincer appelle la «human capital earnings function», le vocable équation de salaire sera dorénavant utilisé pour désigner de manière générale cette fonction.

empiriques ayant marqué le domaine aux cours des dernières années. Plus spécifiquement, la section 5 traitera des articles faisant usage des mesures indirectes du capital humain couramment employées, soit l'expérience totale sur le marché du travail et l'ancienneté dans la firme, alors que la section 6 fera l'inventaire des articles utilisant des mesures directes quant à la formation acquise. Enfin, nous concluons avec la section 7.

2. THÉORIE DU CAPITAL HUMAIN.

2.1) Préambule.

Cette section s'attardera surtout à décrire dans ses grandes lignes la théorie du capital humain telle que développée par Becker dans son article original de 1962 et reprise dans son livre (1975). La contribution de Hashimoto (1981) sera également examinée puisqu'elle est étroitement liée à la théorie de Becker en ce sens que Hashimoto a formalisé de façon rigoureuse le concept de partage des coûts et des rendements découlant de l'investissement en capital humain effectué par la firme et/ou le travailleur. Par ailleurs, à partir de Ben-Porath (1967), un courant connexe de la littérature a cherché à pousser encore davantage l'idée que l'investissement en capital humain découle d'une perspective de cycle vital et qu'en fait le problème doit être abordé d'un point de vue dynamique. De plus, la production de capital humain par un individu est dépendante d'une technologie qui lui est propre, en ce sens que le travailleur cherche à combiner du mieux qu'il peut ses dotations en temps et ressources, incluant sa dotation en capital humain, afin de dériver son programme optimal d'investissement, étant donné les contraintes auxquelles il fait face. L'aspect purement «technologique» a également été abordé par Becker (1967) lors du fameux «Woytinsky Lecture». La perspective dynamique mise de l'avant par Ben-Porath, bien que se prêtant difficilement à un traitement explicitement empirique (on ne peut dériver directement une forme réduite estimable à partir des conditions de premier ordre du programme d'optimisation inter-temporel), nous donne des prédictions précises quant au profil temporel de

l'investissement et du revenu pour un individu. Nul doute que ces prédictions sont à la base de la dérivation de l'équation de salaire de Mincer.²

2.2) Capital humain général et capital humain spécifique.

La distinction fondamentale proposée par Becker (1962 et 1975) quant au type d'investissement a des implications fort différentes en ce qui concerne le financement de l'investissement et le rendement qui en découle. Supposons qu'une firme et un travailleur entament une relation d'emploi. Faisons en outre l'hypothèse que le marché du travail ainsi que le marché des biens fonctionnent en concurrence parfaite. S'il n'y a aucune formation suite à l'embauche de ce travailleur, son salaire sera évidemment égal à sa productivité marginale, et ce, en tout temps. Si, par contre, il y a investissement, la relation entre le salaire et la productivité marginale s'en trouvera modifiée, tel que le souligne Becker:

«Training might lower current receipts and raise current expenditures, yet firms could profitably provide this training if future receipts were sufficiently raised or future expenditures sufficiently lowered. Expenditures during each period need not equal wages, receipts need not equal the maximum possible marginal productivity, and expenditures and receipts during all periods would be interrelated»³

Si l'investissement accroît la productivité du travailleur dans son emploi actuel de même que dans tout autre emploi éventuel, il va de soi que les forces concurrentielles amèneront la firme à payer le travailleur au taux du marché, sans quoi le travailleur quitterait. Pourquoi la firme voudrait-elle alors dispenser un tel programme de formation de caractère purement général si elle n'en reçoit aucun bienfait dans l'avenir et qu'en plus elle doive en assumer les frais

² Parallèlement à Ben-Porath, Becker (1967) a également analysé le profil d'investissement et de revenu.

³ Becker(1975), page 18.

directs? La réponse est que la firme serait prête à donner une formation parfaitement générale à son travailleur si ce dernier payait pour sa formation. Le travailleur sera disposé à le faire puisque la formation acquise lui permet d'accroître son salaire futur, et ce, quel que soit l'emploi occupé.⁴ Par conséquent, durant la période de formation, le travailleur recevra un salaire inférieur à la valeur de sa productivité marginale sur le marché et dès que la formation sera complétée, son salaire sera de nouveau fixé par le marché, c'est-à-dire à sa (nouvelle) valeur de la productivité marginale. Pour la formation générale donc, le marché fait en sorte que tous les coûts sont défrayés par le travailleur et que tous les bénéfices lui échoient. Une firme qui défrayerait en tout ou en partie les coûts de formation (sans être compensée par un salaire inférieur) ne pourrait être viable dans un monde concurrentiel puisqu'elle ne pourrait garder le travailleur qu'en lui payant un salaire équivalent à ce qu'il pourrait obtenir sur le marché, ce qui implique forcément des pertes pour cette firme. Ce mécanisme par lequel les forces du marché agissent sur le salaire du travailleur est illustré dans la figure 1. Pour fins de simplification, il y est supposé que la relation d'emploi ne dure que deux périodes, la première étant consacrée à la formation et la seconde où il n'y a pas de formation.

La courbe A indique la productivité marginale du travailleur s'il ne reçoit aucune formation alors que la courbe B fait de même dans le cas où il est formé par son employeur actuel. On remarque qu'en première période, le travailleur reçoit un salaire inférieur à ce qu'il pourrait obtenir sur le marché alors qu'en deuxième période, grâce à son investissement sa valeur marchande s'accroît au delà de ce qu'elle aurait été s'il n'avait été formé, d'où l'incitation pour lui à accepter un salaire de départ inférieur.

Supposons maintenant que l'investissement soit de nature à n'augmenter la productivité du travailleur que dans l'emploi actuel, laissant inchangée sa productivité marginale sur le marché externe. Par conséquent, le salaire que cet employé peut obtenir sur le marché est complètement indépendant de toute formation dans son emploi présent.

⁴ Nous faisons abstraction, pour l'instant, des questions liées à l'incertitude quant au rendement futur de l'investissement.

S'ensuit-il nécessairement que le salaire versé au travailleur par son employeur est également indépendant de toute formation spécifique? A priori, on pourrait supposer que oui en ce sens que la firme défrayerait tous les coûts liés à la formation (sans payer un salaire inférieur au salaire du marché) et empocherait par la suite tous les rendements. Ce schéma découlerait du refus des travailleurs de financer de quelque façon que ce soit le programme de formation spécifique de peur que l'employeur ne mette fin unilatéralement à la relation d'emploi lorsque viendrait le temps pour le travailleur de recouvrer son investissement par un salaire supérieur au salaire du marché. Le travailleur subirait alors une perte de capital. Cependant, selon l'argument avancé par Becker, le même genre de raisonnement s'applique aussi pour la firme. Supposons que la firme supporte tous les coûts liés à la formation spécifique. Comme la firme paie le travailleur au salaire du marché, elle se trouve à être en tout temps sujette à un départ du travailleur entraînant alors une perte de capital pour la firme qui ne pourra jamais recouvrer son investissement par le biais d'une productivité plus élevée.

«Turnover becomes important when costs are imposed on workers or firms, which are precisely the effects of specific training. Suppose a firm paid all the specific training costs of a worker who quit after completing the training. According to our earlier analysis, he would have been receiving the market wage and a new employee could be hired at the same wage. If the new employee were not given training, his marginal product would be less than that of the one who quit since presumably training raised the latter's productivity. Training could raise the new employee's productivity but would require additional expenditures by the firm. In other words, a firm is hurt by the departure of a trained employee because an equally profitable employee could not be obtained. In the same way an employee who pays for specific training would suffer a loss from being laid off because he could not find an equally good job elsewhere»⁵

⁵ Becker (1975), page 29. La question de la relation entre la spécificité du facteur travail et son roulement a été également étudiée par Oi (1962).

La solution pour Becker est de prendre conscience que le roulement (et la perte de capital qui en découlerait) peut être réduit en payant au travailleur qui vient de compléter sa formation un salaire supérieur à ce qu'il obtiendrait sur le marché. Afin de ramener l'équilibre entre l'offre et la demande de travailleurs en formation, il suffirait alors à la firme de faire partager une partie des coûts de la formation. En bout de ligne, la firme ne paie ni la totalité des coûts non plus qu'elle n'empoche tout le rendement, les deux se trouvant à être partagés avec le travailleur. La figure 2 illustre ce processus. En première période, le travailleur est rémunéré à un salaire inférieur à ce qu'il pourrait obtenir sur le marché puisqu'il paie en partie pour sa formation spécifique. À noter également que le salaire de ce travailleur est supérieur à sa productivité marginale dans la firme, puisqu'il consacre une partie de son temps à être formé. En deuxième période toutefois, la firme recouvre son investissement en rémunérant ce travailleur à un salaire inférieur à sa productivité marginale et l'employé y trouve son compte puisque son salaire est maintenant au dessus de sa meilleure alternative. À ce stade de la discussion, il est sans doute utile de préciser les implications empiriques de la théorie développée par Becker. D'abord, le salaire croît avec l'expérience *et* l'ancienneté parce que la productivité croît également avec l'expérience et l'ancienneté lorsqu'il y a formation générale ainsi que formation spécifique à l'emploi.

Hashimoto (1981) a par la suite développé de façon formelle l'hypothèse de partage suggérée par Becker. Selon Hashimoto, le fait que la firme et le travailleur partagent les coûts et le rendement liés à l'investissement n'a rien à voir en soi avec l'incertitude créée par la possibilité qu'une ou l'autre partie mette fin à la relation d'emploi, mais est plutôt la conséquence de coûts de transactions durant la ou les périodes suivant l'investissement. Ces coûts découlent de la difficulté qu'ont les parties à s'entendre sur la valeur de la productivité marginale du travailleur au sein de la firme et hors de la firme. Ce sont ces coûts qui créent de l'incertitude quant à la «capturabilité» des rendements. Étant donné l'importance de la contribution d'Hashimoto, il semble approprié d'explicitier son raisonnement en détail.⁶

⁶ Ce qui suit reprend presque textuellement son exposé en page 475.

Supposons qu'un travailleur et une firme créent un lien d'emploi d'une durée de deux périodes et que les deux parties sont neutres par rapport au risque. L'emploi implique qu'il y a un investissement en capital humain spécifique à la firme en première période et l'employeur de même que le travailleur décident de la taille de l'investissement ainsi que de son partage. Au début de la deuxième période, toute l'information pertinente est révélée aux deux parties et chacun doit décider si la relation se poursuit ou non. Le travailleur détient H unités de capital humain parfaitement transférables, c'est-à-dire général, et sa fonction de coût pour produire h unités de capital humain est donnée par $C(h)$ avec $C' > 0$ et $C'' > 0$. La valeur pour la firme par unité de h est m , alors que h n'a aucune valeur sur le marché. En d'autres termes, le capital est parfaitement spécifique et par conséquent non transférable. De plus, les deux parties s'entendent sur la valeur de m . Il n'y a donc aucun coût de transaction. En deuxième période, la valeur du produit marginal de ce travailleur dans la firme est donnée par:

$$v = H + mh \quad (1)$$

Étant donné les hypothèses formulées plus haut, la valeur du produit marginal de ce travailleur sur le marché est donnée par:

$$y = H \quad (2)$$

Par conséquent, le rendement de l'investissement, R , sera:

$$R = v - y = mh \quad (3)$$

Le salaire de deuxième période (w) de cet individu est donc composé de son salaire dans la meilleure alternative, y , et d'une prime représentant sa part de l'investissement:

$$w = y + \alpha R, 0 \leq \alpha \leq 1 \quad (4)$$

où α est la part que le travailleur reçoit du rendement sur le capital spécifique. Quant à l'employeur, sa part est donnée par:

$$r = v - w = (1-\alpha)R \quad (5)$$

Supposons maintenant que l'investissement s'avère être une erreur ex post. Peu importe qui prend la décision de cesser la relation, cette décision sera optimale pour toute valeur de α . En effet, l'employeur voudra congédier si $r \leq 0$ alors que l'employé partira si $w \leq y$. Si $\alpha=0$, il s'ensuit que $w=y$ et le travailleur est indifférent entre quitter ou ne pas quitter mais l'employeur, lui, ne l'est pas. Si $\alpha=1$ (i.e. $r=0$), alors à ce moment l'employeur est indifférent à ce que la relation se poursuive ou non, mais l'employé ne l'est pas. Dans un cas comme dans l'autre, une des deux parties prend la bonne décision pour les deux. Si α est entre 0 et 1, les deux parties voudront se séparer. Par conséquent, la valeur de α n'a rien à voir avec le fait de séparer ou non puisque les deux parties connaissent parfaitement et sans coût les valeurs de v et de y . Hashimoto dresse ensuite le parallèle avec le théorème de Coase en soulignant que peu importe la façon dont les droits de propriété sur les rendements de l'investissement sont assignés (α), l'absence de coûts de transaction sur ces droits ne change en rien l'efficacité économique de l'utilisation de la ressource. Il en conclut donc que le coût et le rendement de l'investissement sont partagés en raison de la présence de coûts de transaction, c'est-à-dire des coûts pour s'entendre sur les valeurs de v et de y . Le reste de son article s'attache à dériver le contrat optimal (α) en présence d'incertitude quant aux valeurs de v et de y tel que les deux parties minimisent les pertes d'efficacité dues à des décisions de séparation non optimales. L'implication du point de vue empirique du modèle de Hashimoto reste la même que dans la cas de Becker: le salaire du travailleur augmentera avec l'ancienneté dans la firme.

2.3) Technologie de production du capital humain et profil d'investissement dans le temps.

Jusqu'à présent, notre attention s'est essentiellement portée sur le côté «demande» pour le capital humain. Deux aspects étroitement liés seront maintenant l'objet de notre

attention, soit la technologie de production du capital humain ainsi que le profil temporel optimal d'investissement (avec le profil de revenu qui lui est associé). Ce dernier aspect est important puisqu'il permet de faire la transition entre le côté purement théorique et l'application pratique de la théorie. Cette section sera surtout consacrée à une révision de deux articles fortement apparentés qui cernent bien les questions pertinentes, soit le «Woytinski Lecture» de Becker⁷ ainsi que la contribution de Ben-Porath (1967).

En utilisant le même cadre théorique que celui proposé pour l'investissement en capital physique, Becker démontre qu'il y a une limite à ce qu'un individu peut investir, même en supposant la constance du rendement sur l'investissement pour un individu typique. La figure 3 montre la relation entre la demande pour le capital humain, représentée par le bénéfice marginal d'un dollar additionnel investi et l'offre, représentée par le coût marginal de financer un dollar supplémentaire d'investissement.

Évidemment, si le taux de rendement marginal est supérieur au taux d'intérêt, le revenu de l'individu pourrait être augmenté par un investissement supplémentaire et l'inverse est vraie si le coût est plus grand que le taux de rendement. L'équilibre se trouvera alors à l'intersection des deux courbes. Comme on suppose que l'individu est preneur de prix pour son capital humain, il faut évidemment trouver d'autres facteurs qui expliquent que sa courbe de demande soit de pente négative. Un des facteurs invoqués par Becker est la présence de rendements marginaux décroissants:

«The principal characteristic that distinguishes human capital from other kinds of capital is that, by definition, the former is embodied or embedded in the person investing. This embodiment of human capital is the most important reason why marginal benefits decline as additional capital is accumulated. One obvious implication of embodiment is that since the memory capacity, physical

⁷ Publié à l'origine en 1967 par l'Université du Michigan et repris dans Becker (1975).

size, etc. of each investor is limited, eventually diminishing returns set in from producing additional capital. The result is increasing marginal costs of producing a dollar of returns.»⁸

Un autre facteur dont il faut tenir compte est l'utilisation du temps par l'individu afin d'accroître son stock de capital humain. Or, si l'individu utilise une partie de son temps pour investir, il va de soi qu'il ne peut en même temps utiliser ce temps à travailler d'où une perte de revenu. De plus, comme il est raisonnable de supposer que l'élasticité de substitution entre le temps de l'individu et les autres inputs nécessaires à l'accroissement de son stock de capital est imparfaite, il en découle que ce stock ne sera pas ajusté de façon instantanée et que l'investissement se fera durant un certain laps de temps. On remarque alors que dans l'optique où les agents se comportent de façon optimale et cherchent à maximiser la valeur présente des bénéfices découlant d'un investissement donné, se profile un sentier optimal d'investissement où en chaque point tous les inputs (y compris le temps de l'individu) sont utilisés efficacement. En d'autres termes, il y a une combinaison optimale d'intrants tout comme il y a une trajectoire temporelle optimale d'investissement. Par contre, bien que le stock de capital soit ajusté sur un laps de temps afin de minimiser les coûts, il n'est pas optimal d'allonger indûment la période d'investissement en raison des considérations suivantes:

- (i) la durée de vie étant limitée, investir plus tard implique que la période pendant laquelle les rendements sont réalisés est plus courte, d'où un montant total de bénéfices plus faible.

- (ii) la valeur présente des bénéfices est inférieure lorsque l'investissement est retardé.

⁸ Becker (1975), pages 98-99.

(iii) le fait d'utiliser son propre temps afin d'augmenter son stock de capital implique *ipso facto* que l'individu ne peut utiliser une partie de son capital humain existant afin d'accroître son revenu de travail. On peut présumer que plus l'individu est jeune, moins la valeur de son temps est élevée puisqu'il a eu moins de temps pour accumuler du capital humain. D'où l'incitation à investir davantage lorsque l'on est jeune.

En ce qui concerne la pente de la courbe d'offre, Becker fait l'hypothèse que les individus n'ont pas un accès illimité à un marché du capital homogène. En raison des subventions à l'éducation et des taxes, le marché du capital est très segmenté et les individus font face à une grille de taux d'intérêts qui reflète les coûts croissants des fonds selon leur source. Par exemple, si l'État subventionne complètement le système d'éducation aux niveaux primaires et secondaires, il en coûte peu pour l'individu d'investir en capital humain. Si par la suite cet individu veut aller à l'université et que l'État se charge de payer les intérêts sur ses emprunts, cela lui coûte déjà plus cher de poursuivre son éducation. Plus tard, lorsqu'il est sur le marché du travail, le coût de financer un programme de formation sera possiblement sous la forme d'un salaire plus faible, et ainsi de suite.

La dérivation formelle des résultats touchant à la forme du profil temporel de l'investissement a été traitée par Ben-Porath (1967) dans un article qui fut précurseur d'un pan entier de la littérature traitant du capital humain et de l'offre de travail, bien que ce dernier aspect n'était pas traité par Ben-Porath. À ce titre, il convient donc de donner un exposé détaillé de son modèle.

Le point de départ du modèle est de considérer le stock de capital humain comme s'il s'agissait d'un stock de capital physique, sujet à dépréciation et dont les services sont rémunérés. On suppose que chaque individu ne dispose que d'une fraction infime du stock de capital humain disponible dans l'économie en un point donné et que, par conséquent, il est un preneur de prix en ce qui concerne le taux a_0 auquel il loue les services de son capital. Il peut

donc offrir la quantité qu'il veut. Le revenu potentiel au temps t , Y_t , est par conséquent égal au produit du taux a_0 par la quantité maximale de services en capital humain que l'individu peut offrir, K_t .

$$Y_t = a_0 K_t \quad (6)$$

Si on définit E_t comme étant le revenu disponible en période t , soit le revenu net de toute dépense faite en vue d'accroître son stock de capital humain, le coût de l'investissement I_t sera évidemment égal à $Y_t - E_t$. Tout comme dans le "Woytinski Lecture" de Becker, l'individu a une fonction de production du capital humain:

$$Q_t = \beta_0 (s_t K_t)^{\beta_1} D_t^{\beta_2} \quad (7)$$

où $\beta_1, \beta_2 > 0$ et $\beta_1 + \beta_2 < 1$; Q_t est la quantité de capital humain produite; D_t est la quantité d'intrants achetés au prix P_d nécessaires à la production de capital humain; s_t est la part comprise entre zéro et un du stock existant de capital humain consacré à la production d'autre capital humain. Ben-Porath souligne que s'il n'y a pas de production conjointe de capital humain et de revenu, s_t est également la proportion du temps consacré à la production de capital humain. Cette hypothèse équivaut à éliminer la possibilité que le travailleur apprenne sur le tas. Le taux de changement du stock de capital est donné par:

$$\frac{dK_t}{dt} = Q_t - \delta K_t \quad (8)$$

où δ est le taux de dépréciation du capital. Les coûts d'investissement comportent donc deux composantes:

$$I_t = a_0 s_t K_t + P_d D_t \quad (9)$$

La première composante à la droite du signe d'égalité représente le coût de renonciation ou, si vous voulez, la valeur marchande des services productifs qui servent à augmenter le stock de capital humain. Le deuxième terme du membre de droite représente les coûts directs en

achats de biens et services requis pour produire le capital. À noter que les concepts sont essentiellement les mêmes que ceux de Becker. Toutefois, Ben-Porath inclut explicitement *dans les coûts* les revenus auxquels renonce l'individu lorsqu'il investit alors que Becker les incluait du côté de la demande en les déduisant du bénéfice marginal par dollar investi. Si on minimise les coûts par les choix de s_t et de D_t sous contrainte de la fonction de production donnée par (7), on obtient la condition suivante pour un minimum:⁹

$$\frac{a_0 s_t K_t}{P_d D_t} = \frac{\beta_1}{\beta_2} \quad (10)$$

En substituant (10) et (7) dans (9) on a alors les coûts d'investissement exprimés en termes de quantité de capital humain produit:

$$I_t = \frac{\beta_1 + \beta_2}{\beta_1} a_0 \left(\frac{\beta_1 P_d}{\beta_2 a_0} \right)^{\frac{\beta_2}{(\beta_1 + \beta_2)}} \left(\frac{Q_t}{\beta_0} \right)^{\frac{1}{(\beta_1 + \beta_2)}} \quad (11)$$

L'objectif de l'individu est donc de maximiser, par le choix de s et de D , la valeur présente de son revenu disponible ($Y_t - I_t$):¹⁰

$$W_t = \int_t^T \exp^{-rv} [a_0 k(v) - I(v)] dv \quad (12)$$

sous contrainte de la fonction de production du capital humain, de l'équation du taux de changement du capital ainsi que de la condition sur s_t (qui doit être compris entre zéro et un).

⁹ En faisant l'hypothèse que s_t est compris de façon stricte entre zéro et un.

¹⁰ Les autres hypothèses formulées par Ben-Porath sont les suivantes: (i) l'utilité de l'individu n'est pas fonction des activités requérant du temps comme intrant; (ii) il y a une quantité fixe de temps qui puisse être allouée à des activités qui produisent un revenu ou du capital humain; (iii) le stock de capital humain K_t est homogène et se déprécie à un taux δ exogène; (iv) le stock de capital humain n'est pas un argument de la fonction d'utilité; (v) il est possible d'emprunter et de prêter de façon illimitée à un taux d'intérêt constant r .

Par l'application du principe du maximum de Pontryagin¹¹, Ben-Porath obtient ensuite des résultats de dynamique comparée qui permettent de caractériser le sentier optimal d'investissement et, par conséquent, le profil du revenu de l'individu tout le long de son cycle de vie. Sans trop entrer dans les détails, mentionnons toutefois les résultats suivants:

(i) Par l'équation (11) ainsi que par les conditions de premier ordre du problème intertemporel auquel fait face l'individu, on peut obtenir les courbes de coût marginal de production du capital humain de même que la courbe de demande. Pour ce qui est de la courbe de coût marginal, il suffit de dériver l'équation (11) par rapport à Q d'où l'on obtient:

$$Cm_t = \frac{a_0}{\beta_0 \beta_1} \left(\frac{\beta_1 P_d}{\beta_2 a_0} \right)^{\frac{\beta_1}{(\beta_1 + \beta_2)}} \left(\frac{Q_t}{\beta_0} \right)^{\left(\frac{1}{(\beta_1 + \beta_2)} \right)^{-1}} \quad (13)$$

Remarquons que le coût est croissant de façon monotone en Q_t (les rendements d'échelle étant décroissants) et qu'il est indépendant du stock de capital de l'individu. Par ailleurs, les conditions de premier ordre du problème permettent de dégager la valeur à tout instant t d'acquérir une unité additionnelle de capital humain. Cette valeur actualisée (ou «shadow price») représentant l'ajout au revenu que la partie non dépréciée de cette unité engendrera tout au long de la vie de l'individu à partir de l'instant t jusqu'à T est donnée par:

$$P_t = a_0 \int_t^T \exp^{-(r+\delta)v} dv = \frac{a_0}{r+\delta} (1 - \exp^{-(r+\delta)(T-t)}) \quad (14)$$

On remarque que cette valeur est indépendante du stock de capital existant ainsi que du nombre d'unités de capital ajoutées, ce dernier résultat découlant évidemment de l'hypothèse voulant que l'individu soit un preneur de prix pour son capital. Par contre, comme l'horizon

¹¹ Pour un exposé concis du principe du maximum, voir Intriligator (1971).

temporel de l'individu est fini, la valeur diminue avec le temps puisque son capital sera utilisé sur une période de temps qui va en diminuant. La quantité optimale de capital humain sera donnée par l'intersection de ces deux courbes (voir figure 4), ce qui nous donne:

$$Q_t = N(1 - \exp^{-(r+\delta)(T-t)})^{\frac{(\beta_1+\beta_2)}{(1-\beta_1-\beta_2)}} \geq 0 \quad (15)$$

À noter que les résultats dérivés ci-dessus s'appliquent au cas où s_t est strictement inférieur à un, c'est-à-dire lorsque l'individu consacre à la fois du temps à investir et à louer les services de son stock existant de capital humain ou lorsqu'il n'investit pas du tout.¹²

Pour tout $t < T$, le «shadow price» du capital humain est positif, bien que décroissant en t , et comme la courbe de coût marginal est monotone croissante à partir de l'origine, il y aura toujours une quantité positive de capital humain produite. Pour $t=T$, par contre, il n'y a aucun ajout puisque le capital humain perd toute valeur. Par conséquent, durant toute la vie active, l'individu accroîtra son stock de capital humain, mais à un rythme décroissant jusqu'au moment où il se retirera. À noter que les implications quant au stock de capital humain *net* sont quelque peu différentes de celles pour le stock brut tel qu'explicitées plus haut. Le changement dans le stock net dépendra à la fois du stock brut existant ainsi que du taux de dépréciation δ . Pour le cas usuel où le stock initial n'est pas très important, le stock net de capital humain augmentera durant un certain temps et lorsque les ajouts bruts (décroissants) ne suffiront plus à remplacer les unités dépréciées, le stock net diminuera jusqu'à la fin de la vie.

Évidemment, l'évolution des stocks net et brut se transpose directement sur les profils des revenus disponible et potentiel. Le taux de changement par rapport au temps du revenu disponible est donné par:

¹² Trois phases peuvent être distinguées, la première où l'individu se consacre entièrement à accroître son stock de capital ($s_t=1$), la deuxième où $0 < s_t < 1$ telle que décrite dans le texte, et finalement celle où $s_t=0$ et l'individu cesse d'investir.

$$\frac{dE_t}{dt} = a_0 \frac{dK_t}{dt} - \frac{dI_t}{dt} \quad (16)$$

Définissons par ailleurs le revenu observé comme:

$$\hat{E}_t = a_0 K_t - \frac{\beta_1}{\beta_1 + \beta_2} I_t \quad (17)$$

Sa dérivée par rapport au temps sera alors:

$$\frac{d\hat{E}_t}{dt} = a_0 \frac{dK_t}{dt} - \frac{\beta_1}{\beta_1 + \beta_2} \frac{dI_t}{dt} = a_0 Q_t - \frac{\beta_1}{\beta_1 + \beta_2} (Cm_t) \frac{dQ_t}{dt} - a_0 \delta K_t \quad (18)$$

Le revenu observé sera plus élevé que le revenu disponible en raison des achats d'intrants servant à produire du capital humain (P_dD) et il sera inférieur au revenu potentiel en raison de la part de ce dernier consacrée à produire des unités additionnelles de capital. Par ailleurs, comme $dI_t/dt < 0$, le changement du revenu observé sera toujours plus grand que la variation du revenu potentiel. Par conséquent, tel qu'illustré à la figure 5, lorsqu'un individu atteint son revenu observé maximum, il est déjà dans la phase de déclin de son revenu potentiel. Il n'y aurait pas de déclin, que ce soit pour le revenu potentiel ou le revenu observé, s'il n'y avait aucune dépréciation.¹³ La stricte concavité du profil du revenu observé est assurée si $\delta=0$ alors que pour le cas où il y a dépréciation, la courbe sera convexe dans un voisinage de $t=T$. Toute portion convexe au début de la vie active est exclu par l'hypothèse de rendements

¹³ Soulignons ici que Ben-Porath prend comme exogène l'offre inter-temporelle de travail. Si on considère l'offre de travail comme étant une des variables de contrôle du modèle, il est possible que le revenu observé de l'individu diminue vers la fin de sa vie, même sans dépréciation. À ce sujet, voir Heckman (1976).

décroissants.¹⁴ À noter que ces résultats sont applicables à la phase correspondant à $s_t < 1$. En ce qui concerne la phase où l'individu investit à plein temps ($s_t = 1$), Ben-Porath note que le processus d'optimisation peut impliquer que l'investissement augmente avec le temps lorsque t est petit. En effet, si le stock initial est petit, il se peut que même en combinant tous les intrants, y compris ce stock, la production de capital humain en résultant soit inférieure au niveau optimal. De plus, lorsque l'individu est jeune, la valeur d'une unité de capital humain supplémentaire (son «shadow price») n'est que relativement affectée par le fait que l'horizon temporel soit fini, de sorte que la demande pour le capital humain est plus forte que lorsque l'individu avance en âge. Toutes ces raisons font en sorte que l'individu consacre tout son temps à la production de capital humain lorsqu'il est jeune. On associe habituellement cette phase du cycle de vie avec la période de scolarisation de l'individu.

En guise de conclusion, on notera que par rapport à l'article précurseur de Becker (1962), le modèle de Ben-Porath ainsi que la contribution de Becker (1967) cherchaient tous deux à fournir une argumentation théorique plus complète afin de caractériser le comportement d'un individu voulant accroître son potentiel de revenu futur. Cette caractérisation, surtout en ce qui concerne Ben-Porath, a permis de dériver un profil typique de revenu de l'individu tout au long de son cycle de vie qui s'apparente aux profils observés. Notons toutefois les hypothèses simplificatrices quant à l'offre de travail qu'on suppose prédéterminée ainsi qu'à l'absence d'incertitude. Ces deux aspects ont été traités ultérieurement par Heckman (1976) qui a rendu endogène la décision d'offre de travail et par Macurdy (1981) qui prit en compte l'incertitude.¹⁵ À la lecture de Mincer (1974), Heckman note que les profils empiriques âge-revenu provenant de coupes transversales montrent que ceux-ci ont tendance à atteindre un sommet dans le courant de la vie active et à décliner par

¹⁴ Comme Becker l'a noté dans son «Woytinski Lecture», il est possible qu'il y ait des rendements croissants au début de la vie active si les ajouts de capital humain accroissent la capacité d'en acquérir d'autres à un rythme plus grand qu'ils n'augmentent le coût d'option, de sorte que le bénéfice marginal net serait croissant pour un certain intervalle. Éventuellement cependant, les rendements décroissants du capital en termes de capacité d'apprentissage ferait en sorte que le profil du revenu potentiel serait concave.

¹⁵ Voir Weiss (1986) pour un survol critique de la littérature portant sur les questions de choix inter-temporels d'investissement en capital humain et d'offre de travail.

la suite alors que les profils âge-salaire horaire n'ont pas cette forme concave. Des modèles d'investissement optimal en capital humain, tel que celui de Ben-Porath, qui ne tiennent pas compte du choix endogène des heures travaillées ne peuvent expliquer ce phénomène alors que les modèles qui prennent en considération cette dimension le peuvent sans difficulté.

2.4) L'équation de salaire de Mincer.

S'il est vrai que les modèles à la Ben-Porath permettent de caractériser de façon rigoureuse le profil de revenu dans le temps, leur application empirique pose toutefois des problèmes considérables, comme le note Willis (1986):

«Unfortunately, the optimal human capital models are very difficult to implement rigorously in empirical work. First, they typically do not have a closed form solution so that the precise functional form for life cycle earnings implied by such a model is usually not known. [...] Second, many of the concepts underlying the model, including human capital itself, are unobservable (or, at least, not usually measured in available data). In addition to human capital, the list of unobservables includes the rental rate on human capital, the rate of discount, the functional form of the human capital production function, the inputs of time and purchased goods used in the investment, and the individual-specific parameters of the production function which may be interpreted as representing the interactions of individuals' "learning ability" with the home, school, and work environments where learning takes place.»

Cette section mettra en relief la façon dont Mincer s'y est pris pour tenter de trouver un heureux compromis entre la rigueur théorique et l'applicabilité empirique. Notons au départ que la problématique de la formation sur le tas («on-the job training») et son impact sur la distribution des revenus dans l'économie avait été traitée antérieurement par Mincer

dans deux articles parus en 1958 et en 1962 dans le *Journal of Political Economy*. Le premier article s'attardait spécifiquement à démontrer qu'un modèle où les individus acquéraient des niveaux différents d'habiletés avec le temps passé dans leur emploi pouvait expliquer la distribution observée des revenus dans l'économie.¹⁶ Quant au second article, il était explicitement consacré à tenter d'estimer le montant des investissements en formation sur le tas et à calculer des taux de rendement sur ces investissements. C'était, en somme, une première tentative visant à mesurer l'impact de la formation ainsi que de la scolarité sur les revenus. La façon dont Mincer s'y est pris à l'époque fut de comparer les profils de revenus moyens de deux groupes qui se différencient par le niveau de scolarité tel que les diplômés de «high school» et les diplômés collégiaux:

«Taking this comparison as an example, the procedure involves year-by-year estimation of training costs which a high-school graduate must incur in order to acquire a college education and the additional amount of training on the job which is, on the average, characteristic of college graduates. Such estimates are obtained on the assumption that the rate of return is the same on each year's investment whether at school or on the job. In any given year j after high-school graduation, those who go on to, or have graduated from, college would have earnings (Y_j) which equal the earnings of high-school graduates (X_j) plus the income earned on differential investment in training made since graduation from high-school [...] Costs of (incremental) training in year j are, therefore, measured by the difference between Y_j and X_j augmented by the (foregone) return on the previous (incremental) costs.»¹⁷

¹⁶ À noter qu'en ces premiers temps où la formation sur le tas intéressait les économistes, Mincer liait de façon étroite le choix de l'occupation avec l'acquisition d'habiletés, différentes occupations supposant différents niveaux de formation.

¹⁷ Mincer (1962) pp.53-54.

On remarque donc que la démarche de Mincer repose sur une approche ultra pragmatique si on peut dire, mais tout de même cohérente: dans son article de 1958, il établit que la distribution (asymétrique) observée des revenus dans l'économie pouvait être expliquée par un modèle simple d'acquisition de capital humain basé sur un comportement «économique» d'individus effectuant des choix conformes à leur préférences. Il résoud d'une certaine façon le paradoxe mis en évidence par Pigou (1932) selon lequel il était difficile de combiner l'hypothèse somme toute raisonnable que les habiletés étaient distribuées normalement dans la population alors que la distribution des revenus ne l'était pas. Par la suite, s'appuyant en partie sur ses travaux ainsi que sur ceux de Becker (1962), il va directement à la conclusion que les différences de revenus observées sont le résultat (exclusif) de différences dans les montants (et le temps) investis pour acquérir des habiletés génératrices de revenus. Les calculs des taux de rendement et des coûts de l'investissement s'effectuent alors mécaniquement.

Dans son livre publié en 1974, Mincer pousse davantage ses investigations en s'appuyant cette fois sur les développements théoriques liés à la théorie du capital humain qui se sont produits entre-temps. Signalons ici la contribution de Ben-Porath, évidemment, mais aussi celle de Becker et Chiswick (1966) qui s'avéra importante du point de vue empirique et qui consistait à faire ressortir (sous certaines hypothèses) la correspondance qui existe entre les montants d'argent investis en capital humain et qui ne sont pas observables et le temps consacré à investir qui, lui, est observable.

Le point de départ de Mincer menant à son équation de salaire est ce qu'il dénomme le «schooling model» qui vise à dériver de façon analytique l'effet de l'investissement en terme de scolarité sur le revenu. Par la suite, il généralise cette procédure afin d'inclure l'investissement fait durant l'emploi.

Soit

n = durée de vie active sur le marché du travail (fixe).

Y_s = revenus annuels d'un individu ayant à son actif s années de scolarité.

V_s = valeur présente au début de la période de scolarité des revenus d'un individu tout au long de sa vie.

r = taux d'actualisation.

t = temps en années.

d = différence de scolarité en années.

Alors,

$$V_s = Y_s \int_s^{n+s} \exp^{-rt} dt = \frac{Y_s(\exp^{-rs} - \exp^{-r(n+s)})}{r} \quad (19)$$

De même pour un individu ayant $s-d$ années de scolarité:

$$V_{s-d} = Y_{s-d} \int_{s-d}^{n+s-d} \exp^{-rt} dt = \frac{Y_{s-d}(\exp^{-r(s-d)} - \exp^{-r(n+s-d)})}{r} \quad (20)$$

Le rapport entre les revenus annuels après s années de scolarité et les revenus après $s-d$ années de scolarité, $k_{s,s-d}$, est donnée par l'égalisation des deux valeurs présentes nettes:

$$k_{s,s-d} = \frac{Y_s}{Y_{s-d}} = \frac{\exp^{-r(s-d)}}{\exp^{-rs}} = \exp^{rd} \quad (21)$$

Si on définit $k_{s,0} = Y_s/Y_0 = k_s$, on alors $k_s = \exp^{rs}$. En prenant le logarithme naturel des deux côtés on obtient:

$$\ln Y_s = \ln Y_0 + rs \quad (22)$$

Dans sa forme la plus primitive, pour reprendre l'expression de Mincer, l'équation (22) montre que l'accroissement en pourcentage du revenu est strictement proportionnel au temps

passé à l'école et que le taux de rendement est donné par le coefficient de proportionnalité.¹⁸ Comme le souligne Willis (1986), l'équation (22) n'est qu'une tautologie découlant de la définition de la valeur présente alors que si on adjoint l'indice supplémentaire i représentant l'individu à chacun des membres de l'équation, on se trouve alors à inclure implicitement une hypothèse sur la technologie de production du capital humain, $\ln Y_{oi}$ étant la capacité «de départ» de l'individu i à engendrer des revenus alors que r_i serait son taux de rendement personnel lié au niveau de scolarité s .¹⁹

S'inspirant des travaux de Becker (1967) et de Ben-Porath (1967), Mincer généralise ce modèle afin de tenir compte du fait que l'acquisition d'habileté ne cesse pas nécessairement avec l'entrée sur le marché du travail. Ainsi, si l'on dénote par C_j les ressources consacrées par l'individu à l'acquisition d'habiletés supplémentaires au temps j alors qu'il est au travail, son revenu net serait alors obtenu en déduisant C_j de son revenu brut ou potentiel E_j . Si on généralise ce processus de formation à plusieurs périodes, le revenu net à tout instant j est alors égal à:

$$Y_j = Y_s + \sum_{t=0}^{j-1} r_t C_t - C_j = E_j - C_j \quad (23)$$

où Y_s est le revenu potentiel initial de l'individu au moment où il quitte l'école. Quant à l'évolution du revenu avec le temps, elle est donnée par:

$$\Delta Y_j = Y_{j+1} - Y_j = r_j C_j - (C_{j+1} - C_j) \quad (24)$$

Le revenu net de l'individu augmentera donc avec l'expérience tant et aussi longtemps qu'il y aura un investissement net positif qui diminue avec le temps ou, à tout le moins, qui

¹⁸ À noter que dès 1958, Mincer avait utilisé un modèle identique en tout point si ce n'est que le mot «schooling» était remplacé par «training». Aussi, il s'était abstenu de faire la transformation logarithmique menant à (22), sans doute parce qu'à l'époque, il n'était pas dans ses intentions d'estimer une telle équation.

¹⁹ Il est à noter que Mincer prend en compte du mieux qu'il peut (lire du mieux que les données le lui permettent) les différences d'«habiletés» entre individus à peu près partout dans son livre. Il en analyse même les conséquences possibles lorsqu'il traite des résidus de régression.

augmente à un taux plus faible que le taux de rendement. En ce qui concerne le revenu potentiel de l'individu, il augmente avec l'expérience si il y a investissement positif:

$$\Delta E_j = r_j C_j \quad (25)$$

À ce stade de son analyse, Mincer s'inspire des conclusions de Becker et de Ben-Porath pour suggérer que le profil temporel du revenu potentiel ainsi que du revenu brut devrait être concave, c'est-à-dire:

$$\Delta^2 E_j = r \Delta C_j < 0 \quad (26)$$

$$\Delta^2 Y_j = r \Delta C_j - \Delta^2 C_j < 0 \quad (27)$$

À noter cependant que si les investissements décroissent à un rythme croissant pendant un certain temps, l'équation (27) est alors de signe positif. Toutefois, dans une perspective de long terme, les investissements doivent éventuellement diminuer ce qui imprime une forme concave au profil de revenu net.

La figure 6 nous montre l'évolution du revenu net Y_j en fonction du nombre d'années d'expérience. On remarquera que durant les premières années d'expérience, le revenu des travailleurs qui investissent sont inférieurs au revenu qu'ils pourraient engendrer avec leur scolarité sans investir, soit Y_s . Toutefois, par leurs investissements, le revenu augmente et à un certain point dépasse Y_s . C'est ce que Mincer appelle son point de renversement («overtaking»). L'année où se produit le renversement peut être dérivée à partir de l'équation (23). En faisant l'hypothèse que les montants investis sont égaux dès l'instant où l'individu fait son entrée sur le marché du travail jusqu'au point de renversement, on a:

$$\begin{aligned} Y_j &= Y_s + r \sum_{t=0}^{t-j-1} C_t - C_j = Y_s \\ \rightarrow r \sum_{t=0}^{t-j-1} C_t &= C_j \rightarrow r \hat{j} C_j = C_j \rightarrow \hat{j} = \frac{1}{r} \end{aligned} \quad (28)$$

La prochaine étape qui vise à rendre opérationnelle l'analyse théorique consiste à transformer en unités de temps les montants investis en capital humain puisque ces derniers ne peuvent, sauf exception, être observés.²⁰ Cette conversion en unités de temps s'effectue en utilisant le même procédé que celui utilisé par Becker et Chiswick (1966) pour analyser le rendement de l'éducation. Soit k_j , le ratio des coûts d'investissement par rapport au revenu potentiel E_j en période j . Ce ratio peut être vu comme représentant la fraction du temps que l'individu consacre à accroître son revenu potentiel. Nous aurons donc:

$$C_j = k_j E_j \quad (29)$$

et

$$E_j = E_{j-1} + rC_{j-1} = E_{j-1}(1 + rk_{j-1}) \quad (30)$$

En résolvant l'équation (29) par récursivité, on obtient alors:

$$E_j = E_0 \prod_{t=0}^{j-1} (1 + rk_t) \quad (31)$$

Il s'agit ensuite de prendre l'approximation logarithmique de l'équation (31) pour obtenir:

$$\ln E_j = \ln E_0 + \sum_{t=0}^{j-1} rk_t \quad (32)$$

²⁰ Il est bien sûr possible que les firmes comptabilisent certaines dépenses liées à la formation. Toutefois, ce genre d'information sous estime probablement l'ampleur des coûts totaux. À ce sujet, voir Mincer (1962).

Comme $Y_j = E_j(1-k_j)$ et que l'on suppose que $k_j = 1$ durant la période scolaire, le log du revenu s'exprime alors comme suit:

$$\ln Y_j = \ln E_0 + r_s s + r_p \sum_{t=0}^{j-1} k_t + \ln (1-k_j) \quad (33)$$

où r_s indique le taux de rendement de l'éducation et r_p le taux de rendement lié à l'investissement post-scolaire. Si on fait l'hypothèse supplémentaire que ce dernier est constant d'une période à une autre, alors on obtient la forme «finale» qui exprime le logarithme du revenu net en fonction du temps passé à l'école et du temps passé sur le marché du travail:

$$\ln Y_j = \ln E_0 + r_s s + r_p K_j + \ln (1-k_j) \quad (34)$$

où K_j est le temps cumulé consacré à investir en capital humain avant l'année j . En réalité, ni le revenu brut et ni le revenu net ne sont observés. Toutefois, Mincer considère que le revenu observé s'apparente beaucoup plus au revenu net, de sorte que pour l'estimation, il considère que l'équation (34) offre un bon compromis. Par conséquent, si nous pouvions disposer d'un échantillon d'individus comprenant toute l'information pertinente requise, nous pourrions estimer une équation comme celle-ci:

$$\ln Y_{i,t} = \ln E_0 + r_s s_i + f(t|k_0, r_{pi}) + \epsilon_i \quad (35)$$

où $f(\)$ est une fonction quelconque de l'expérience totale sur le marché du travail et ϵ_i est un terme d'erreur résiduel qui inclut les facteurs non observables. Toutefois, comme le souligne Mincer:

«Of course, the availability of such information is not even conceivable. A more modest research objective is to abstract from individual variation in initial earning capacity ($\ln E_0$) and in rates of return on investments, and consider only the effects of the *volume* of investment on earnings. Average

parameters $\ln E_0$, r_s , and r_t would then appear in the statistically estimated coefficients [...].»²¹

De plus, les différences d'investissement post-scolaire d'un individu à un autre qui sont reflétées par les k_0 en plus des différences en terme d'expérience accumulée, font en sorte qu'il faille supprimer les indices i à l'intérieur de la fonction $f()$.

Si on fait l'hypothèse que le ratio k_t diminue de façon linéaire par rapport au temps à partir d'un ratio initial k_0 , c'est-à-dire:

$$k_t = k_0 - \frac{k_0}{T}t \quad (36)$$

où T est le temps total où il y a investissement positif, il suffit d'intégrer la version continue de l'équation (33) pour se retrouver avec une fonction concave en t , l'expérience totale cumulée sur le marché du travail:

$$\ln Y_{i,t} = \ln E_0 + r_s s_i + r_t k_0 t_i - \frac{r_t k_0}{2T} t_i^2 + \ln \left(1 - k_0 \frac{k_0}{T} t_i\right) \quad (37)$$

et cette fonction, dans sa forme estimable, devient:²²

$$\ln Y_{i,t} = a + b_1 s_i + b_2 t_i + b_3 t_i^2 + v \quad (38)$$

On voit donc que la fameuse «équation de Mincer» est le produit assez ingénieux d'un effort de concilier la rigueur théorique (puisque la forme estimable découle de considérations

²¹ On peut noter que l'apparition ultérieure de banques de données longitudinales permettent d'attaquer de front les problèmes causés par la présence d'hétérogénéité non observée.

²² En supprimant tous les indices i des coefficients, tout en reconnaissant que les individus ont des k_0 et des r_t différents, Mincer se trouve en quelque sorte à spécifier un modèle à coefficients aléatoires où une partie de l'aléa non observable est attribuable (par exemple) aux différences de rendement d'un individu à un autre.

théoriques) avec l'applicabilité empirique. Dans la prochaine section, nous ferons l'inventaire des premiers résultats d'estimation d'une telle équation. À noter toutefois qu'à ce stade d'évolution de l'analyse, aucun effort particulier n'est déployé afin de tenir compte du degré de spécificité du capital. Une généralisation en ce sens viendra plus tard. Comme Mincer était attentif à ce que l'équation estimable soit cohérente avec les analyses de Becker (1967) et de Ben-Porath (1967) concernant le profil temporel d'investissement optimal et que ces modèles ne faisaient aucune distinction entre les divers degrés de spécificité du capital humain, cette omission n'est guère étonnante.

3. PREMIERS RÉSULTATS EMPIRIQUES.

3.1) Remarques préliminaires.

Il est important de noter que l'approche pragmatique développée par Mincer découle en bonne partie du fait qu'il est impossible de trouver de bonnes mesures directes d'investissement en capital humain. Il faut donc chercher à se débrouiller du mieux qu'on peut avec l'information disponible, et souvent cette information se résume au nombre d'années qu'un individu a passé sur le marché du travail. Toutefois, même cette information n'est pas toujours disponible de sorte qu'on évalue l'expérience d'un individu par la différence entre son âge et le nombre d'années qu'il a passé à l'école. Par contre, certaines enquêtes posent des questions directes sur la formation acquise durant (et même avant) l'emploi courant. L'éventail des questions concernant la formation va de l'aspect qualitatif (et subjectif) du genre «votre emploi vous permet-il de développer de nouvelles habiletés?» à des mesures plus directes comme le nombre de semaines passées en formation. Le but de cette section est de faire l'inventaire des travaux utilisant les mesures indirectes usuelles. Les articles ayant recours aux mesures directes de l'acquisition de capital humain seront traités dans une section ultérieure.

3.2) Résultats de Mincer.

Les résultats montrés dans le tableau 1²³ nous font voir différentes spécifications de l'équation de base. On remarquera que le taux de rendement lié à la scolarité s'est croissant par rapport au nombre de variables incluses afin de tenir compte des différences d'expérience acquise sur le marché du travail. Comme l'avait noté Mincer, si l'on admet que les travailleurs investissent à des degrés divers en capital humain lorsqu'ils commencent à travailler et que les personnes moins scolarisées ont eu plus de temps pour accumuler du capital humain lié à l'emploi (c'est-à-dire de l'expérience) que les gens plus scolarisés, ces derniers pourraient très bien avoir un revenu plus faible en début de carrière comparativement à des travailleurs du même âge qui sont moins scolarisés. De sorte que si l'on veut vraiment mesurer le taux de rendement de l'éducation à partir d'un échantillon aléatoire et que l'on ne dispose d'aucune mesure de l'expérience passée sur le marché du travail, il faut à tout le moins regrouper les individus par niveaux de scolarité. Par ailleurs, dans l'éventualité où l'on possède une mesure de l'expérience, il suffit de l'inclure dans l'équation de salaire pour qu'émerge le «vrai» taux de rendement. C'est ce que nous montre le tableau 1 où, à mesure que la spécification se raffine, le taux de rendement de l'éducation s'accroît.

À noter que les résultats ne tiennent pas compte des biais possibles causés par l'hétérogénéité individuelle non observée. Pour être plus précis, en supposant que chaque individu puisse être caractérisé par un paramètre α fixe inconnu de l'économètre, s'il se trouve que ce paramètre est corrélé avec les variables explicatives dont nous voulons mesurer l'effet, les coefficients associés à ces variables s'en trouveront biaisés. Dans l'éventualité où on ne dispose que d'une coupe instantanée, ce problème ne comporte pas de solution permettant d'éliminer l'effet de α . Le mieux que l'on puisse faire, c'est de tenter d'avoir si possible un échantillon aussi homogène que possible de façon à atténuer l'inconvénient causé par ce problème de variable non observée. Il faut tout de même reconnaître que Mincer était parfaitement conscient de ce problème, d'abord dans la dérivation même de l'équation de salaire (voir équation (35)) et ensuite dans l'attention qu'il porte à l'analyse des résidus où tous

²³ Reproduit partiellement à partir de Mincer (1974), page 92.

les facteurs non observables tels que l'intensité avec laquelle les individus investissent ainsi que leur taux de rendement propre se retrouvent et ont un impact.

Afin de voir de façon plus précise de quel impact il s'agit, reprenons l'analyse de Mincer et considérons le log du revenu d'un individu à trois moments distincts dans sa carrière:

$$\begin{aligned} \ln Y_{si} &= \ln E_{si} + \ln(1-k_{0i}) \\ \sigma^2(\ln Y_s) &= \sigma^2(\ln E_s) + \sigma^2(\ln(1-k_0)) + 2\rho_1 \sigma(\ln E_s) \sigma(\ln(1-k_0)) \end{aligned} \quad (39)$$

$$\ln Y_{ji} = \ln E_{si} ; \quad \sigma^2(\ln Y_j) = \sigma^2(\ln E_s) \quad (40)$$

$$\begin{aligned} \ln Y_{pi} &= \ln E_{si} + rK_{Ti} \\ \sigma^2(\ln Y_p) &= \sigma^2(\ln E_s) + r^2 \sigma^2(K_T) + 2\rho_2 \sigma(\ln E_s) \sigma(K_T) \end{aligned} \quad (41)$$

où Y_{si} est le revenu de l'individu i avec s années de scolarité au moment où il entreprend sa carrière, Y_{ji} est son revenu au point de renversement («overtaking») et Y_{pi} est son revenu maximal. Le changement dans la variance du log du revenu dépendra alors de la dispersion en termes d'investissements ainsi que de la corrélation entre ces investissements et le revenu potentiel de départ, $\ln E_s$ (qui dépend de la scolarité). Une corrélation positive entre l'investissement K_T et $\ln E_s$ implique qu'il y a corrélation négative entre $\ln(1-k_0)$ et $\ln E_s$ ($\rho_1 < 0$). Si ρ_1 et ρ_2 sont nuls, alors le profil de variance des résidus sera en U, le minimum de la variance se produisant au moment du renversement. Le profil sera monotone décroissant si ρ_2 est suffisamment négatif alors que si ρ_2 est suffisamment positif, le profil de variance sera monotone croissant. De fait, Mincer trouve que pour les diplômés du secondaire, le profil a une forme en U. Pour ceux qui détiennent un diplôme d'études supérieures, le profil de variance sera plutôt croissant alors que pour les non diplômés du secondaire, il est décroissant. Cette question de la structure de covariance temporelle du terme d'erreur sera

examinée par Hause (1980) pour un groupe relativement homogène de jeunes Suédois faisant leur entrée sur le marché du travail. En testant les restrictions implicites suggérée par la théorie du capital humain sur la structure de covariance du terme résiduel, Hause ne peut rejeter la présence d'une forme en U du profil temporel de la variance.

Un autre facteur de biais que l'on doit prendre en considération tient au fait que lorsqu'on dispose d'un échantillon de travailleurs à différents stades dans leur carrière, il faut tenir compte du fait que le processus qui a amené ces travailleurs à occuper leur emploi actuel n'est pas exogène. En effet, par leurs activités de furetage, ces individus ont été amenés à «expérimenter» différents emplois et à effectuer plusieurs tirages à partir de la distribution des salaires. Or si l'on admet que plus un individu est expérimenté, plus il a de chance de trouver un partenaire qui lui convient, le coefficient associé à l'expérience se trouvera en fait à mesurer à la fois le rendement sur le capital humain acquis ainsi que le rendement au furetage. Sous-jacent à ce processus se trouve évidemment l'hypothèse que l'information n'est pas parfaite et qu'elle est coûteuse à acquérir. Afin de généraliser la formulation de Mincer, il nous faudra donc introduire la notion d'emploi au modèle. Cette extension permet de faire d'une pierre deux coups: d'abord, elle nous permet de tenir compte du concept de «matching» développé par Jovanovic (1979a,b et 1984), de même qu'elle nous permet d'introduire la notion de capital spécifique à la firme.

3.3) capital spécifique à la firme.

Jusqu'à présent, la distinction fondamentale faite par Becker (1962) n'a pas été prise en compte dans l'estimation de l'équation de salaire. Seule l'expérience totale accumulée sur le marché du travail a servi à mesurer l'acquisition de capital humain au courant d'une carrière. Pourtant, un fait empirique relevé par Mincer et Jovanovic (1981) est que la probabilité de quitter un emploi entre l'instant t et l'instant $t+1$ étant donné qu'on l'a occupé jusqu'à t , soit le taux de «hasard», est une fonction décroissante de l'ancienneté dans la firme. L'explication classique à ce phénomène est l'accumulation de capital humain spécifique à la firme en

question. D'ailleurs, ce lien négatif entre l'ancienneté d'une part et le taux de hasard de l'autre est tout à fait compatible avec l'hypothèse de partage de Becker. En effet, la raison pour laquelle les deux parties voulaient partager les coûts et le fruit d'un investissement en capital spécifique, selon Becker, était justement afin de réduire le roulement (pour une analyse rigoureuse de la question du partage et de son effet sur les départs et les mises à pieds, voir Parsons (1972)). Dans la mesure où le capital spécifique s'accroît avec l'ancienneté dans la firme, il s'ensuit donc qu'un travailleur accumulant de plus en plus d'habiletés d'usage plus ou moins exclusif à son emploi sera de moins en moins susceptible de le quitter. Le tableau 2 montre les résultats de l'étude de Mincer et Jovanovic. À noter l'impact d'ajouter l'ancienneté dans la firme sur le coefficient associé à l'expérience ainsi que sur le pouvoir explicatif total de la régression.

Toutefois, comme le notent les auteurs, le résultat concernant la fonction de hasard et l'équation de salaire sont susceptibles d'être dûs en partie à des effets d'hétérogénéité individuelle. En effet, les individus diffèrent probablement dans leur propension intrinsèque à changer d'emploi et si cette dimension n'est pas prise en compte, la relation entre l'ancienneté la mobilité s'en trouve vraisemblablement biaisée. De même, il est plausible que plus le pairage entre un travailleur et une firme est bon, plus le salaire du travailleur sera élevé *dès le départ*. Ceci aura comme conséquence de faire en sorte que le travailleur demeurera plus longtemps avec son employeur comparativement à un autre dont le pairage n'est pas d'aussi bonne qualité. Une régression du salaire sur l'ancienneté produira alors un lien positif en partie fortuit entre les deux. Mincer et Jovanovic tentent de corriger ce biais possible en incluant dans l'équation de mobilité ainsi que dans l'équation de salaire le nombre d'emplois préalables détenus par le travailleur. Notons que l'inclusion de cette variable affecte peu la pente du profil ancienneté-salaire pour les jeunes travailleurs alors que pour les plus âgés, il semble que des changements d'emploi fréquents représentent un indice d'«instabilité» qui fait en sorte qu'une relation entre une firme et ce travailleur soit peu propice à des investissements spécifiques. De fait, l'inclusion de cette variable dans l'équation de salaire réduit de moitié le coefficient associé à l'ancienneté. Cette question d'hétérogénéité des

pairages entre travailleurs et employeurs fut explicitement traitée de façon théorique par Jovanovic dans une série d'articles parus entre 1979 et 1984. La prochaine sous-section met surtout l'accent sur le premier de ces articles.

3.4) théorie du matching en bref.

Le modèle néo-classique standard, dont la théorie d'accumulation optimale du capital humain est un exemple, fait l'hypothèse que toute l'information pertinente est connue ou, à tout le moins, non coûteuse à acquérir. En réalité, il semble beaucoup plus plausible de supposer que a) à tout instant t , les gens ne savent pas «où» se situe l'emploi qui leur convient le mieux et qu'il leur faut donc déployer des énergies à fureter afin d'acquérir cette information, et b) même lorsqu'ils acceptent une offre d'emploi, bien souvent ils ne connaissent qu'imparfaitement les attributs de cet emploi et ce n'est qu'en «expérimentant» avec celui-ci qu'ils apprendront s'il leur convient ou non. La théorie du matching développée par Jovanovic se concentre principalement sur la question de l'apprentissage des attributs de l'emploi ou, si on préfère, de la qualité du pairage entre le travailleur et son employeur.

Dans sa version originale (Jovanovic 1979a), la théorie du matching suppose qu'au moment où une firme et un travailleur forme un match (suite à un processus de furetage de la part des deux parties, par exemple), ni l'un ni l'autre ne savent si «ça va marcher» ou non. Tout ce qu'ils savent, c'est que leur rencontre est le résultat d'un tirage provenant d'une distribution a priori qui leur est connue. Ce n'est qu'en observant l'output du travailleur dans cet emploi que les deux parties apprendront graduellement de période en période la qualité de leur match, c'est-à-dire leur emplacement dans la distribution. En juxtaposant un processus d'apprentissage faisant usage du théorème de Bayes afin de mettre à jour les croyances des deux parties concernant la qualité du match (l'information, bien qu'imparfaite, est symétrique) avec un contexte de concurrence parfaite entre les firmes, Jovanovic montre que le contrat

d'équilibre consiste pour la firme à offrir un salaire égal à la productivité espérée du travailleur étant donné que celui-ci a la stratégie optimale suivante: tant que son évaluation a posteriori de la qualité du match se situe en deçà d'un certain seuil, il demeure avec son employeur pour une autre période, sinon il quitte et effectue un nouveau tirage avec un autre employeur et ce tirage est indépendant des matches précédents.

On voit donc ici que, comme nous l'avions souligné plus haut, si on tente d'estimer l'impact de l'ancienneté sur le salaire afin de mesurer le rendement sur l'investissement spécifique à la firme, on trouvera un lien positif qui pourrait s'avérer tout à fait fortuit, puisque le fait que les gens qui ont davantage d'ancienneté aient également un salaire plus élevé ne signifie pas nécessairement que le salaire *augmente* avec l'ancienneté. Bref, il y a un effet de sélection par lequel les bons matches durent plus longtemps et sont mieux rémunérés. Dans une section ultérieure, nous verrons comment les études plus récentes ont tenté d'atténuer les biais d'estimation causés par ces effets de sélection. La prochaine section, quant à elle, s'avérera un bref retour à la théorie. En effet, deux articles empiriques de Medoff et Abraham (1980, 1981) ont soulevé des doutes quant au postulat qui veut que les travailleurs mieux payés soient nécessairement plus productifs. En effet, en examinant les dossiers de rémunération et d'évaluation de performance des employés de plusieurs firmes, Medoff et Abraham ont noté qu'il ne semblait avoir aucun lien entre les augmentations de salaires et la performance des travailleurs telles qu'évaluée par la firme. Quant à Lazear (1979 et 1981), il a, d'un point de vue théorique, montré l'existence d'un lien positif entre l'ancienneté et le salaire qui découle de considérations incitatives. L'idée de base de Lazear est que dans un contexte de risque moral du côté des travailleurs, il peut être optimal de payer les travailleurs à un niveau moindre que leur productivité marginale lorsqu'ils commencent leur séjour dans la firme et de les rémunérer à un niveau supérieur à leur productivité lorsqu'ils acquièrent de l'ancienneté. De cette façon, les débutants déploient davantage d'efforts au travail avec la promesse implicite qu'ils seront mieux payés plus tard. Il est optimal pour eux de fournir l'effort en question car plane sur leur tête la menace d'un licenciement s'ils trichent. La relation positive entre le salaire et l'ancienneté n'a alors rien à voir avec la productivité des individus.

Il s'agit donc de voir comment la théorie du capital humain peut faire face aux interrogations soulevées par ces auteurs.

4. DÉVELOPPEMENTS RÉCENTS DE LA THÉORIE DU CAPITAL HUMAIN.

En reprenant sensiblement la même structure qu'Hashimoto, Carmichael (1983) parvint à démontrer qu'il peut être optimal d'avoir un contrat prévoyant une augmentation de salaire pour les travailleurs ayant plus d'ancienneté même si la productivité de ces derniers n'augmente pas. Tout comme dans le modèle de Hashimoto, il suppose qu'après une période de formation au début de la relation, l'information n'est révélée que de façon asymétrique. Ainsi, seule la firme connaît la valeur de la productivité marginale du travailleur alors que seul ce dernier connaît son degré de satisfaction dans son emploi. On suppose en outre que l'information ne peut être échangée *ex post* par les parties en raison de coûts de transaction, ce qui implique que les décisions de séparation sont prises de façon indépendantes par la firme et le travailleur. Celui-ci quitte son emploi dans l'éventualité où son niveau de satisfaction se trouve en-deçà d'un certain seuil alors que la firme licencie le travailleur si sa productivité révélée s'avère insuffisante. L'idée de base de Carmichael est que le bien-être peut être amélioré en structurant le contrat de telle sorte que les deux parties tiennent compte («internalisent») des pertes totales de bien-être résultant d'une séparation plutôt que de simplement tenir compte de leur part respective du rendement comme dans le modèle d'Hashimoto, et ce peu importe lequel des deux partenaires provoque la séparation. Afin d'atteindre un tel optimum contraint, Carmichael modifie le modèle d'Hashimoto en incorporant des promotions basées sur le niveau d'ancienneté dans la firme.

Supposons qu'il y ait deux types d'emploi, les emplois de types 1 et les emplois de types 2. La productivité des travailleurs est identique dans les deux types d'emploi. Suite aux décisions de séparations faites au début de la deuxième période, tous les travailleurs encore à l'emploi de la firme occupent des emplois de type 1 où ils reçoivent un salaire w_2 . Plus tard

dans la deuxième période, certains de ces travailleurs sont promus à des emplois de type 2 sur la base de leur ancienneté et ils reçoivent un salaire égal à w_2+B où B représente un boni déterminé *ex ante* au moment de la signature du contrat. Par hypothèse, il y a un nombre fixe d'emplois de type 2. Par conséquent, une mise à pied par la firme ne lui fait épargner que w_2 puisqu'un travailleur junior comblera le poste de type 2 nouvellement libéré. L'incitation pour la firme à congédier un travailleur dans un emploi de type 2 s'en trouve donc réduite. En fait, les travailleurs juniors agissent comme tierce-partie au contrat en étant les bénéficiaires des congédiements de travailleurs seniors puisque ces licenciements augmentent leurs chances d'obtenir un poste de type deux. Le salaire de deuxième période étant déterminé, celui de première période est ajusté de façon à ce que la valeur espérée du contrat soit égale à ce que peut obtenir le travailleur sur le marché. Carmichael montre que ce système de promotions basées sur l'ancienneté rend les décisions de séparations plus efficaces que dans le contrat de Hashimoto.

Deux implications du modèle de Carmichael valent la peine d'être soulignées. D'abord, les travailleurs plus anciens vont typiquement recevoir un salaire plus grand que la valeur de leur productivité marginale. Donc, à cet égard, la théorie du capital humain est en mesure de rationaliser les résultats de Medoff et Abraham et de fournir une explication concurrente à celle de Lazear. Deuxièmement, bien que Carmichael ne considère que l'investissement en capital humain spécifique à la firme, son contrat optimal implique que les travailleurs vont payer la totalité du coût de l'investissement par des salaires de départ plus bas. Notons enfin que Carmichael fait l'hypothèse implicite que la firme peut s'engager à respecter le contrat dérivé *ex ante*. Cette hypothèse est importante car elle implique que la firme ne voudra pas changer le nombre d'emplois de type 2 bien que cela serait dans son intérêt de le faire dès qu'un travailleur junior s'apprête à être promu.

En considérant de façon explicite la possibilité qu'il y ait renégociation, MacLeod et Malcomson (1993a,b) montrent que les firmes effectuant des investissements spécifiques n'ont pas à offrir de salaires supérieurs à ceux du marché à leur travailleurs. Comme il est

BLSH

impossible ou trop coûteux d'écrire des contrats complets *ex ante* qui seraient robustes à la renégociation, il y aura des situations où les deux parties voudront renégocier leur contrat. Cela se produira uniquement dans le cas où l'une des deux parties préférerait répudier le contrat bien qu'il serait efficace que la relation se poursuive. Le résultat sera que l'autre partie consentira à céder une fraction de son surplus afin d'inciter son partenaire à ne pas répudier le contrat.

De façon plus précise, supposons qu'une firme et un travailleur forment un partenariat en première période et que la firme effectue des investissements spécifiques qui affecteront les gains de deuxième période alors que le travailleur ne procède qu'à des investissements de nature générale. Des chocs aléatoires affectant la productivité du travailleur dans la firme ainsi que son salaire sur le marché sont dévoilés aux deux parties à la fin de la première période. Par exemple, si la valeur marchande du travailleur s'avère plus grande que le salaire négocié au début de la relation, la firme voudra ajuster le salaire afin de l'égaliser à la valeur marchande du travailleur, en supposant bien sûr qu'il soit efficace que la relation se poursuive. De manière tout à fait identique et étant donné que la relation d'échange est efficace, le travailleur acceptera une baisse de salaire afin d'éviter la séparation si la firme préférerait le congédier plutôt que le payer au taux convenu dans le contrat. Naturellement, il y aura des situations où ni la firme ni le travailleur ne voudront renégocier. Dans de tels cas, les deux parties préfèrent échanger au taux fixé par le contrat et ce taux peut impliquer qu'il y a partage du rendement de l'investissement de la firme. Toutefois, il est intéressant de noter que le modèle prédit que dans l'éventualité où les deux parties signent un contrat qui assure que les décisions d'investissement seront efficaces (c'est-à-dire un contrat qui ne souffre pas du problème de «holdup» (voir Williamson et al (1975)), le travailleur recevra un salaire égal à sa valeur marchande, ce qui implique que le salaire n'augmente pas avec l'ancienneté. Par ailleurs, si les deux parties procèdent à des investissements spécifiques (et non plus seulement la firme), MacLeod et Malcomson démontrent que si le travailleur accepte les termes du contrat efficace, alors il y aura un lien positif entre l'ancienneté et le salaire.

5. RÉSULTATS RÉCENTS.

5.1 Remarques préliminaires.

Jusqu'à présent nous n'avons fait mention que des travaux empiriques de Mincer (1974) ainsi que de Mincer et Jovanovic (1981). Bien d'autres auteurs ont utilisé une formulation à la Mincer. Signalons, par exemple, Borjas (1981) ainsi que Bartel et Borjas (1981), qui s'attardaient à étudier les questions de l'impact de la mobilité sur l'évolution salariale. De même, Mincer et Higuchi (1988) utilisèrent cette formulation afin de comparer les structures salariales et les différences quant à la mobilité des travailleurs au Japon et aux États-Unis. Toutefois, afin d'abrégier la présentation, nous devons passer outre ces contributions et nous concentrer sur les articles qui nous semblent les plus significatifs, à la fois sur le plan méthodologique et sur le plan des résultats. Dans cette section, notre attention se portera sur trois articles parus au cours des dernières années, soit ceux d'Abraham et Farber (1987), Altonji et Shakotko (1987) ainsi que Topel (1991).

5.2 Le modèle de base et les résultats d'estimation.

Les trois premiers articles sus-mentionnés prennent comme point de départ l'équation suivante:

$$\ln w_{ijt} = \beta_1 \text{Expérience} + \beta_2 \text{Ancienneté} + \alpha_i + \gamma_{ij} + \epsilon_{ijt} \quad (42)$$

où α_i est un terme d'erreur spécifique à l'individu représentant la dimension non observée par l'économètre des attributs du travailleur, γ_{ij} est une composante spécifique au pairage de l'individu i avec l'emploi j , représentant la qualité non-observée du match alors que ϵ_{ijt} est un bruit blanc résiduel. Toutes les autres variables explicatives, y compris les termes quadratiques pour l'expérience et l'ancienneté sont omis afin d'alléger la présentation.

Les problèmes de biais d'estimation découlent de la corrélation entre les variables observées et les composantes d'erreur non-observées, plus particulièrement le terme reflétant la qualité du match, γ_{ij} . Par exemple, la qualité du match découlant en partie d'un processus de furetage, il est raisonnable de supposer que plus les gens ont accumulé d'expérience, plus ils ont eu l'occasion de localiser de meilleurs matches. Le taux de rendement lié à l'expérience, β_2 , comprendra alors en fait un rendement au furetage en plus d'un rendement sur le capital humain général. Une façon de remédier à ce problème serait d'avoir une variable observable qui s'avérerait un bon substitut pour α_i et γ_{ij} . C'est cette approche qui fut retenue par Abraham et Farber. En exploitant l'aspect longitudinal de leur données, ils connaissent la durée ultime des emplois occupés par l'individu, sauf celui qui est encore en cours au moment de la dernière entrevue. Abraham et Farber adoptent donc la spécification suivante:

$$\ln w_{ijt} = \beta_1 \text{Expérience} + \beta_2 \text{Ancienneté} + T_{ij}^u \theta + \epsilon_{ijt} \quad (43)$$

où T_{ij}^u représente la durée ultime de l'emploi j . L'idée est que plus le match est de bonne qualité et plus l'individu est «habile», plus longue sera la durée du match. Quant au dernier emploi occupé par l'individu (qui s'avère souvent l'emploi le plus longtemps occupé), Abraham et Farber estiment sa longévité par un modèle de durée. Au moment de la dernière entrevue, on sait déjà que l'individu l'a occupé pendant t années. Il s'agit d'estimer la durée résiduelle de cet emploi et d'ajouter cet estimé à t pour obtenir T_{ij}^u .

L'approche retenue par Altonji et Shakotko fut plutôt de tenter d'éliminer la corrélation entre γ_{ij} et l'ancienneté en utilisant la méthode des variables instrumentales. S'il était possible de trouver un bon instrument pour l'ancienneté qui aurait en plus la propriété de ne pas être corrélé avec la composante de pairage, le problème de biais de β_2 serait résolu. Dans la mesure où l'effet de pairage est vraiment fixe, l'instrument retenu par Altonji et Shakotko pour l'ancienneté dans l'emploi j , soit la déviation de l'ancienneté par rapport à sa durée moyenne dans cet emploi, s'avère effectivement posséder tous les attributs souhaitables. En effet, il est fortement corrélé avec l'ancienneté tout en étant, par construction, orthogonal à γ_{ij} . Notons ici que les auteurs ne se préoccupent pas de la corrélation possible entre α_i et

l'expérience et/ou l'ancienneté. Si l'on admet que les individus mieux pourvus sont moins susceptibles d'avoir vu leur carrière interrompue par des périodes de chômage, à ce moment il y aura vraisemblablement corrélation entre l'expérience totale accumulée sur le marché et α_i . De même, si davantage d'expérience implique davantage d'occasions de fureter pour trouver un meilleur emploi, alors, par ricochet, α_i sera corrélé avec l'ancienneté.

Topel, par contre, choisit d'éliminer les effets fixes en utilisant une procédure à deux étapes.²⁴ Il obtient d'abord un estimé convergent de la somme de β_1 et β_2 en utilisant les premières différences de l'équation (43) à l'intérieur d'un même emploi:

$$\ln w_{ijt} - \ln w_{ijt-1} = (\beta_1 + \beta_2) + \Delta(\text{autres variables}) + \Delta \epsilon_{ijt} \quad (44)$$

Notons que les effets fixes sont éliminés et qu'il ne nous reste qu'une différence de bruits blancs. Ensuite, Topel estime une équation en niveau de façon à obtenir un estimé de β_2 . Soit X_0 , l'expérience initiale au moment de commencer un emploi, on aura $X_0 = \text{expérience-ancienneté}$. De plus, appelons b l'estimateur convergent obtenu en première étape. On a alors:

$$\ln w_{ijt} - b(\text{Ancienneté}) = X_0 \beta_2 + e_{ijt} \quad (45)$$

Comme le β_2 estimé sera biaisé à la hausse, la différence entre b et l'estimé de β_2 représentera une borne inférieure pour β_1 , le rendement lié à l'ancienneté (c'est-à-dire au capital humain spécifique).

Les résultats de ces trois études sont présentés au tableau 3. Notons que tous ces articles ont utilisé la même banque de données, ce qui rend ces estimés directement comparables.

On peut noter la différence entre, d'une part, les résultats de Topel et, d'autres part, ceux d'Abraham et Farber et d'Altonji et Shakotko. Toutefois, comme le souligne Topel, les

²⁴ Il utilise également une spécification faisant appel à des termes d'ordre 4 pour l'expérience et l'ancienneté afin d'avoir une forme fonctionnelle davantage conforme aux profils expérience-salaires empiriques.

données sur l'ancienneté contenues dans l'échantillon retenu du Panel Study of Income Dynamics (PSID) sont truffées d'erreurs de mesure. Il utilise donc une variable d'ancienneté corrigée, différente de celle utilisée par les deux autres articles. Du moment qu'il utilise l'ancienneté corrigée avec la méthodologie d'Altonji et Shakotko, il obtient sensiblement les mêmes résultats que dans son article. En ce qui concerne les différences entre Abraham et Farber et Topel, ce dernier note que l'équation (43) estimée par Abraham et Farber peut se réécrire comme suit:

$$\ln w = X_0\beta_2 + (T-T^*)(\beta_1+\beta_2) + T^*(\beta_1+\beta_2) + T^\theta + \xi \quad (46)$$

où T représente l'ancienneté dans l'emploi courant et T^* est sa durée de vie moyenne. Or, estimer l'équation (43) par les moindres carrés est équivalent à estimer l'équation (46) en imposant que les coefficients associés à $T - T^*$ et à T^* soient identiques. Comme X_0 et T^* sont fixes à l'intérieur d'un emploi, $T - T^*$ leur est orthogonal, par conséquent si on estime l'équation (46) sans imposer de restriction sur les coefficients, on obtiendra un estimé convergent de $\beta_1 + \beta_2$. Toutefois, le processus de sélection implique que des variables comme T sont corrélées avec les composantes d'erreur. Par conséquent, le coefficient associé à T sera biaisé. Donc, imposer la restriction mentionnée plus haut implique qu'on obtient un estimé biaisé de $\beta_1 + \beta_2$.

6. MESURES DIRECTES DE L'ACQUISITION DE CAPITAL HUMAIN.

6.1 Remarques préliminaires.

Rarissimes il y a peine cinq ou six ans, les études faisant usage de banques de données comportant des mesures directes du capital humain ou, plus exactement, des mesures directes quant à la formation reçue avec l'employeur actuel ou dans des emplois antérieurs, tendent à se multiplier ces jours-ci. Vouloir faire une recension de toutes ces contributions s'avérerait un projet digne d'un article en soi. Aussi, de manière à cerner l'essentiel de cette production,

nous nous contenterons de mettre en relief trois aspects de la formation reçue en milieu de travail.

- (i) dans quelle mesure la formation augmente-t-elle la productivité et le salaire du travailleur?
- (ii) la formation reçue s'avère-t-elle transférable d'un employeur à un autre?
- (iii) les travailleurs défraient-ils indirectement les coûts de leurs programmes de formation en acceptant des salaires de départ plus faibles?

Par ailleurs, comme l'éventail des mesures du capital humain contenues dans ces banques de données en ce qui concerne la formation est assez disparate, il serait sans doute utile d'en faire l'inventaire en identifiant le ou les articles faisant usage de la banque de données correspondante (voir tableau 4).

- (i) Impact de la formation sur le salaire et la productivité.

Comme tous les articles étudiés utilisent une spécification à la Mincer, l'impact de la formation sur le salaire de l'employé est comparable d'une étude à l'autre bien que comme on a pu le voir ci-haut, les questions concernant la formation étaient assez hétérogènes. En général, l'effet de la formation sur le salaire avec l'employeur actuel est plutôt substantiel et statistiquement significatif, comme on peut le voir au tableau 5. À noter toutefois les travaux de Lynch (1992) et Blanchflower et Lynch (1994) qui utilisent les données du NLSY. Cette banque de données contient les épisodes de formation pour chaque répondant et ce, par catégorie de formation, de sorte qu'il est possible de reconstruire le temps cumulé passé à être formé que ce soit avec l'employeur courant ou avec les employeurs précédents. Avec des échantillons assez similaires (y sont exclus ceux qui obtiennent un diplôme d'études collégiales ou universitaires), les résultats sont plutôt tièdes en faveur d'un impact positif de la formation sur le salaire.

En revanche, Duncan et Hoffman (1979) de même que Brown (1989), avec la même banque de données (PSID), trouvent un impact positif important de la formation, surtout dans le cas de Brown qui porte une attention spéciale aux biais possibles pouvant être causés par la simple formulation des questions. Ainsi, pour la spécification qu'il préfère, il trouve un taux de rendement de plus de 11%. De plus, Brown ainsi que Duncan et Hoffman constatent que le rendement à l'ancienneté est étroitement associé aux épisodes de formation. Brown en conclut donc que l'effet d'ancienneté découle principalement de considérations liées à l'accumulation de capital humain et n'a que peu à voir avec les théories concurrentes à la Lazear qui voient dans le lien positif entre l'ancienneté et le salaire un phénomène contractuel lié aux asymétries d'information.

Bartel (1992), quant à elle, a examiné les dossiers de personnel d'une grande entreprise manufacturière pour la période allant de 1986 à 1990. Cet échantillon longitudinal lui permettant d'exercer un certain contrôle quant aux biais d'hétérogénéité, elle trouve elle aussi un impact important de l'ordre d'au moins 13%. De plus, le fait d'introduire l'information concernant les jours passés en formation réduit le rendement lié à l'ancienneté de 18%, montrant par là qu'il y a réellement un lien entre l'acquisition d'habileté et l'ancienneté. Évidemment, bien que cette banque de données comporte certains avantages du point de vue des erreurs de mesures, il n'en reste pas moins qu'il ne s'agit que d'une seule firme et qu'il est plutôt délicat de tirer des conclusions générales à partir de ce qui n'est en somme qu'une seule observation.

Lillard et Tan (1992), dans leur étude exhaustive (ils ont fait usage de trois banques de données), en sont également venus à la conclusion que la formation formelle dispensée par la compagnie («company provided training») est un facteur important de croissance salariale avec un impact de plus de 15%. Un tel résultat est en opposition directe avec Lynch ainsi que Blanchflower et Lynch dont les données mesuraient également l'aspect formel des programmes de formation.

En ce qui concerne l'impact sur la productivité, Bishop (1994) et Barron, Berger et Black (1993) trouvent que la formation dispensée par la compagnie, qu'elle soit de nature formelle ou informelle, a généralement un impact important. Bien que Bishop s'attarde surtout à mesurer l'effet de la formation acquise avec un ou des employeurs précédents, un sujet sur lequel nous reviendrons plus loin, les deux articles estiment (avec des données en partie identiques) qu'en général l'effet sur la productivité est beaucoup plus important que l'effet sur le salaire. Ainsi, Barron, Berger et Black soutiennent que la formation a environ dix fois plus d'impact sur la productivité que sur le salaire, et ce résultat est valable pour les deux banques de données dont ils ont fait usage. Les auteurs estiment que ce résultat est plutôt surprenant même si, théoriquement, rien n'empêche que la part du rendement qui va au travailleur soit très faible si le capital humain produit par l'investissement est entièrement spécifique à la firme. Dans le modèle d'Hashimoto, par exemple, si le degré d'incertitude concernant la valeur sur le marché du capital humain acquis avec l'employeur présent est nul, alors la part d'équilibre du travailleur sera de zéro. Toutefois, Barron, Berger et Black estiment que le taux de roulement élevé que l'on observe dans l'économie américaine devrait faire en sorte que les travailleurs partagent une partie des coûts et du rendement.

Notons que les mesures de productivité utilisées par Bishop ainsi que par Barron, Berger et Black sont construites à partir de réponses données par les employeurs aux questions suivantes:

"Please rate your employee on a productivity scale of zero to 100, where 100 equals the maximum productivity rating any of employees [in this] position can attain and zero is absolutely no productivity by your employee. What is the productivity of [the last worker hired] during (his/her) first two weeks of employment?"-EOPP.

"Please rate (name of last worker hired) on a productivity scale of zero to 100, where 100 equals (name's) productivity when (he/she) is fully trained and

zero is absolutely no productivity by (name)...What was (name's) productivity on this scale during his/her first two weeks of employment?"-SBA.

Les auteurs ont ensuite construit un indice salarial des employés tout juste embauchés par rapport au salaire des travailleurs avec deux ans d'expérience dans cet emploi (en supposant que ces derniers soient pleinement qualifiés après deux ans). Ce sont ces deux indices que les auteurs comparent ensuite pour en arriver à leurs conclusions.

Bien qu'un certain consensus se dégage à l'effet que la formation telle que mesurée dans ces diverses banques de données ait un impact positif sur le salaire, on peut toutefois noter des divergences assez importantes, comme le montre une comparaison de l'étude de Brown et celle de Barron, Berger et Black, par exemple. Deux explications plausibles s'offrent pour justifier de telles divergences, les erreurs de mesures ainsi que la présence d'hétérogénéité non observée. Étant donnée la nature très subjective de certaines questions (pensons surtout aux indices de productivité), il est possible que les mêmes questions concernant les mêmes personnes mais posées à des superviseurs différents auraient produit des résultats différents. Il est même possible que les mêmes questions posées aux mêmes personnes mais à une date différente n'auraient pas amenées les mêmes réponses. Étant donné la taille réduite des échantillons étudiés (moins de 1,000 répondants pour Barron et al. avec le SBA, un peu plus de mille avec le EOPP), quelques dizaines de répondants supplémentaires pourraient avoir un impact sur les conclusions.

Par exemple, lorsque Brown passe d'une spécification où aucun correctif n'est apporté pour tenir compte des biais possibles dus aux erreurs de mesures liées aux réponses à la spécification qu'il préfère, le coefficient du terme linéaire de formation (sa spécification de base est quadratique) passe de 2% à 27%.

Pour ce qui est de l'hétérogénéité non observée, notons que lorsque Lynch (NLSY) passe d'une estimation simple pour une coupe transversale à une estimation en première

différence, l'impact positif qu'elle trouvait pour la formation formelle avec l'employeur actuel disparaît complètement. Lorsque Blanchflower et Lynch font la même chose en utilisant un échantillon quelque peu différent du NLSY (encore que similaire eu égard à la scolarité puisque les deux études se concentrent sur les jeunes non diplômés d'institutions collégiales ou universitaires), l'effet du «company training» s'accroît de 8 à 12 % et devient à toutes fins utiles significatif!

Cette sensibilité apparente des résultats est assez troublante en soi car elle suggère d'une certaine façon qu'avant de tirer des conclusions solides sur la population des travailleurs en formation, il faudra accumuler d'autres résultats provenant d'études faisant usage d'un nombre beaucoup plus grand d'observations afin que les diverses erreurs de mesure qui sont potentiellement présentes aient une chance de s'annuler, si bien sûr ces erreurs de mesures sont «classiques» au sens statistique du terme.

(ii) La formation est-elle transférable d'un employeur à un autre?

Les conclusions générales à ce sujet sont qu'à tout le moins, la formation reçue avec des employeurs précédents semble avoir un impact très important sur la productivité actuelle (voir tableau 6), comme le souligne Bishop:

"Formal training received on the job from a previous employer has no effect on starting wage but increases initial productivity by 9.5 percent of the wage and reduces training requirements by 17.3 percent...Formal training received off the job [...] has no initial effect on anything, but it increases the index of suggestions by 37 percent and current productivity by 15.9 percent. Formal off-the-job training does not increase current wage rates however, so

profitability increased by 13.8 percent of the wage at six months of tenure and by 18.6 percent of the wage at the time of the interview."²⁵

Les résultats de Bishop quant à l'impact salarial de la formation «off-the-job» est en contradiction avec ceux de Lynch qui trouve un impact positif d'environ 10% et ce, même en prenant les premières différences. Toutefois, Blanchflower et Lynch trouvent un effet plus faible de l'ordre de 4-5% à peine significatif. Quant à la formation de nature formelle de type «on-the-job», les deux études utilisant le NLSY concluent à l'absence totale d'impact sur le salaire.

Lillard et Tan, quant à eux, trouvent des effets importants avec les données du CPS. Ainsi, la formation de nature informelle accumulée avec des employeurs précédents est associée avec une prime salariale de 17.6% dans l'emploi actuel alors que la formation plutôt formelle suivie dans ces emplois se transforme en une prime de près de 20%. À noter que ces deux types de formation sont de nature «on-the-job» ou sur le tas.

Encore ici, on peut se rendre compte de la divergence des résultats quant à l'impact salarial de la formation préalable à l'emploi. Cette divergence est d'autant plus grande qu'en ce qui concerne Lynch et Lillard et Tan, notamment, les types de formation étudiées rendent en principe les résultats comparables.

(iii) Dans quelle mesure les travailleurs paient-ils leurs programmes de formation par des salaires de départ inférieurs?

Ici, le constat est simple: personne n'a trouvé d'effet négatif économiquement significatif. Comme on peut le constater au tableau 7, seuls Barron, Berger et Black (1993)

²⁵ Lorsque Bishop fait mention de la productivité actuelle («current»), il fait référence à la productivité au moment de l'entrevue. Donc, bien que la formation n'ait aucun impact initial, elle a tout de même ultérieurement un impact important.

ont réussi à isoler un effet négatif (et statistiquement significatif) de la formation sur le salaire de départ, mais cet effet est très faible. Bien qu'ils admettent que des problèmes d'hétérogénéité non observée puissent masquer en partie l'effet réel (l'idée étant que les gens plus habiles sont davantage susceptibles d'être à la fois formés et d'avoir un salaire de départ plus élevé, ce dernier effet annulant en tout ou en partie l'effet négatif de la formation), ils en concluent même que l'absence d'un tel effet peut refléter la présence de marchés duals. Dans les «bons» emplois, il y aurait à la fois des salaires plus élevés et plus de formation, alors que les «mauvais» emplois seraient caractérisés par des salaires faibles et peu de formation.

7. CONCLUSION

Quelques questions ouvertes demeurent en ce qui concerne la vérification empirique de la théorie du capital humain. Notons d'abord que très peu d'attention a jusqu'ici été portée sur le degré de spécificité du capital par rapport à l'industrie dans laquelle opère l'entreprise. Cette question n'est pas sans intérêt car pratiquement tous les articles qui, jusqu'à présent, se sont penchés sur la question de déterminer si il y a un effet d'ancienneté dans la firme ou non, ont vraisemblablement surestimé la spécificité par rapport à la firme du capital humain accumulé. Or, s'il advenait que l'effet d'ancienneté (on parle bien sûr d'effet sur le salaire) soit négligeable ou même nul, alors toute la notion de partage de rente qui fait presque figure de postulat serait remise en question.

Par ailleurs, toutes les études faisant usage des mesures indirectes d'accumulation du capital humain ont fait l'hypothèse que lorsqu'il y a un problème d'hétérogénéité non observée, toute l'hétérogénéité se retrouve dans l'ordonnée à l'origine des profils de salaire. Or, il se peut très bien que les individus diffèrent également dans les pentes des profils. Si oui, on devrait observer une variance non nulle de ces pentes. Cette avenue n'a jamais été vraiment exploré (à l'exception de Hause (1980)): tout au plus les auteurs se bornent à estimer l'équation de

salaires par classes d'occupation ou par industries. Pourtant, la dérivation de l'équation de salaire par Mincer fait explicitement référence à l'hétérogénéité dans les pentes (ou taux de rendement).

Par ailleurs, à part Lynch (1991) et Mincer (1988) avec des données limitées, la question même du degré de spécificité du capital humain a été peu étudiée. Pour être plus précis, dans quelle mesure la formation reçue dans un emploi est-elle un facteur positif ou négatif de mobilité? Les problèmes causés par l'hétérogénéité non observée requièrent l'utilisation de données longitudinales afin de pouvoir y apporter certains correctifs, sinon il sera toujours difficile de déterminer si les effets mesurés sont présents en raison de variables qu'on ne peut mesurer et qui sont corrélées avec les variables qui nous intéressent.²⁶

Enfin, notons que tous les efforts déployés afin de montrer que les travailleurs paient au moins en partie leurs programmes de formation par un salaire de départ plus faible ont été pratiquement vains jusqu'à présent. Encore là, étant donné la variabilité relativement grande des résultats faisant usage de mesures directes de la formation, il faudra accumuler davantage d'études afin de pouvoir avancer des conclusions définitives sur ce sujet.

²⁶ En fait, Meyer (1990), dans le cadre d'un modèle de durée a développé une méthodologie afin de tenir compte de l'hétérogénéité non observée et ce, même si on utilise des données en coupe instantanée. À noter toutefois que, par hypothèse, il suppose l'absence de corrélation entre le paramètre d'hétérogénéité et les variables explicatives.

Tableau 1. Premiers résultats de Mincer.

Formes fonctionnelles	R carré
$\ln Y = 7.58 + .070(S)$	0.067
$\ln Y = 6.20 + .107(S) + .081(X) - .0012(X \text{ carré})$	0.285
$\ln Y = 4.87 + .255(S) - .0029(S \text{ carré}) - .0043(XS) + .148(X) - .0018(X \text{ carré})$	0.309

Notes- Expérience=X; scolarité=S; Revenu annuel d'hommes américains de race blanche en 1959 (exluant les agriculteurs)=Y.

Tableau 2. Résultats de Mincer et Jovanovic (1981).

Formes fonctionnelles	R carré
$\ln W = .2437 + .0741(S) + .0467(X) - .0007(X \text{ carré})$	0.263
$\ln W = .2351 + .0732(S) + .0372(X) - .0006(X \text{ carré}) + .0305(A) - .0007(A \text{ carré})$	0.295

Notes- Expérience=X; scolarité=S; ancienneté=A; salaire réel (W) d'hommes américains en 1975-76 échantillonnés à partir du «Panel Study of Income Dynamics de l'Université du Michigan».

Tableau 3. Résultats comparés d'Altonji et Shakotko (1987),
Abraham et Farber (1987) et Topel (1991).
(Variable dépendante: log du salaire réel)

Auteurs	Expérience	Ancienneté
Altonji & Shakotko	0.0589 (0.0044)	-0.0043* (0.0019)
Abraham & Farber	0.0397 (0.0044)	* 0.006 (0.0019)
Topel	0.0713 (0.0181)	0.0545 (0.0079)

Notes-Les données proviennent du PSID et couvrent la période allant de 1968 à 1983. Les écarts-types sont indiqués entre parenthèses. Tous les termes d'ordre supérieur à un sont omis.

* Notez qu'Altonji et Shakotko estiment que l'effet d'ancienneté est d'environ 4.8% dans la première année mais est pratiquement nul ou légèrement négatif par après.

Tableau 4. Répertoire des banques de données contenant des mesures explicites du capital humain.

Banque de données	Mesure du capital humain ou questions s'y rapportant	Auteurs
Panel Study of Income Dynamics (PSID)	<p>(i) "On a job like yours, how long would it take an average person to become fully qualified?"</p> <p>(ii) "Are you learning new skills on the current job which could lead to a better job or promotion?"</p>	<p>Brown (1989) Ducan et Hoffman (1979) Mincer (1988)</p>
National Longitudinal Surveys (NLS)	<p>(i) "Do you receive or use additional training (other than schooling training) on your job?"</p> <p>(ii) "What was the longest type of training you have received since the last interview?"</p>	<p>Lillard et Tan (1992) Parsons (1989)</p>
Small Business Administration (SBA)	<p>"... We asked detailed information about the training this worker received in the first three months of employment... There are five measures of training: off-site formal training programs, formal training programs offered by the firm on site, informal training by the worker's supervisor, informal training by co-workers, and time that the worker spent watching others perform tasks. In addition, we asked ... "How many weeks does it take a new employee hired for (name's) type of position to become fully trained and qualified if he or she has no previous experience on this job...?"</p>	<p>Barron, Berger et Black (1993)</p>

(suite du tableau 4)

Employment Opportunity
Pilot Program, Individual
and Employer Survey (EOPP)

(i) "Describe up to 4 training events occurring
between 1/1/79 and the interview date in 1980?"

Barron, Black et
Lowenstein (1989)
Barron, Berger
et Black (1993)
Bishop (1994)
Lillard et Tan (1992)

(ii) "Number of hours typically spent by a new
employee in the position last filled watching
other people doing the job rather than doing
it himself during the first three months of
employment?"

(iii) "Number of hours a new employee in the
position spends in formal training?"

Levine (1993)

(i) "My job makes me keep LEARNING NEW THINGS",
(1="Strongly disagree", 5="Strongly agree").

(ii) "the TIME REQUIRED TO TRAIN someone
for the job" (0="A few hours", 6="5 years or more").

(iii) "the IMPORTANCE OF FORMAL TRAINING in this
company as a source of skills (same codes as formal training)."

(iv) "...At the establishment level the OJT measure
is ORIENTATION, a dummy equal to one if the firm has
an orientation procedure for new hires." 2

Lynch (1991 et 1992)
Blanchflower
et Lynch (1994)

(i) "In addition to your schooling, military and
government-sponsored training programs, did you
receive any other types of training for more than one month?"

(ii) "Which category best describes where you
received this training?" 3

National Longitudinal
Survey of Youth (NLSY)

Enquête de 1982-83 portant
sur des entreprises
manufacturières des régions
d'Indianapolis aux États-Unis
et d'Atsugi au Japon.

(suite du tableau 4)

Current Population
Survey (CPS)

Lillard et Tan (1992)

"What training was needed to get the current or last job and what training is needed to improve skills on the current job?"

National Federation of
Independent Business Survey
(NFIB)

Bishop (1994)

"... Please think of the last person hired for this job (job X) by you prior to August 1986 regardless of whether that person is still employed by your firm. Call this individual person A. The individual hired for job X immediately before person A is called person B. Do not include rehires of former employees..."

(i) "How many hours did you or an employee spend training or closely supervising A or B?"

(ii) "How many additional hours [...] did A/B spend learning the job by watching others rather than doing it?"

(iii) "How many hours did A/B spend reading manuals, etc., in order to learn the job?"

58

1. Extrait de Barron, Berger et Black (1993), page 5. À noter que la mesure commune aux cinq types de formation est le nombre total d'heures passées en formation ainsi que le nombre d'heures par semaine passées en formation.

2. Extrait de Levine (1993), page 729. OJT signifie «on-the-job training».

3. On demandait aux répondants de spécifier la ou les dates de début des programmes de formation suivis de même que la ou les dates où ils prenaient fin, et ce jusqu'à concurrence de trois ou quatre programmes selon l'année de l'entrevue.

BLSH

Tableau 5. Effet de la formation sur le salaire et la productivité.

Banque de données	Auteur(s)	Variable dépendante	Effet (par type ou source de formation s'il y a lieu)
PSID	Duncan et Hoffman(1979)	Log du revenu horaire	5 à 8%
	Mincer(1988)	Croissance du salaire réel	4 à 6%
	Brown(1989)	Log du salaire horaire	18.2% terme linéaire -0.024 terme au carré
NLS-Young Men	Lillard et Tan(1992)	Log du revenu annuel ou du salaire hebdo. (selon l'année)	11.9%-École technique 16.9%-Fourni par cie 16.6%-Gestion 14.2%-Technique ou profess 9.6%-Semi-qualifié
EOPP	Barron, Black et Loewenstein (1989)	Croissance du salaire réel	1.5%
		Taux de crois. de la productivité	3%
	Barron, Berger et Black(1993)	Log du salaire d'un nouvel employé sur le salaire d'un employé après 2 ans	1%
		Taux de crois. de la productivité	5%
SBA	Barron, Berger et Black(1993)	Log du salaire d'un nouvel employé sur le salaire d'un employé après 2 ans	1%
		Taux de crois. de la productivité	5%

BLSB

(suite du tableau 5-Effet de la formation sur le salaire...)

CPS	Lillard et Tan(1992)	Log du revenu annuel	27%-formel, fourni par la cie 5.6%-informel (sur le tas)
NLSY	Lynch(1992)	Log du salaire	10.4%-formel, fourni par la cie (non signif.) 5à10%-cours d'apprenti (non significatif) -1%-"Off-the-job" (non significatif) 10%-"Off-the-job" (effets fixes, signif.)
	Blanchflower et Lynch(1994)	Log du salaire réel	8.2%-formel, fourni par la cie (non signif.) 6%-cours d'apprenti (non significatif) -2%-"Off-the-job" (non significatif)
Enquêtes portant sur des usines américaines et japonaises			
	Levine (1993)	Log du revenu annuel	+2%+- "Time required train"(TRT) 1.8%(USA),0%(Japon)- "Learning new things" 6.6%(USA), 0%(Japon)- Interaction TRT et ancienneté>=3 ans.
Données d'une entreprise manufacturière américaines			
	Bartel (1992)	Croissance du salaire	13%

Notes- Le niveau déterminant si un résultat est significatif ou non est fixé à 5%.

Tableau 6. Effet de la formation préalable à l'emploi sur le salaire et la productivité.

Banque de données	Auteur(s)	Variable dépendante	Effet (par type ou source de formation s'il y a lieu)
CPS	Lillard et Tan (1992)	Log du revenu annuel	17.6%-Formel, fourni par la cie. 19.8%-informel (sur le tas)
NFIB	Bishop (1994)	Indice du salaire de départ	0%-formel, fourni par la cie. 0%-«Off-the-job»
		Indice de productivité initiale	9.5%-formel, fourni par la cie. 0%-«Off-the-job»
		Indice de productivité au moment de l'entrevue	0%-formel, fourni par la cie. 15.9%-«Off-the-job»
EOPP	Bishop (1994)	Indice du salaire de départ	4%-Institutions privées.
		Indice de productivité initiale	20%-Institutions privées.
		Indice de productivité après 3 mois net des coûts de formation	20%-Institutions privées.

BLON

(suite du tableau 6-Effet de la formation préalable à l'emploi...)

NLSY	Lynch(1992)	Log du salaire	0%-formel, fourni par la cie. 21%-cours d'apprenti 10.4%«Off-the-job»
	Blanchflower et Lynch(1994)	Log du salaire réel	-3%-formel, fourni par la cie (non signif.) 19%-cours d'apprenti 4%«Off-the-job»

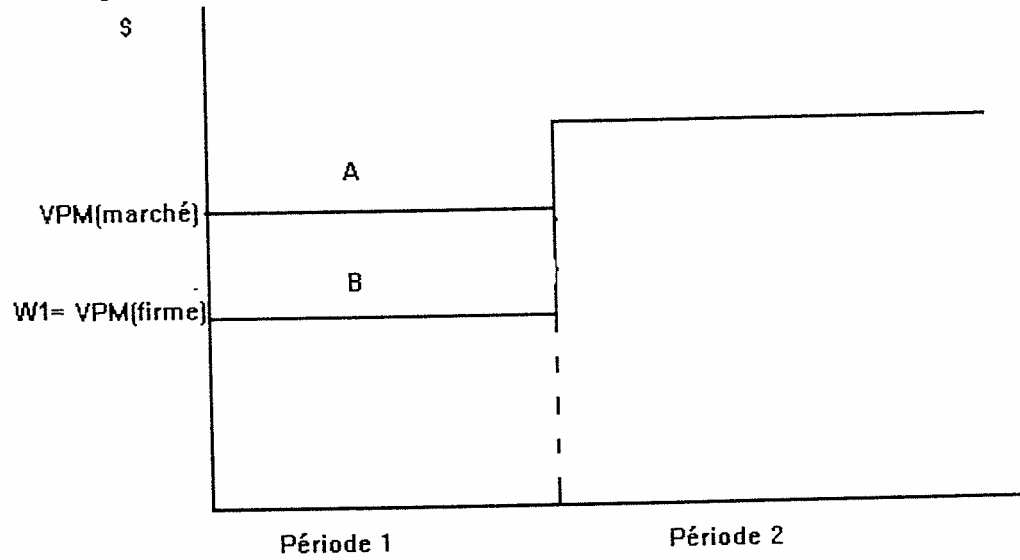
Notes- Le niveau déterminant si un résultat est significatif ou non est fixé à 5%.

1979

Tableau 7. Effet de la formation sur le salaire de départ.

Banque de données	Auteur(s)	Variable dépendante	Effet (par type ou source de formation s'il y a lieu)
EOPP	Barron, Black et Loewenstein (1989)	Log du salaire de départ	0%
SBA	Barron, Berger et Black (1993)	Log du salaire de départ	-2.8% (évalué au nombre moyen d'heures de formation déclarées.
Enquêtes portant sur des usines américaines et japonaises	Levine (1993)	Log du revenu annuel	Aucun effet significatif et ce, peu importe le type de formation mesuré.

Figure 1. Capital humain général.



VPM: Valeur de la productivité marginale W: Salaire

LQ19

Figure 2. Capital spécifique à la firme.

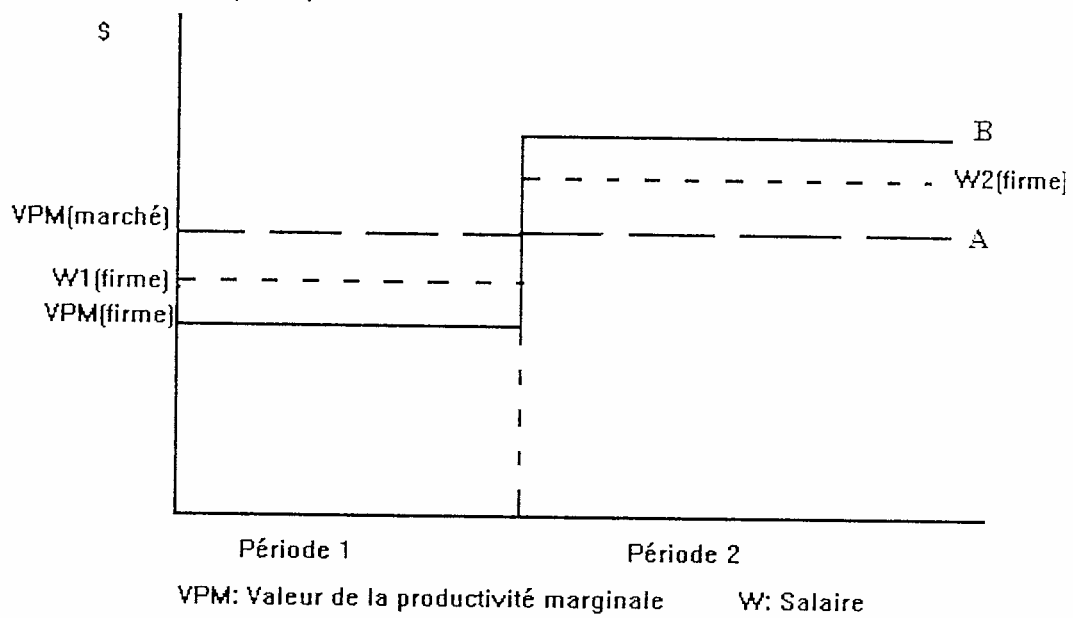
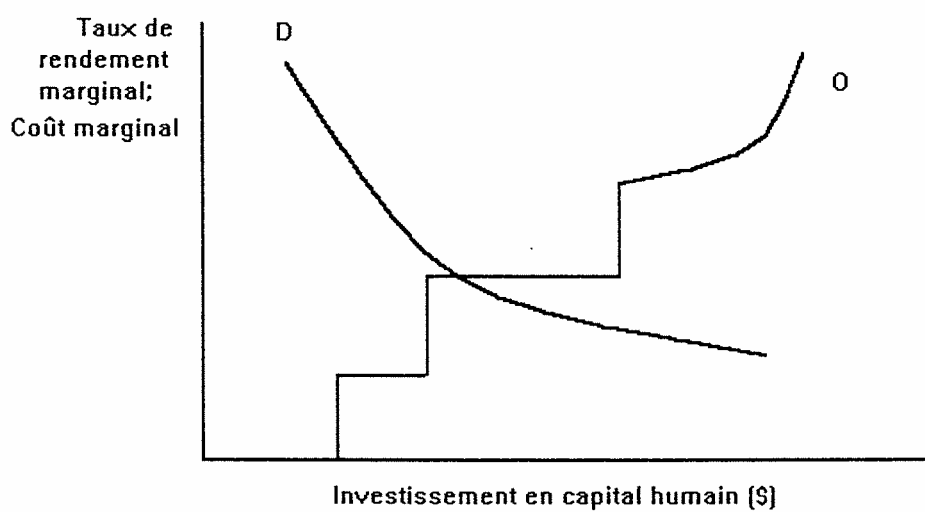


Figure 3. Offre et demande de capital humain.



UCLON

Figure 4. Stock d'équilibre de capital humain.

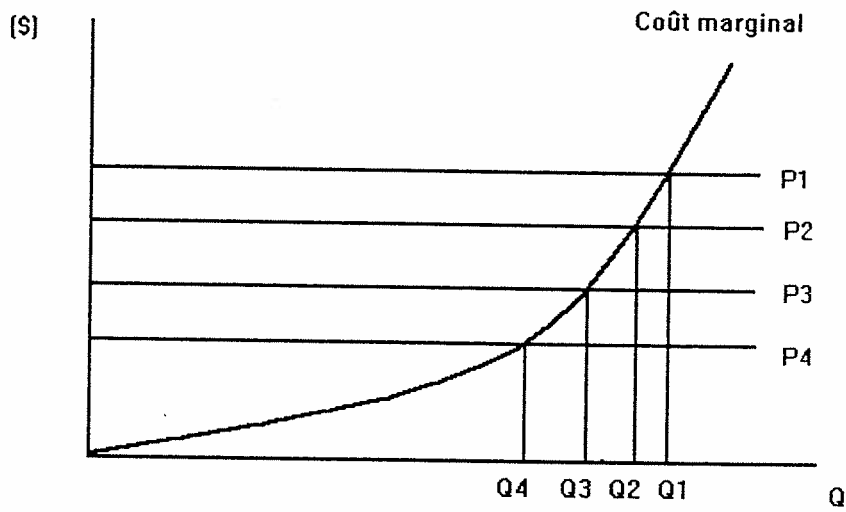
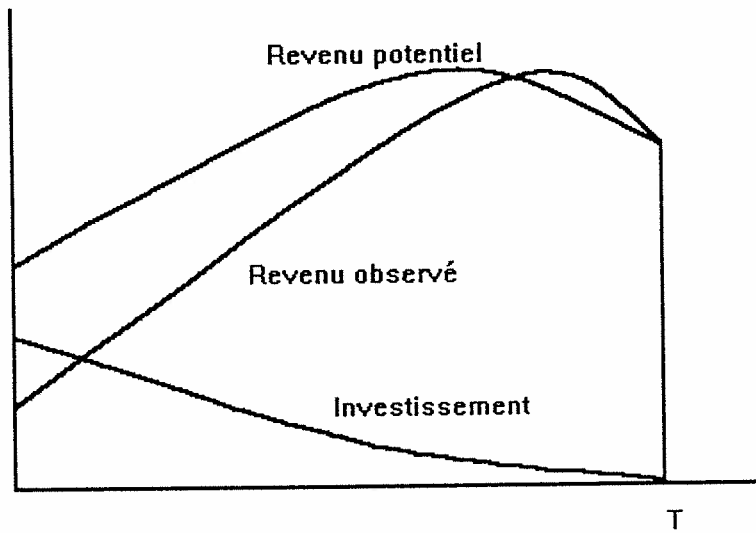
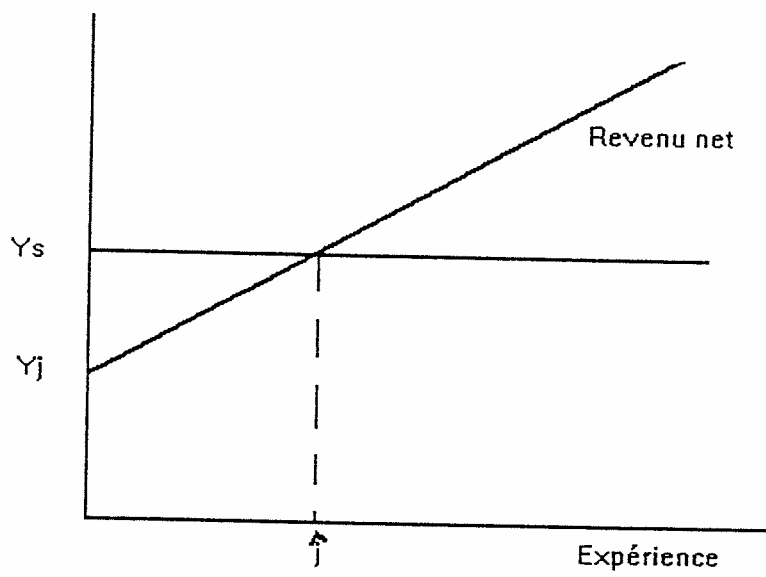


Figure 5. Profil d'investissement et de revenu.



BLSM

Figure 6. Point de renversement.



Deuxième étude

Wages and Mobility: the Impact of Employer-Provided Training

ICTD

I. Introduction

Much of the debate surrounding the issue of whether wages rise with years of seniority has focused on the magnitude of the tenure effect and on the possible explanations for this effect, including the theory of human capital. Only recently have there been attempts at measuring directly the effect of accumulating human capital through training. Mincer (1988), Brown (1989), Barron, Black and Loewenstein (1989), Barron, Berger and Black (1993), and Levine (1993) have attempted to do so with training data that were of a qualitative and subjective nature while Lynch (1992) and Blanchflower and Lynch (1994) used a data set, the National Longitudinal Survey of Youth (NLSY), that has the actual number of weeks spent training.¹

The objective of this paper is to provide yet another look at this question using the NLSY for the period spanning the years 1979 to 1991. The longitudinal aspect of the larger data set is then exploited to analyze the impact of explicitly taking into account unobserved heterogeneity on (i) the effect of various forms of training on the wage profile, ii) the extent to which workers changing jobs benefit from having been trained by previous employers, and (iii) the propensity to change jobs. Also, the question of whether workers pay for their training by accepting lower starting wages will be examined by using a variable, on-the-job training still in progress, that seems best suited for picking up that effect, if indeed workers do pay for the training expenses (which are shown to be incurred largely by employers).

The estimating strategy is the following: after estimating with ordinary and generalized least squares a standard wage equation augmented with the training variables, these variables as well as the tenure and experience variables will be instrumented in order to alleviate correlation problems between them and the individual and job-match fixed effects, following the procedure proposed by Altonji and Shakotko (1987). The results show that there are

¹ See also Altonji and Spletzer (1991) for another example of a study which makes use of data on the duration of training spells. Their paper, however, focuses on the determinants of employer-provided training.

substantial returns to training, more particularly on-the-job training, whether it be with the current employer or with subsequent employers. In fact, the results indicate that the human capital acquired through training appears to be just as valuable (to the worker) with subsequent employers than it is with the current employer. This apparent portability of the skills acquired through on-the-job training suggests that we should be observing lower starting wages for workers who are undergoing training. Indeed, the results show that workers implicitly pay for these returns by having lower starting wages. This is shown by estimating a starting wage equation with ordinary least squares where some control is provided for unobserved heterogeneity. Next, a proportional hazard model is estimated and the results show a substantial degree of firm-specificity for human capital acquired through training received with the current employer. This result holds even if control is provided for unobserved person-specific heterogeneity through the use of fixed-effect methods. Note, however, that when I control for such heterogeneity, training received with previous employers is shown to have no significant positive impact on mobility, contrary to the case in which there is no control for heterogeneity.

These results point toward the following conclusion: for this sample of young workers, employers react to market pressures by setting wages at the market level with no premium paid for the firm-specific productivity of the trained workers. To be more precise, the results from the wage equation and those from the hazard model indicate that formal on-the-job training in the current job contains at once a general component which the employer rewards up to its market value and a firm-specific component which reduces mobility but which does not appear to be rewarded. Thus, there is little evidence that these workers are paid much in excess of their outside option.

The rest of the paper is organized as follows. Section II briefly reviews the theoretical models underpinning the estimation. Section III provides a thorough description of the data set and of the various definitions of the training variables used in the estimation procedures. Particular attention is paid to the timing of training over the course of an employment

relationship and on who pays the direct costs of training, observation that will help in the interpretation of the results. Section IV presents the estimating framework as well as a discussion of the results. Section V concludes the paper.

II. Theoretical Framework

According to the theory of general human capital as formulated in Becker (1975), workers who invest in general on-the-job training should pay the full costs and reap the full return from their investment. If firms pay the direct costs of general training, then workers should be expected to reimburse the firm by accepting a lower starting wage. In other words, during the training period, workers are paid below their value on the market. When training is completed, their value to the firm as well as their market value increase and competitive forces insure that they are paid at their market value. An immediate implication is that wages should rise with experience in the labor market since productivity increases with time in the labor market for those who get general training. Moreover, the theory predicts that investment in on-the-job training should decline as the workers age for basically two reasons: 1) the period over which the rewards will accrue to the workers gets shorter and 2) the opportunity cost of training increases as the wage increases over time. On the other hand, if workers invest in entirely firm-specific on-the-job training, the theory suggests that the firm and the workers should share in the returns and the costs. The motivation for sharing the investment stems from the uncertainty surrounding the parties' post-investment behavior. For instance, were the workers paying the full costs of the investment, they would face the possibility of an employer-initiated separation in the period following the investment. They would therefore be incurring a capital loss since, by definition, the value of marginal product outside the firm is the same whether training has occurred or not. As for the firm, if it were to bear the full costs of training, the possibility that the workers would leave while the firm was in the process of harvesting the return on its investment would imply a capital loss for the latter. Consequently, both parties share the costs and the returns on the investment. The empirical implications are that 1) wages should be rising with seniority because productivity

increases with seniority and 2) turnover should decrease when there is investment in firm-specific human capital.

Hashimoto (1981) subsequently provided a formalization of Becker's sharing hypothesis. He showed that sharing arises in response to transaction costs, not because of uncertainty concerning the behavior of parties in the post-investment years. More precisely, since it is costly for both parties to evaluate and agree on the worker's productivity in the firm and outside the firm in the years following the investment, there will be situations where the worker may quit or the firm may dismiss the worker even if separation entails a net loss. Therefore, parties will determine the optimum sharing ratio prior to undertaking the investment, thus minimizing the loss from non-optimal separations. Hashimoto shows that if there were no transaction costs, the sharing ratio would be irrelevant to the separation decisions. Although the main reason for sharing is different than in Becker's original argument, the empirical implication for the wage profile is the same, namely that wages should rise with seniority.

An even more efficient contract was derived by Carmichael (1983). Like Hashimoto, he assumed that after an initial period of training, information is revealed asymmetrically: only the firm knows the value of marginal product while only the worker knows his degree of job satisfaction. Since information cannot be exchanged ex post due to transaction costs, both parties make their separation decision independently. The worker leaves if his job satisfaction is below a certain threshold while the firm fires the worker if his revealed productivity is below a critical level. Carmichael argues that welfare can be improved by making both parties internalize the entire expected losses from a separation, regardless of who initiates it, instead of merely internalizing their share of the return, as in Hashimoto's contract. To achieve such a constrained optimum, Carmichael modified Hashimoto's structure by introducing seniority-based promotions. Suppose that there are two jobs in the second period, type 1 jobs and type 2 jobs. The productivity of the workers is identical in the two types of jobs. After separation decisions are made at the start of the second period, all workers still with the firm enter type

1 jobs where they are paid w_2 . Later in period, some of these workers are promoted to type 2 jobs on the basis of their seniority where they are paid $w_2 + B$ where B is a bonus which is determined ex ante when the parties agree on the contract. By assumption, there is a fixed number of type 2 jobs, therefore a layoff can only save the firm w_2 since a junior worker will immediately fill the vacant position. Thus, the incentive for the firm to fire any high wage worker is reduced. In effect, junior workers are acting as a third party by benefitting from any separation since the layoff of a senior worker increases their own seniority, thus improving their chance of eventually getting a type 2 job. Having determined the second period wage schedule, the first period wage is adjusted so that the expected value of the contract equals that available to the worker in the market. Carmichael shows that this system of promotions based on seniority makes the separation decisions more efficient. Two implications of Carmichael's model are worth emphasizing. The first one is that the more senior workers will typically be earning more than their value of marginal product to the firm. Thus the theory of specific human capital can respond to the challenges posed by Lazear (1979) concerning mandatory retirement and by Medoff and Abraham (1980) who provided evidence that wages might rise with seniority even if productivity did not. Secondly, his optimal contract implies that workers pay the full cost of the investment in the form of lower starting wages. It should be noted Carmichael implicitly assumes that the firm can commit to the wage schedule derived ex ante. This is an important assumption because it assures that the firm will not change the number of type 2 jobs available, even though it would be profitable to do so once a junior worker is due for a promotion.

By explicitly taking into account the possibility of contract renegotiation, Macleod and Malcomson (1993a,b) show that firms making specific investments need not offer above market clearing wages to their workers. Since it is impossible or too costly to write a complete contract ex ante that would prevent renegotiation, there will be situations in which the parties renegotiate an incomplete contract. This happens if and only if one party prefers to break the contract though it is still efficient for the parties to continue their match. In this case, one party is prepared to give up some of its gains in order to induce the other party not

to breach. To be more precise, a firm and a worker have formed a match in period one and the firm has made a specific investment that affects second-period payoffs, and the worker invests only in general skills. Random shocks to the productivity of the worker and to his outside market wage become known to both parties at the end of the first period. If the draw the worker receives makes the outside option preferable to the negotiated wage at the beginning of the relation, assuming that it is efficient to continue the match, the firm renegotiates the worker's wage to the level of the outside option. Similarly, if the firm prefers to fire the worker rather than pay her the contract wage, the worker is prepared to take a pay cut in order to avoid a separation. Naturally, there are situations where no renegotiation occurs after the realization of the random shocks. In that case, both parties prefer to trade at the contract wage that involves some sharing of the return on the firm's investment. When only the firm invests in the worker, a contract which avoids the "holdup" problem (see Williamson et al. (1975)), is one in which the worker receives her market alternative each period, and hence wages do not rise with tenure. When both parties make relation-specific investments, Macleod and Malcomson show that if the worker accepts the terms of the efficient contract, then the measured effect of tenure on wages is positive.

III. The Data

The National Longitudinal Survey of Youth data set surveyed 12,686 young males and females who were between the age of 14 and 21 in 1979². It contains detailed employment histories of the respondents thereby permitting the construction of relatively error-free variables for tenure as well as for the total experience accumulated since the beginning of one's full time transition to the labor market. Of particular interest is the existence of a detailed account of the training histories of the respondents, with the starting and ending dates of up to four spells³ of training lasting at least four weeks since the last

² The response rate was at 71% in 1991.

interview. While the data give a good account of the formal aspect of skill acquisition it is missing out on the informal aspect and, possibly, on formal programs of short duration. The training spells are broken down into up to 12 categories (the actual number varies from year to year, although it stays at twelve from 1988 onward) which were then classified in three groups in the same fashion as Lynch (1992). Categories 8 and 9 (company-provided training and seminars or training programs at work but not run by employer) were classified as on-the-job training, category 3 represents apprenticeship programs and categories 1 (business colleges), 2 (nurses programs), 4 (vocational-technical institutes), 5 (barber and beauty), 6 (flight school), 7 (correspondence courses), 10 (seminars or training programs outside of work), 11 (vocational rehabilitation center) and 12 (other) were classified as off-the-job training. It is also known whether the programs were completed, not completed or still in progress at the time of the interview. Since these workers can be followed as they move from employer to employer, it is then straightforward to calculate the accumulated weeks of training with all previous employers⁴ by each type of training. I also know the identity of the party who incurred the direct costs of the training programs for the years 1988-1991. Even though this information is not available for the whole period considered, the four years for which it is available provide strong evidence that employers are paying for OJT.

Lynch (1992) could only rely on a 1983 cross-section of the data set to get most of her results while Blanchflower and Lynch (1994) followed 1979, 1980 and 1981 subsamples of 18-year-olds until they were 25 years old. Here the major restriction that I impose is in terms of entering the labor market on a full-time basis.⁵ The people who were considered as meeting that criterion were (i) those whose primary activity was either working full-time, on

³ Actually, three spells from 1979 to 1986 and four spells from 1988 to 1991. The fact that no questions concerning training were asked in the 1987 interview does not pose any particular problems (except possibly accuracy of recall) because the starting and ending dates given in 1988 allow for the appropriate assignments with respect to the 1987 and 1988 interview dates.

⁴ Note that all training undergone by the individuals while they were not employed has been excluded from the calculations.

⁵ We also limit the sample to workers who are at least 18 years of age.

a temporary lay-off or looking actively for a job, (ii) those who did not return to school on a full-time basis within six years⁶ and (iii) those who had worked at least half the year since the last interview and who were working at least 20 hours per week. Individuals excluded from the sample are those younger than 18, those that had been in the military at any time, the self-employed, the ones whose jobs were part of a government program and the ones working without pay, those who were in the farming business and also all public sector employees. Another restriction was to exclude those that occupied two jobs (or more) on a full-time basis because it is impossible to determine with which of these employers the training is taking place, thereby possibly falsifying the results. For the earnings equation, I am then left with 29,020 observations while the starting wage model has 13,394 observations. For the hazard model, only those that had at least two completed spells of employment were retained to allow the implementation of a fixed-effect estimation methodology, which leaves me with 8,097 spell-workers.⁷

Some summary statistics of the sample are provided in table 1 while table 2 gives us a breakdown of the number of employees who were trained by types of training and by industrial aggregates.

Next, table 3 shows some summary statistics comparing those that have received training (and who have completed their programs) with those that have not. To be more precise, the sample is restricted to workers who had positive values for the training variables at any moment in the employment relationship. For example, workers those that had no training in the first two years, but who received training thereafter are included in the sample.

⁶ The choice of six years as a cutoff point is arbitrary, and hence debatable. The idea is to exclude those that make "quasi-permanent" transitions and who might be considering returning to school a few years down the road. It seems reasonable to assume that few people would enter the labor market while planning to leave it in six years or more to go back to school. The same could not be said if we were considering a one to three year (say) horizon. In any event, the results were left unchanged if all school returners were excluded.

⁷ Note that people were allowed to work less than 20 hours per week only if they were undergoing training. Also, to avoid left-censoring, individuals who had started working prior to January 1, 1978 were eliminated.

This way of calculating the means avoids any upward bias that would have resulted had these people been excluded. Beside the wage differentials, it is very interesting to note that the trained workers have enjoyed longer careers with their respective employers compared to those that have never been trained. Also, it is worth noting that the additional experience that the trained workers have is usually accounted for by the additional tenure. Therefore, taking these numbers at face value, it would seem that all types of training are of a fairly specific nature.

If all training costs are paid by the employers and the skill enhancement programs are, at least to some degree, portable, then you would expect the workers to bear some portion of the costs by receiving lower starting wages. As is shown in table 4, it appears that employers are bearing much of the costs of all types of training programs, especially the programs corresponding to on-the-job training where between 1988 and 1991, 96.5% of the direct costs were paid by the employers. Although that information is not available for the whole length of the panel, it seems very unlikely that a dramatic shift would be observed in the event that it was. Interestingly, even the categories of training programs which are labelled as off-the-job were paid for in large part by the employers, suggesting that many of these programs were considered as relevant to the job, thereby straining a little bit the usefulness of distinguishing between on-the-job and off-the-job programs.

The issue of job relevance of the training programs can also be examined by looking at the timing of the programs i.e. how soon after starting working with their current employer do workers begin their training. One expects that the sooner the worker receives training after being hired, the more relevant to the current job is the program. A firm making specific investments maximizes its profits by pocketing the returns from these investments as soon as possible. Figures 1 to 4 allow us to see the extent to which training starts early in the employment relationship. In figure 1, the number of months elapsed since the beginning of the job has been plotted against the number of workers undergoing their first (and only the first) on-the-job program with their new employer⁸. It is clear from this graph that much of the

training occurs very early in the relationship. The same can be said for off-the-job and apprenticeship programs although to a lesser extent regarding OFT (see figures 2 and 3). This pattern of early training is important for estimation purposes since most if not all studies that have used training data of a qualitative and subjective nature have been unable to find convincing evidence of workers paying for their training through lower starting wages. As Levine (1993) recognizes, it may be that such data are too coarse to pick up that effect if it is there. In the next section, I find evidence for this hypothesis by focusing on training that is still going on at the moment of the interview.

The tenure-training (OJT) profile is illustrated in figure 4. To construct the training profile, the sample is restricted to workers who had (OJT) training with their current employer. The average number of weeks of completed training for these workers by tenure intervals of one year are then computed⁹. As the theory of human capital predicts (e.g. Ben-Porath (1967)), all investment is heavily frontloaded. The profile fluctuations following the first year reflect the turnover of workers in the sample.

IV. Results

IV.1 Earnings Equation Estimates

Consider the following log wage equation augmented with training variables:

$$\ln w_{ijt} = \beta_1 T_{ijt} + \beta_2 Exp_{it} + \beta_3 TCJ_{ijt} + \beta_4 TPJ_{it} + \alpha_i + \theta_{ij} + \epsilon_{ijt} \quad (1)$$

where w_{ijt} represents the real hourly wage of person i in job j at time t , T is tenure, Exp is total labor market experience, TCJ is training in current job, and TPJ is training in previous jobs.

⁸ Note that not all workers undergoing their first program are represented here; I have set the cutoff point at 28 months.

⁹ For the first year of tenure, it is also the case that by the six-month mark, workers have on average over 12 weeks of completed OJT training.

All other covariates, including higher order terms for experience and tenure, are ignored for ease of presentation. The unobserved heterogeneity components can be decomposed into an individual effect (α_i) and a job-match effect (θ_{ij}).¹⁰ The person-specific effect can be seen as representing unmeasured aspects of each individual's earning ability while the job-match component represents the unknown (to the econometrician) quality of the employment relationship stemming from search activity, for example. Both of these effects are assumed to be time-invariant. It is important to reiterate that the training variables available in the NLSY reflect the formal aspect of the process by which workers accumulate human capital, in that only training programs which last at least one month are accounted for. No doubt that we miss out on the informal aspect of the training as well as on the formal aspect for short (i.e. less than one month) programs.

The training covariates for each category of programs (on-the-job, off-the-job and apprenticeship) are partitioned into two parts: training received with the current employer and all training received with previous employers, and all these variables have been converted from weeks to years (and fractions thereof). Column (1) of table 5 shows the results obtained by using OLS.¹¹ As can be seen, all three types of training are rewarded by the current employer with completed OTT (OTTC) being rewarded the most. It also appears that human capital accumulated through formal training programs with previous employers is quite portable from firm to firm. However, the same sort of individual and job-match heterogeneity biases which affect the coefficients on tenure and experience are likely to be present regarding the training variables. More able individuals (those with high α 's) may have enjoyed careers that were interrupted less frequently by unemployment spells and they may have received more training, both with the current employer and with previous employers, while better matches (high θ 's) are likely to be formed if you have more experience and more training (due to human capital effects and search effects). Also, tenure and training with the current

¹⁰ See Chamberlain (82) for a discussion of nuisance parameters in panel data.

¹¹ All results are calculated using the weighted sample.

employer are possibly correlated with the job-match component. Dropping for the moment the assumption of a time-invariant job-match component, let's suppose that it can be written as

$$\theta_{ijt} = \psi_1 T_{ijt} + \psi_2 Exp_{ijt} + \psi_3 OJTC_{ijt} + \psi_4 OJTCP_{ijt} + \omega_{ijt} \quad (2)$$

where ω_{ijt} is assumed to be orthogonal to the regressors. The other training variables are not shown for ease of presentation but the same analysis applies to them. The discussion above suggests that ψ_2 and ψ_4 are positive. In the context of maximizing behavior on the part of a worker who faces a wage distribution, Topel (1991) argues that ψ_1 is negative once we control for experience, assuming there is a tenure effect. If there is no tenure effect, then ψ_1 equals zero. Regarding ψ_3 , given that we control for both experience and tenure, it seems reasonable to argue that the job-match component of wages is non-negatively correlated with on-the-job skill acquisition. Substituting equation (2) into equation (1) we get

$$\ln w_{ijt} = (\beta_1 + \psi_1) T_{ijt} + (\beta_2 + \psi_2) Exp_{ijt} + (\beta_3 + \psi_3) OJTC_{ijt} + (\beta_4 + \psi_4) OJTCP_{ijt} + \alpha_i + \omega_{ijt} + \epsilon_{ijt} \quad (3)$$

We see from equation (3) that although we are interested in the β 's, using ordinary least squares will produce estimates of composite effects and the regressors would still be correlated with α . To provide some correction for these problems, the instrumental variable (IV) methodology of Altonji and Shakotko (1987) is used and is extended to the training variables¹². All variables pertaining to training with the current employer, as well as tenure, are instrumented with their deviations from job-match means, whereas experience and training with previous employers are instrumented with their deviations from individual means. The instruments for training with the current employer and tenure are, by construction, uncorrelated with the job-match component while the instruments for training with previous employers and experience are, also by construction, uncorrelated with individual component.

¹² See also Finnie (1993) for an extension of the method of Altonji and Shakotko to the experience variable.

This methodology is preferred to first-differencing on the ground that differencing has greater potential to further enhance any measurement errors present in the data as compared with using deviations from means.

Another estimation problem pertains to the fact that residuals are likely to be serially correlated given that the same individuals are followed over time. Consequently, to provide correction for this problem, equation (1) is reestimated with generalized least-squares (under the assumption of a fixed individual effect along with an i.i.d. term), both with and without the IV methodology proposed. The results are reported in columns (2) and (3) of table 5.

Focusing on on-the-job training first, we can see that most of the results are qualitatively the same as with OLS, except that the coefficients are smaller both with GLS and IV-GLS. Therefore, the returns estimated with OLS cannot be attributed solely to job-match or individual heterogeneity. Curiously, for off-the-job training programs, GLS and IV-GLS estimates are much larger than with OLS. Thus, it seems that correcting for heterogeneity has the effect of removing the downward bias that poorly matched or low-ability workers cause to the estimated coefficients. This contrasts with the results on apprenticeship programs where the results suggest that well matched and high-ability individuals are selected into the programs. Given the few workers who undertake apprenticeship (63), a word of caution seems warranted here regarding the results pertaining to this type of training. The basic conclusion that is to be drawn from these results is the portability of formal on-the-job and off-the-job training programs.¹³ It appears that all completed programs of OT and OFT are fairly general. Also, since the OTCC (OFTCC) and OTCP (OFTCP) coefficients are not significantly different, we are led to conclude from these results that current employers set the wage of their trained workers at pretty much the market level following completion of the programs. Already the summary statistics in table 3 provided a clue that employers reacted to market pressures. There we saw that workers who had received OTT stayed with their employers a full year and a half longer on average. This fact and the fact that OTT seems to

¹³ In fact, a simple F-test on the equality of the coefficients associated with OTCC and OTCP could not reject the hypothesis that the two are indeed equal.

¹⁴ See Kiefer (1988) for a survey of hazard models or Lancaster (1990) for a detailed textbook exposition.

Then, it can be shown that the partial log-likelihood function is equal to:

$$\lambda_{ij}(t) = \exp(\beta' X_{ij}(t) + \alpha_i) \lambda_{i0}(t), \quad j = 1, \dots, n_i; \quad i = 1, \dots, N. \quad (4)$$

To gain a better perspective on the degree of firm-specificity of the human capital acquired with the current employer through formal training, a proportional hazard model is estimated by using Cox's partial likelihood approach.¹⁴ The advantage of this approach is that it avoids the need to specify a baseline hazard function (actually, all common factors, including the baseline hazard, cancel from the partial likelihood). Another advantage of Cox's methodology is that it is possible to eliminate all individual-specific factors, including individual-specific baseline hazards, provided that multiple spells of employment are available for each worker. To be more specific, let's suppose that for worker i we have n_i spells (ordered by their increasing length) and that the duration for each spell is denoted t_{ij} , where j stands for the spell number. Assuming all spells for the same person are independently distributed given her heterogeneity parameter, I can write the hazard functions as:

IV.2 Specification of the Hazard Model

be, at least to some degree, general (thereby increasing outside job opportunities) is not really contradictory; rather, it just means that employers manage to keep their trained workers by paying them their worth on the outside market.

$$L^p = \frac{\sum_{i=1}^N \sum_{j=1}^{i-1} \ln \left(\frac{\sum_{k=j}^i \exp(\beta' X_{it}^{(k)}(t_{it}^{(k)}))}{\exp(\beta' X_{it}^{(i)}(t_{it}^{(i)}))} \right)}{\quad} \quad (5)$$

where the denominator corresponds to the risk set of worker i . Note that both α_i and λ_{i0} do not appear in equation (5). I should point out that a rigorously specified hazard model dealing with time-varying covariates such as the training variables ought to condition on the complete time-paths (past and future) of the covariates, especially if these covariates are considered endogenous to the leaving process. This issue is a difficult one to treat and I approximate the paths by using the current values (as of t) for these variables.

The results are presented in table 6. The effect of correcting for unobserved individual heterogeneity, while changing the magnitude of the coefficients (they are actually larger in absolute value), does not alter the basic conclusion that on-the-job training with the current employer seems to have a sizeable impact on the conditional probability of leaving.¹⁵ Consequently, it can be inferred from these results that the skills accumulated are fairly firm-specific. On the other hand, it seems that skills acquired with previous employers are not a significant factor of increased mobility once control is provided for unobserved individual heterogeneity. However, we should not be so surprised at such a result. It is a well known empirical fact that mobility tends to decrease with experience, presumably because workers, as time goes by, are able to find increasingly better matches.¹⁶ Yet, labor market experience serves as a proxy for general human capital in a standard earnings equation. Given that skills accumulated with previous employers, as measured by the number of weeks spent training, are positively correlated with experience, then it is likely that this variable is correlated with the quality of the match in much the same way as experience is. Therefore, even if the

¹⁵ Note that the effect of failing to control for heterogeneity is to bias the coefficients downward in a partial likelihood framework (see Lancaster (1990)). Therefore, it is not possible to conclude that the increase in the coefficient associated with OJTCC is due solely to the fact that I control for the presence of low-ability individuals.

¹⁶ See Topel and Ward (1992) on that subject.

The way to capture that effect is by setting up a starting wage equation and by using a variable indicating that training is still going on. OJT which has not been completed (workers quit on the programs or the program was not completed for other reasons) will not do for the following reason. Assuming that workers should be paying for their training by having lower wages at the start, the results above show that market pressure forces employers

The NLSY does not contain any information concerning the costs of these training programs and hence we cannot determine if the workers pay the totality of the costs in the form of lower starting wages. However, given that the trained workers seem to be paid at a level approaching their market-wide marginal product and given that from table 4, it is known that between 1988 and 1991 employers paid over 95% of the direct costs of OJT programs, we should be finding that the workers indirectly pay for part of these programs in the form of lower starting wages.

IV.3 The Impact of Training on the Starting Wage

The results in table 6 coupled with those in column (2) of table 5 indicate that formal on-the-job training in the current job contains at once a general component which the employer rewards up to its market value and a specific component which reduces worker mobility. It is then plausible to suggest that firms tend to keep their trained workers longer as compared with other workers because they are more productive.

concludes that skills acquired through previous training were to be negative and significant, it cannot be concluded that skills acquired through previous training are specific to the current firm. By focusing on the probability of exiting from the first job, thereby circumventing to a degree the problem created by the positive correlation between match quality and experience, Lynch (1991) finds evidence of a positive impact of prior off-the-job training on mobility. Thus, she concludes the skills acquired through this type of training is fairly general.

to pay their trained workers at their marginal product following completion of OJT.¹⁷ Then, there is no reason to believe that things are much different for workers who quit on a program. More precisely, their wage should be reset at the value of marginal product just like it would have been had they not undertaken any training program.¹⁸ If I use instead the number of weeks of OJT training undertaken by the worker and still going on at the time of the interview, there lies the best hope of capturing any effect. There remains the problem of unobserved heterogeneity. Authors (e.g. Barton, Black, Loewenstein (1989)) have generally acknowledged that the effect of job-match or individual heterogeneity biases will be to underestimate the impact on the starting wage, even possibly masking it totally, the reason is that more able persons may receive more training and may be paid more even if they are undergoing training as compared with workers who are not being trained.¹⁹ Therefore, any result found should probably be seen as a lower bound on the true result. However, some control is provided for unobserved heterogeneity with a variable indicating the maximum amount of training that person will have received by the end of her relationship with the employer.²⁰ If it is true that higher ability individuals are likely to be paid more and to receive more training, then that variable should be picking up some of that effect. Just to make clear the methodology employed here, I used only the first wage observations on each job and then

¹⁷ To quote Mincer (1962) p.53: "A direct computation of foregone earnings of workers engaged in on-the-job training would be possible if data were available on their earnings during and after the period of training, and on earnings of a comparison group of workers who have the same amount of formal schooling and are otherwise similar to the trainees, but do not receive any on-the-job training. Presumably, the latter would have a flatter age-earnings profile than the former."

¹⁸ Uncompleted OJT never came out as significant in any of the regressions although it did show up as negative whether it was instrumented or not. Perhaps this reflects some sluggishness on the part of the wage as it may not be reset immediately after quitting training. Another possible explanation is that employers revise (downward) their expectation regarding the productivity of the workers. In any event, these problems do not occur if we use ongoing training at the time of the interview.

¹⁹ Actually, this hypothesis goes back to Mincer (1962). On page 51 (footnote 4) he writes "...Greater learning from experience is characteristic of workers with greater motivation and ability, and their earnings at the early stages of the career may be as high or higher than those of other workers. But such finding that people with greater ability have higher productivity than others at any given stage of experience does not negate the existence of investment in on-the-job training, though it may bias the estimation of its magnitude."

²⁰ See Abraham and Farber (1987) for another application of a similar control.

used OLS. Results are reported in table 7. Even though only 134 observations had a positive value for OTT still going on (OTTOC)²¹, the effect is relatively large and significant: workers do seem to pay for their OTT through lower starting wages. The same cannot be said for OFT and APP but this is not really that surprising given that over 25% of OFT programs undertaken between 1988 and 1991 were not paid by the employer (35% for apprenticeship) thereby introducing more noise in the model. Also, when the control for unobserved heterogeneity is added to the equation, the effect is larger by about one percent.

To summarize the results I have obtained, it seems that OTT and OFT with the current employer is neither entirely general (otherwise there would be no reason for it to reduce mobility) nor is it entirely specific (since then, having been trained by previous employers should have no impact on the current wage), but is rather a blend of these two extremes. Also, there is no evidence that employers pay their trained workers above their market-wide marginal product. Finally, workers appear to implicitly pay for the degree of portability of their newly acquired skills by having lower starting wages.

V. Conclusion

The idea that general skills accumulated through training should be paid for by the employee is a central prediction of the theory of human capital. However, all efforts directed at verifying that prediction have been practically fruitless, so much so that some researchers (e.g. Barron, Berger and Black (1993)) have started interpreting the lack of evidence of workers paying for their training through lower starting wages as possible evidence of dual labor markets. While the latter might indeed exist, this paper has provided direct evidence that workers do in fact have lower starting wages while they are going through training. Also, it has been shown that firms respond to market pressures by setting the wage for trained workers at the market level even if, as the hazard model results show, a sizeable portion of

²¹ Given that much of the training takes place very early in the employment relationship (see fig. 1), we should not be surprised to find so few respondents still in training at the time of the interview.

the investment in on-the-job training is specific. Therefore, the evidence concerning training is not consistent with the firm and the worker sharing the rent on the specific part of the investment. I should mention again that these results hold for formal training programs which naturally represent only a portion of the process by which workers accumulate skills on their current job. It could be that the informal process of skill acquisition makes the workers more indispensable to the firm which could then ill-afford to lose them and would consequently, perhaps, be willing to share the rent over the return on the specific capital embodied in these workers.

Table 1

MEAN SAMPLE CHARACTERISTICS (Weighted)-NLSY

5.75	Real Hourly Wage (\$1979)
41.7	Hours Worked
2.44	Tenure
5.82	Experience
12.44	Years in School
890	Number of Workers with On-the-job Training
13.68	Number of Weeks Trained (OJT-Completed) *
721	Number of Workers with Off-the-job Training
15.85	Number of Weeks Trained (OFT-Completed) *
63	Number of Workers with Apprenticeship Training
31.89	Number of Weeks Trained (APP-Completed) *
12.6	Percentage Nonwhite
44.7	Percentage Married
45.4	Percentage Female
25.1	Age
29,020	Number of Observations
5,649	Number of Individuals
13,590	Number of Jobs

Note-*Conditional on having positive value for these variables.

NUMBER OF JOB-WORKER OBSERVATIONS WITH
TRAINING BY INDUSTRIAL AGGREGATES

INDUSTRIES	ON-THE-JOB	OFF-THE-JOB	APPRENTICESHIP
------------	------------	-------------	----------------

Forestry and Fisheries (282)	7	21	0
Mining (398)	24	16	0
Construction (2,129)	43	77	115
Manufacturing (7,838)	590	494	94
Transportation, Communication and Public Utilities (1,890)	354	163	19
Wholesale and Retail Trade (7,233)	474	482	42
Finance, Insurance and Real Estate (2,282)	382	226	8
Business and Repair Services (1,969)	160	143	24
Personal Services (956)	56	67	5
Entertainment and Recreation Indus. (317)	20	19	1
Professional and Related Services (3,726)	425	492	11

Note-Number of observations indicated in parentheses.

SUMMARY STATISTICS-TRAINING (COMPLETED) VS NO TRAINING

TABLE 3

	OJT	No OJT	OFT	No OFT	APP	No APP
Real Wage (\$1979)	7.52	5.47	6.92	5.61	8.63	5.71
Hours Worked	42.23	41.57	41.68	41.65	42.65	41.64
Tenure	3.62	2.25	3.58	2.29	4.58	2.41
Experience	7.07	5.63	6.71	5.71	7.18	5.80
Training Weeks	13.68	0	15.85	0	31.89	0

Note-All statistics are weighed.

TABLE 4

TRAINING PROGRAMS UNDERTAKEN BETWEEN 1988 AND 1991

	ON-THE-JOB	OFF-THE-JOB	APPRENTICESHIP
Programs Undertaken	1159	734	54
Number of Programs Paid by Employers	1119	532	35
Percentage Paid By Employers	96.5	72.5	65.0

TABLE 5

EARNINGS FUNCTIONS ESTIMATES: THE IMPACT OF TRAINING
 (Dependent Variable: Log of Real Hourly Wages (\$1979))

Independent Variable	(1)	(2)	(3)
Tenure	0.0508 (0.0032)	0.0430 (0.0030)	0.0210 (0.0042)
Tenure Squared	-0.0029 (0.0003)	-0.0029 (0.0003)	-0.0019 (0.0003)
Total Experience	0.0532 (0.0044)	0.0768 (0.0044)	0.0945 (0.0049)
Total Experience Squared	-0.0011 (0.0003)	-0.0021 (0.0003)	-0.0025 (0.0003)
OJTCC	0.1692 (0.0255)	0.1666 (0.0274)	0.1216 (0.0372)
OJTCP	0.1984 (0.0278)	0.2049 (0.0355)	0.1319 (0.0439)
OFTCC	0.0750 (0.0212)	0.0931 (0.0216)	0.1420 (0.0300)
OFTCP	0.0089 (0.0173)	0.0820 (0.0247)	0.1260 (0.0335)
APPCC	0.0637 (0.0242)	0.0414 (0.0253)	0.0187 (0.0316)
APPCP	0.2146 (0.0224)	0.1464 (0.0308)	0.0733 (0.0380)
Industr. Dummies	YES	YES	YES
Occup. Dummies	YES	YES	YES
R-Squared	0.3431	0.8014	0.7949
F-TEST:			
OJTCC Coef.=OJTCP Coef.	0.59	0.84	0.04
p-value	0.44	0.36	0.84

Notes-Standard errors are in parentheses. Sample size is 29,020. Other covariates not shown in table include race, gender, marital status, union membership, SMSA, urban/rural, health, an intercept, unemployment rate in corresponding region, number of jobs ever held by respondent, four region and three education dummies. Aggregate effects are controlled for by having a dummy variable for each year. OJTCC: on-the-job training completed with current employer. OJTCP: on-the-job training completed with previous employers. The same definitions apply to off-the-job (OFT) and apprenticeship (APP) programs.

Table 6

HAZARD MODEL RESULTS
(Cox Partial Likelihood Approach)

Independent Variable	(1) No Control for Heterogeneity	(2) With Control for Heterogeneity
Experience (0-1)	0.3082 (0.0638)	0.3576 (0.1030)
Experience (1-2)	-0.1592 (0.1809)	-0.1958 (0.2898)
Experience (2-3)	-0.1858 (0.1799)	-0.1048 (0.2885)
Experience (>3)	-0.7406 (0.1819)	-0.2882 (0.2955)
OJTCC	-1.3516 (0.2350)	-2.9917 (0.5236)
OJTCP	0.2599 (0.1257)	-0.1513 (0.3741)
OFTCC	-1.5107 (0.2056)	-2.4151 (0.4172)
OFTCP	0.4439 (0.1068)	0.3932 (0.2971)
APPCC	-0.2926 (0.2274)	-0.0781 (0.3863)
APPCP	0.0258 (0.1660)	0.2231 (0.3205)
Number of Previous jobs	0.1186 (0.0033)	0.2058 (0.0128)
Indst. Dummies	YES	YES
Occup. Dummies	YES	YES
Log Likelihood	-54,738.01	-4,717.03

Notes-Sample size: 8,097 spells of which 1,019 are censored. Standard errors in parentheses. Unshown covariates are the same as those in table 5.

Table 7

OLS REGRESSION ESTIMATES OF THE EFFECT OF ONGOING TRAINING ON STARTING WAGES
 (Dependent Variable: log of Real Hourly Wages (\$1979))

Independent Variable	(1)	(2)
	Coefficients (Standard Errors)	Coefficients (Standard Errors)
Tenure	0.0438 (0.0047)	0.0440 (0.0047)
Prior Experience	0.0255 (0.0151)	0.0257 (0.0015)
OJTCC	-0.1699 (0.0594)	-0.1811 (0.0595)
Maximum OJT	-	0.1413 (0.0467)
During Relationship		
OJTCC	0.3013	0.1454
OJTCP	(0.0774)	(0.0928)
OJTCC	0.2261	0.2274
OJTCC	(0.0351)	(0.0351)
OJTCC	-0.0641	-0.0622
OJTCC	(0.0495)	(0.0495)
OJTCC	0.0201	0.0232
OJTCC	(0.0725)	(0.0724)
OJTCC	0.0845	0.0833
OJTCC	(0.0268)	(0.0268)
OJTCC	0.0797	0.0769
OJTCC	(0.0609)	(0.0608)
OJTCC	0.0293	0.0308
OJTCC	(0.1489)	(0.1489)
OJTCC	0.0051	0.0051
OJTCC	(0.0008)	(0.0008)
Indust. Dummies	YES	YES
Occup. Dummies	YES	YES
R-Squared	0.2640	0.2645

Notes-Sample size is 13,394. Unshown covariates are the same as those in table 5. OJTCC: Number of weeks of ongoing OJT with the current employer. The same definition applies to OJT and APP.

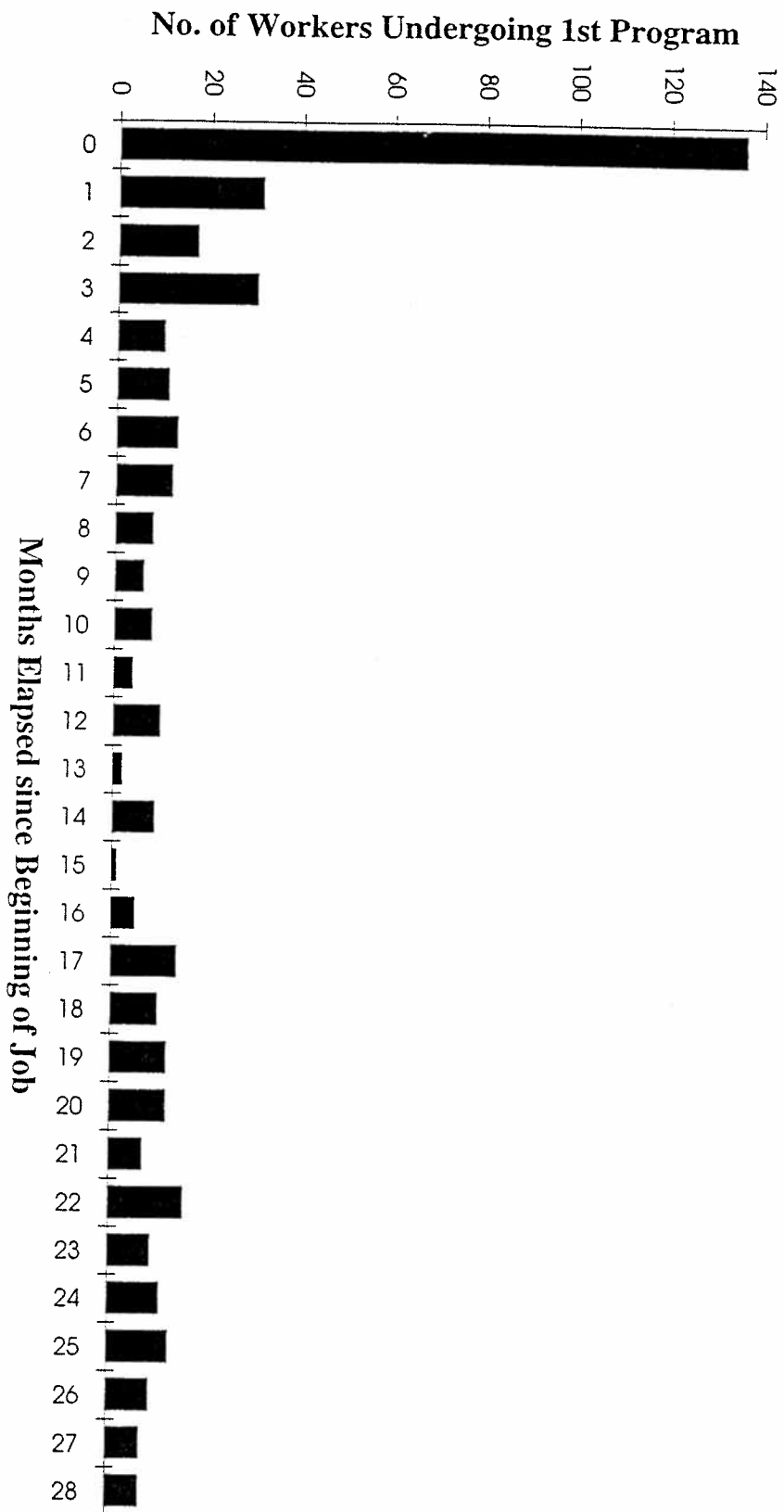


Figure 1. Timing of First Training Program (OJT)

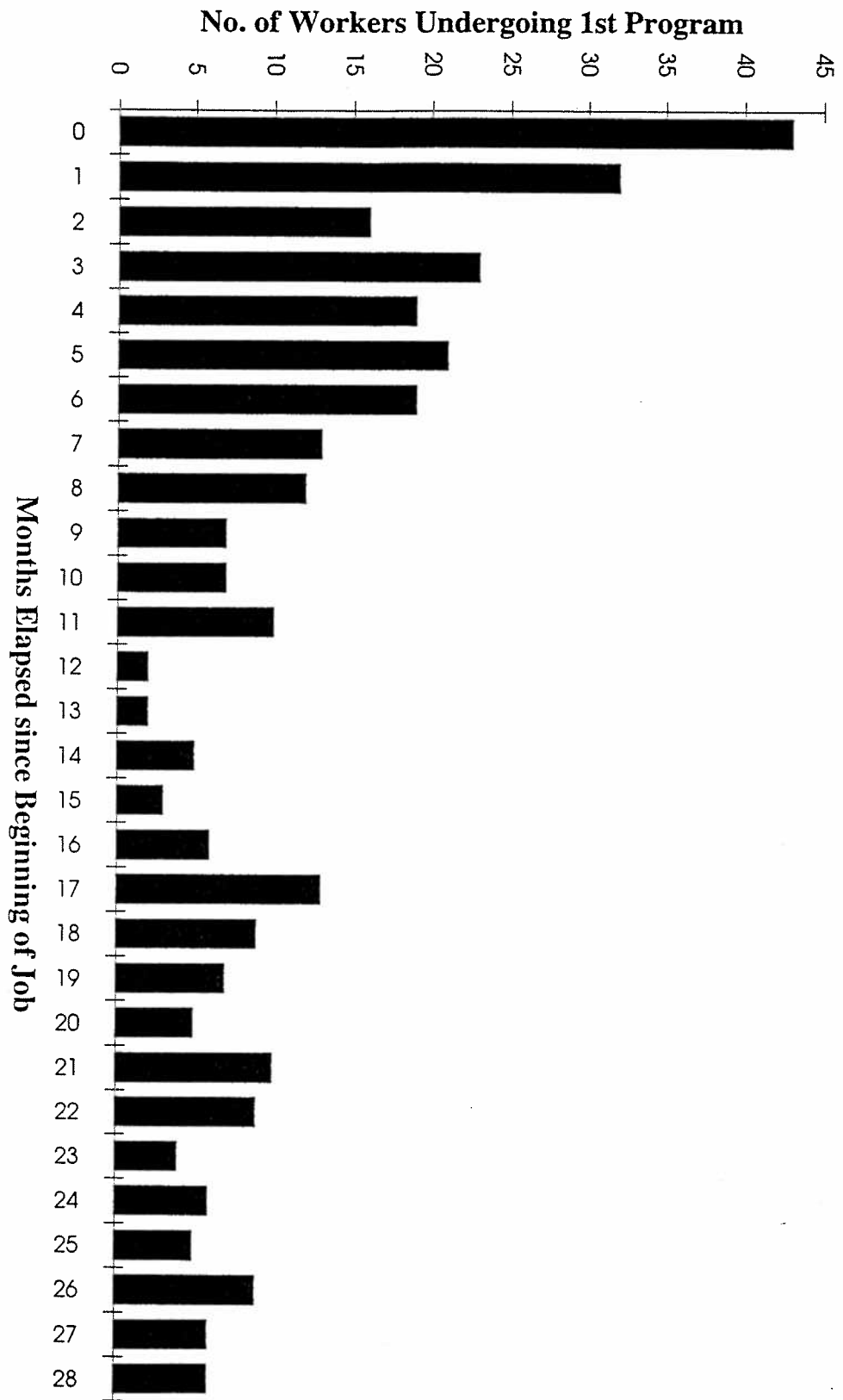


Figure 2. Timing of First Training Program (OFT)

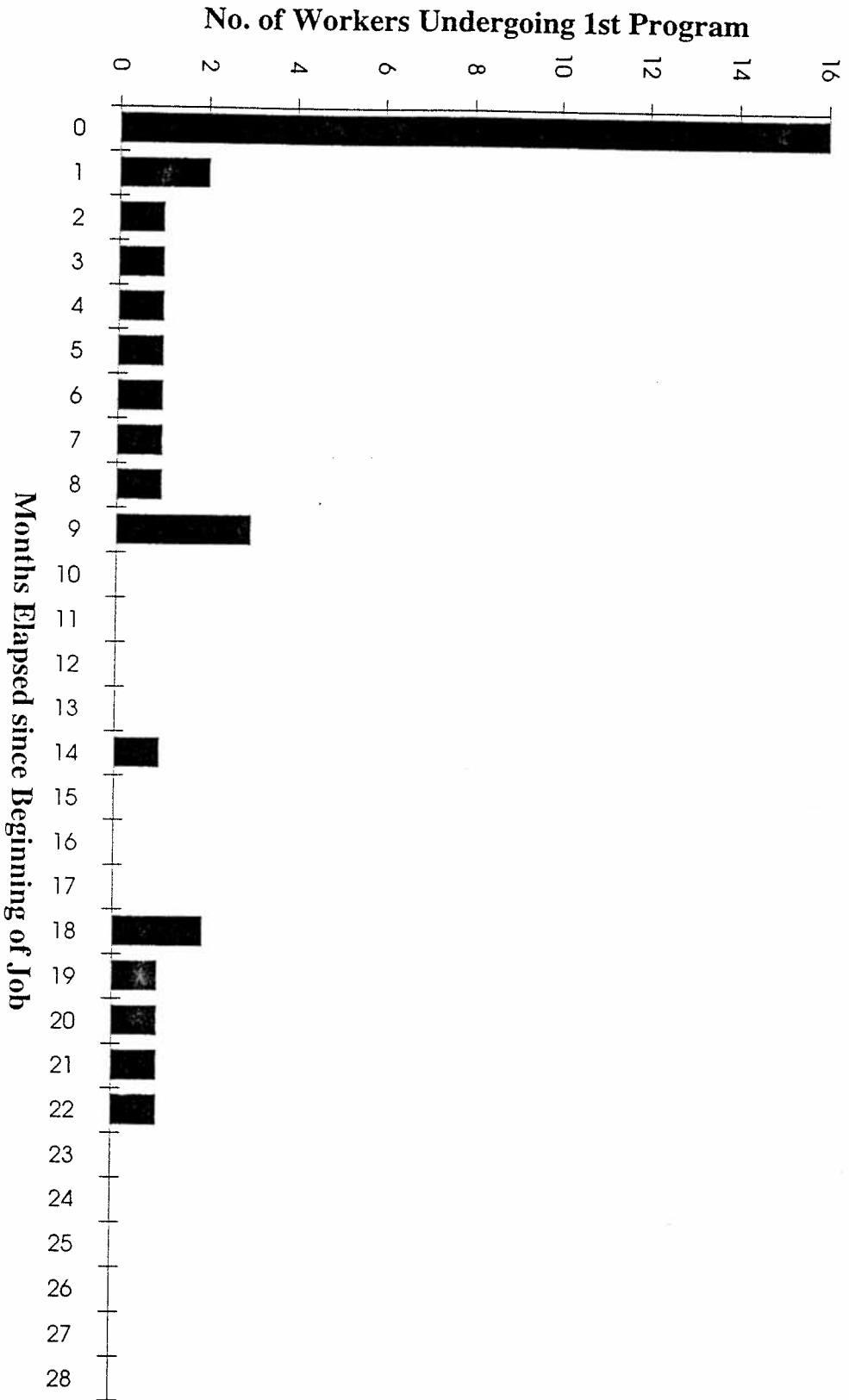
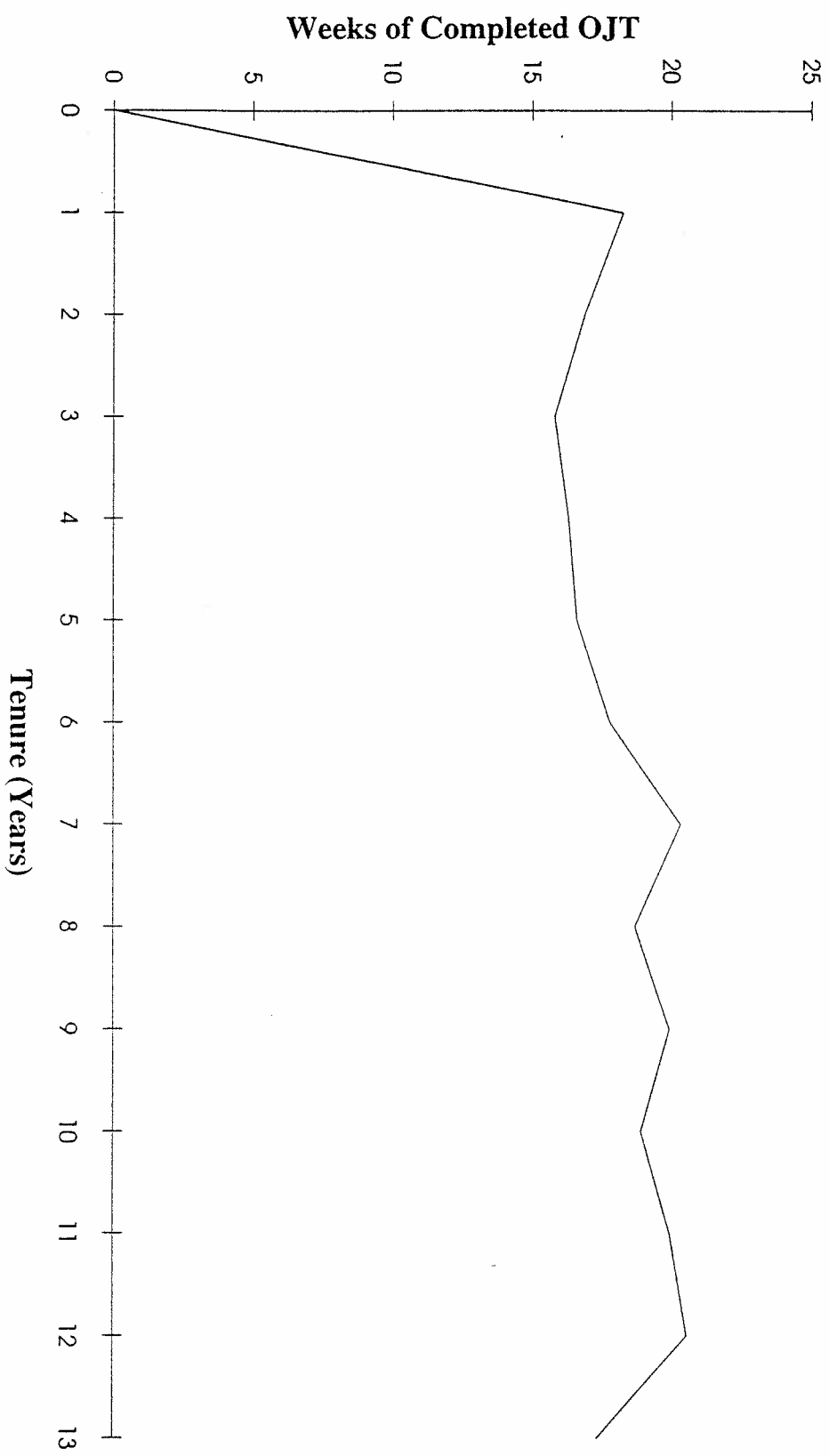


Figure 3. Timing of First Training Program (APP)

Figure 4. Tenure-Training Relationship



Industry-Specific Capital and the Wage Profile: Evidence from the
NLSY and the PSID

Troisième étude

I. Introduction

The extent to which wages rise with years of seniority with the same employer has been the subject of some controversy over the last few years (e.g. Topel (1991); Altonji and Shaikoiko (1987); Abraham and Farber (1987); Abowd, Kramarz and Margolis (1994)). Much of the debate surrounding this issue has focused on the appropriate econometric methods to be used to handle the issue of the endogeneity of the tenure variable. However, virtually no attention has been paid to the question of whether it is appropriate to decompose a worker's total labor market experience into only two components, tenure with the current employer and total prior experience (or total experience including tenure if one wants to obtain the latter's net effect). With data from the National Longitudinal Survey of Youth (NLSY) and from the Panel Study of Income Dynamics (PSID), it is shown that simply by adding total experience in the current industry as an additional explanatory variable, the net tenure effect vanishes almost completely. This suggests that past studies (most notably Topel (1991)) have overlooked an important factor in analyzing the effect of tenure on wages. It is worth noting that this result holds when the analysis is carried out either at the one-digit, two-digit, or three-digit level. Therefore, it seems that what matters most for the wage profile in terms of human capital is not so much firm-specificity but industry-specificity.¹

These results lead to the following basic conclusion: for these two samples of workers, the wage formation process seems to be very competitive with no solid evidence of rent sharing over the return on firm-specific capital. Or, put in the language of bargaining theory, there is little evidence that the workers represented in these two samples are paid much in excess of their outside option.²

¹ See also Neal (1993) for a different approach to assessing the degree of industry-specificity.

² For a theoretical model that describes the wage (price) formation process in the presence of renegotiation and relation-specific investments, see MacLeod and Malcomson (1993a,b). Under certain conditions and when proper allowance is made for the possibility of contract renegotiation, they show that employers need not offer their workers above market-clearing wages.

II. The Data

The National Longitudinal Survey of Youth data set surveyed 12,686 young males and females who were between the age of 14 and 21 in 1979³. It contains detailed employment histories of the respondents thereby permitting the construction of relatively error-free variables for tenure as well as for the total experience accumulated since the beginning of one's full time transition to the labor market. At the time this project was started, data were available from 1979 to 1991.

The people who were considered as having entered the labor market on a full-time basis were (i) those whose primary activity was either working full-time, on a temporary lay-off or looking actively for a job, (ii) those who did not return to school on a full-time basis within six years⁴ and (iii) those who had worked at least half the year since the last interview and who were working at least 20 hours per week. Individuals excluded from the sample are those younger than 18, those that had been in the military at any time, the self-employed, the ones whose jobs were part of a government program and the ones working without pay, those who were in the farming business and also all public sector employees. We are then left with 29,020 observations. Some summary statistics of the sample are provided in table 1.

Turning now to the Panel Study of Income Dynamics (PSID), the sample consists of heads of households aged 18 to 64 with positive earnings for the period spanning the years 1981-1987.⁵ The question of whether people have entered the labor market on a full-time

³ The response rate was at 71% in 1991.

⁴ The choice of six years as a cutoff point is arbitrary, and hence debatable. The idea is to exclude those that make "quasi-permanent" transitions and who might be considering returning to school a few years down the road. It seems reasonable to assume that few people would enter the labor market while planning to leave it in six years or more to go back to school. The same could not be said if we were considering a one to three year (say) horizon. In any event, the results were left unchanged if all school returners were excluded.

basis is obviously less of a concern for this sample of older workers. Summary statistics are provided in table 2.

III. Results

III.1 Basic Model.

Consider the following log wage equation:

$$\ln w_{ijt} = \beta_1 T_{ijt} + \beta_2 Exp_{it} + \beta_3 Expind_{ikt} + \alpha_t + \theta_{jt} + \gamma_{tk} + \epsilon_{ijt} \quad (1)$$

where w_{ijt} represents the real hourly wage of person i in job j in industry k at time t , T is tenure, Exp is total labor market experience and $Expind$ is total experience in the current industry. Specifically, the variable experience in the industry gives the consecutive number of years one has been in the same industry excluding tenure with the current employer. For example, if a worker leaves her first employer to take another job in the same industry, she adds experience in that industry, whereas if she takes another job in a different industry, her seniority in the industry is accordingly reset to zero. An implicit assumption is that the worker stays in the same industry throughout her employment relationship with a particular employer (the data show that this is not always the case: there are workers who change industry while not changing employers). Still, despite its shortcomings, this variable should give a fairly good idea of the industry effects embedded in the tenure variable. All other covariates, including squared terms for tenure, total experience and industry experience, are ignored for ease of presentation. As in previous studies, unobserved heterogeneity can be decomposed into an individual effect (α_i) and a job-match effect (θ_{jt}). The person-specific effect can be seen as representing unmeasured aspects of each individual's earning ability while the job-match component represents the unknown (to the econometrician) quality of the employment relationship stemming from search activity, for example. Both of these effects are assumed

⁵ This PSID extract was kindly supplied by Robert Valetta who used it in his paper with David Brownstone (Brownstone and Valetta (1993)) on the modeling of measurement error bias in wage equations.

to be time-invariant. Another unobserved heterogeneity component, γ_{ik} , which serves the same purpose as the job-match component, is added to represent the unobserved quality of the match between the individual and the industry in which he works.

As emphasized in the literature, the problem in estimating equation (1) with ordinary least-squares is that the unobserved components are likely to be correlated with tenure, total labor market experience and also, in our case, total experience in the industry. Those with high α 's may have enjoyed careers that were interrupted less frequently by unemployment spells, while better matches (high θ 's and high γ 's) are likely to be formed if you have more experience due to human capital and search effects. Also, tenure and total experience in the current industry are likely to be correlated with their corresponding match quality components. Dropping for the moment the assumption of time-invariant job-match and industry-match components, let's suppose that we can write them as

$$\theta_{ijt} = \psi_1 T_{ijt} + \psi_2 Exp_{ijt} + \omega_{ijt}$$

$$\gamma_{ikt} = \phi_1 Exp_{ikt} + \phi_2 Exp_{ind:kt} + \eta_{ikt} \quad (2)$$

where ω_{ijt} and η_{ikt} are assumed to be orthogonal to the regressors.⁶ The discussion above suggests that ψ_2 and ϕ_2 are positive. In the context of maximizing behavior on the part of a worker who faces a wage distribution, Topel(1991) argues that ψ_1 is negative once we control for experience, assuming there is a tenure effect. If there is no tenure effect, then ψ_1 equals zero. Presumably, the same sort of considerations apply to ϕ_3 . That is, provided that we control for total labor market experience, the quality of the match in the industry should be negatively correlated with the number of years one has been in the same industry. Substituting equation (2) into equation (1) we get

⁶ For simplicity, I also assume that total experience in the industry is not correlated with the job-match component while tenure is not correlated with the unobserved quality of the match in the industry. One could argue that more experience in the industry may help you find a better job-match because of superior information in comparison to a worker who has never worked in the industry.

$$\ln w_{ijt} = (\beta_1 + \psi_1)Ten + (\beta_2 + \psi_2 + \phi_2)Exp + (\beta_3 + \phi_3)ExpInd + \alpha_1 + \omega_{ijt} + \eta_{ikt} + \epsilon_{ijt} \quad (3)$$

We see from equation (3) that although we are interested in the β 's, using ordinary least-squares will produce estimates of composite effects and the regressors would still be correlated with α . To provide some correction for these problems, I use the instrumental variable (IV) methodology proposed by Altonji and Shakotko (1987)⁷. Tenure is instrumented with its deviations from job-match means whereas experience is instrumented with its deviations from individual means. In the same spirit, total experience in the industry is instrumented with its deviations from industry-match means. The instruments for tenure and experience in the industry are, by construction, uncorrelated with their respective match quality components, while the instrument for experience is, also by construction, uncorrelated with the individual component. First differences (as in Topel (91)) were not used to obtain a consistent estimate of the sum of the tenure and experience coefficients, because this would enhance any measurement errors present in the data as compared with using deviations from means. This is further justified by Topel's observation that much of the discrepancy between his results and those of Altonji and Shakotko stemmed from measurement errors pertaining to the tenure variable.

Finally, since the same individuals are followed over time, residuals will be serially correlated due to the presence of a fixed individual effect. To provide correction for this problem, all regressions are done using generalized least-squares under the assumption that the error term contains an individual-specific component.

III.2 Earnings Equation Estimates.

⁷ See also Finnite (1993) for an extension of the method of Altonji and Shakotko to the experience variable.

I now turn to the question of disentangling industry effects from purely firm-specific effects.⁸ Note that the analysis is carried out at the 1-digit, 2-digit, and 3-digit levels. If the tenure effect is entirely firm-specific, then it should not matter whether you change industry or not: the tenure coefficient should not budge at all. On the other hand, if a portion of the tenure effect reflects the specificity of the human capital acquired on the current job relative to the industry in which the firm operates, then adding such a control should decrease the coefficient on tenure. Using the NLSY data, the results shown in tables 3 and 4 seem to validate the latter explanation: in the GLS specification of the 1-digit case (column (4) in table 3), over 50% of the effect of tenure is accounted for by industry effects. The tenure effect further decreases in the 2-digit case and disappears completely in the 3-digit case (columns (3) and (2), respectively). Once the instrumental variable specification is adopted (see table 4), there is no evidence of a substantial and statistically significant positive tenure effect, even at the 1-digit level. The entire effect is picked up by the variable representing experience in the industry. Note that total experience is only slightly affected by the inclusion of the new variable.

Table 5 provides a breakdown by occupational category. It could be that for managers or professionals, firm-specific investments are more important and that firms would prefer to pay these highly skilled individuals above their outside option rather than to face the prospect of losing them, especially if these workers are in short supply. The results indicate that there is a return to tenure for the category consisting of professionals, technical workers, managers and administrators. In the GLS specification, even after adding experience in the industry, the tenure effect is substantial and significant. In fact, in the IV-GLS specification, the effect is larger after adding the control for industry effects. For the next three categories of workers, the results are similar to those obtained for the total sample: the industry effect is large and significant while the firm effect is not. Again, it is interesting to note that total experience is

⁸ All results are obtained using the weighted samples.

not markedly affected by the added variable. Total experience and experience in the industry really do seem to provide complementary explanatory power to the wage formation process.

III.3 Comparison with Data from the PSID

Given that the NLSY is composed of young persons making their transition to the labor market (the oldest individuals in 1991 were 33 years old), it could very well be that the results above are peculiar to that data set. To be more precise, assuming that skills which are truly firm-specific are associated more with older workers than their younger counterparts, then the results above may not hold with a sample of workers who have more mature careers.⁹ Therefore, to check whether results are robust across data sets, we have estimated the same type of equation with data from the Panel Study of Income Dynamics (PSID). The sample consists of heads of households aged 18 to 64 with positive earnings for the years spanning the period 1981 to 1987. Results are shown in tables 6 and 7. As shown in table 6, the impact of including the additional explanatory variable is qualitatively the same as in the case of the NLSY. With GLS and IV-GLS, adding the control for industry-specific capital has the effect of reducing the tenure coefficient to close to zero. Results by occupational categories are shown in table 8. Contrary to the results obtained with the NLSY, there is no evidence of a return to tenure for professionals, managers and administrators. For clerical and unskilled workers, there is even evidence of a negative net return to firm seniority. However, for service workers, the estimated return to tenure is sizeable and significant, which is not the case for the NLSY sample. Interestingly, a common denominator of the results with the PSID is that the total experience coefficients are much smaller in comparison with those estimated with the NLSY. Thus, it appears that as workers' careers evolve, the skills they acquire are more narrowly defined. Of course, as workers gain experience in the labor market,

⁹ If we admit that investment in firm-specific skills is complementary to the quality of the match, then it follows that older workers who have had more time to sample the job offer distribution would be better candidates for such investments. On the question of complementarity between match quality and firm-specific capital, see Jovanovic (1979).

the search process leads them to more satisfactory matches. Presumably then, the opportunities for investment in more specific skills are enhanced in these "better" jobs.

IV. Conclusion

With data from the NLSY, it has been shown that by controlling for workers who change industry when they change jobs, the tenure effect is reduced by over 50% at the 1-digit level if generalized least-squares are used, while it disappears completely at the 3-digit level, whether it be with GLS or with the instrumental variable methodology borrowed from Altonji and Shakotko (1987). These results generally hold when I subdivide the sample by occupational categories, although there is some evidence of a tenure effect for professionals, technical workers and managers. Also, by using data from the PSID, I have shown that these results are robust across data sets and are not peculiar to the young NLSY workers making their full-time transition to the labor market. The basic conclusion to be drawn from these results is that the wage formation process seems to be very competitive for most of the workers in these two data sets.

Having established that the average tenure slope is close to zero, an interesting avenue for future research would be to determine whether there is substantial variance in the slopes. Results by occupational categories suggest that there is some variability in the estimated tenure slopes. However, all research up to now has used the assumption of fixed parameters. A more refined analysis would call upon the use of a random coefficient model to study the covariance structure of a log-earnings equation with proper allowance made for the randomness in the tenure and experience profiles.¹⁰ Then we could determine with more confidence whether there is evidence of workers receiving wages above their outside option.

¹⁰ See Hause (1980) for a study linking the covariance of earnings to the theory of human capital.

Table 1

MEAN SAMPLE CHARACTERISTICS (Weighted)-NLSY

	1-Digit	2-Digit	3-Digit
Real Hourly Wage (\$1979)	5.75	41.7	2.44
Hours Worked	5.82	12.44	12.6
Tenure	5.82	12.44	12.6
Experience	5.82	12.44	12.6
Years in School	12.44	12.6	44.7
Percentage Nonwhite	12.6	44.7	45.4
Percentage Married	44.7	45.4	25.1
Percentage Female	45.4	25.1	29,020
Number of Observations	29,020	5,649	13,590
Number of Individuals	5,649	5,649	5,649
Number of Jobs	5,649	5,649	5,649
Number of Job Changes Involving:	5,649	5,649	5,649
a) A Change of Industry	5,085	5,613	6,552
b) No Change of Industry	2,818	2,290	1,351

TABLE 2

MEAN SAMPLE CHARACTERISTICS-PSID

	1-Digit	2-Digit	3-Digit
Real Hourly Wage (\$1979)	8.39	40.0	9.4
Hours Worked	21.8	12.8	10.3
Tenure	21.8	12.8	10.3
Experience	21.8	12.8	10.3
Years in School	12.8	10.3	66.9
Percentage Nonwhite	12.8	10.3	66.9
Percentage Married	10.3	66.9	18.6
Percentage Female	10.3	66.9	18.6
Age	18.6	66.9	34.6
Number of Observations	34.6	66.9	15,480
Number of Individuals	15,480	2,750	2,750
Number of Jobs	2,750	2,750	2,750
Number of Job Changes Involving:	2,750	2,750	2,750
a) A Change of Industry	1,033	1,143	1,337
b) No Change of Industry	1,071	961	767

EARNINGS FUNCTIONS ESTIMATES- NLSY: INDUSTRY VS TENURE EFFECT
 (Dependent Variable: Log of Real Hourly Labor Income (\$1979))

TABLE 3

Independent Variable	(1)	(2)	(3)	(4)
Tenure	0.0436	0.0013	0.0121	0.0173
Tenure Squared	(0.0029)	(0.0064)	(0.0049)	(0.0044)
Tenure Squared	-0.0030	0.0001	-0.0007	-0.0011
Experience in Current Industry(3-digit)	-	0.0480	-	-
Experience in Current Industry(2-digit)	-	-	0.0403	-
Experience in Current Industry(1-digit)	-	-	-	0.0364
Experience in Industry Squared(3-digit)	-	-0.0034	-	-
Experience in Industry Squared(2-digit)	-	-	-0.0027	-
Experience in Current Industry(1-digit)	-	-	-	-
Experience in Industry Squared(1-digit)	-	-	-	-0.0025
Total Experience	0.0779	0.0720	0.0695	0.0683
Total Experience Squared	(0.0044)	(0.0045)	(0.0045)	(0.0045)
Industry Dummies	YES	YES	YES	YES
Occup. Dummies	YES	YES	YES	YES
R-Squared	0.8001	0.8018	0.8015	0.8014

Notes-Other covariates include education dummies, race, sex, regional, union coverage, marital status, occupation, industry and year dummies. Standard errors are shown in parentheses (rounded to 0.0001 when smaller). Sample size is 29,020.

TABLE 4

EARNINGS FUNCTIONS ESTIMATES-NLSY: INDUSTRY VS TENURE EFFECT
 (Dependent Variable: Log of Real Hourly Labor Income (\$1979))

Independent Variable	(1)	(2)	(3)	(4)
Tenure	0.022	-0.0093	-0.0012	0.0028
Tenure Squared	-0.0020	0.0004	-0.0002	-0.0004
Experience in Current Industry(3-digit)	-	0.0354	-	-
Experience in Current Industry(2-digit)	-	-	0.0295	-
Experience in Current Industry(1-digit)	-	-	-	0.0266
Experience in Industry Squared(3-digit)	-	-0.0026	-	-
Experience in Industry Squared(2-digit)	-	-	-0.0021	-
Experience in Industry Squared(1-digit)	-	-	-	-0.0020
Total Experience	0.0955	0.0914	0.0894	0.0883
Total Experience Squared	(0.0049)	(0.0050)	(0.0051)	(0.0052)
Indust. Dummies	YES	YES	YES	YES
Occup. Dummies	YES	YES	YES	YES
R-Squared	0.7945	0.7936	0.7949	0.7949

Notes-Unshown covariates are the same as those in table 3. Standard errors in parentheses. Sample size is 29,020.

INDUSTRY EFFECT (1-digit) VS TENURE EFFECT BY OCCUPATIONS-NLSY
 (Dependent Variable: log of Real Hourly Wages (\$1979))

PROFESSIONALS, TECHNICAL WORKERS, MANAGERS AND ADMINISTRATORS

CLERICAL AND UNSKILLED WORKERS				
Independent Variable	(1)	(2)	(3)	(4)
Tenure	0.0312	0.0159	0.0067	-0.0067
Tenure Squared	(0.0066)	(0.0090)	(0.0103)	(0.0136)
Tenure Squared	-0.0016	-0.0008	0.0006	0.0013
Total Experience	0.0926	0.1009	0.0839	0.0928
Total Experience	(0.0007)	(0.0008)	(0.0011)	(0.0012)
Total Exp. Squared	(0.0107)	(0.0134)	(0.0111)	(0.0141)
Total Exp. Squared	-0.0027	-0.0029	-0.0021	-0.0025
Experience in Industry	(0.0007)	(0.0007)	(0.0007)	(0.0012)
Experience in Industry	-	-	0.0328	0.0302
Exp. in Industry Squared	-	-	(0.0105)	(0.0138)
Exp. in Industry Squared	-	-	-0.0028	-0.0026
R-Squared	0.6808	0.6619	0.6821	0.6630
Number of Observations	6,470			
PROFESSIONALS, TECHNICAL WORKERS, MANAGERS AND ADMINISTRATORS				
Independent Variable	(1)	(2)	(3)	(4)
Tenure	0.0264	0.0219	0.0159	0.0291
Tenure Squared	(0.0065)	(0.0082)	(0.0089)	(0.0114)
Tenure Squared	-0.0028	-0.0030	-0.0019	-0.0032
Total Experience	0.0980	0.1141	0.0689	0.1193
Total Experience	(0.0006)	(0.0007)	(0.0009)	(0.0010)
Total Exp. Squared	(0.0113)	(0.0132)	(0.0090)	(0.0141)
Total Exp. Squared	-0.0024	-0.0027	-0.0015	-0.0028
Experience in Industry	(0.0006)	(0.0007)	(0.0006)	(0.0007)
Experience in Industry	-	-	0.0155	-0.0121
Exp. in Industry Squared	-	-	(0.0094)	(0.0120)
Exp. in Industry Squared	-	-	-0.0013	0.0003
R-Squared	0.6794	0.6692	0.6783	0.6698
Number of Observations	6,788			

TABLE 5-continued

CRAFTSMEN AND KINDRED WORKERS, OPERATIVES AND LABORERS

Independent Variable	(1)	(2)	(3)	(4)
Tenure	0.0430	0.0157	0.0170	-0.0071
Tenure Squared	(0.0043)	(0.0059)	(0.0067)	(0.0094)
Tenure Squared	-0.0029	-0.0014	-0.0008	0.0006
Total Experience	(0.0004)	(0.0005)	(0.0007)	(0.0008)
Total Experience	0.0825	0.1068	0.0741	0.0990
Total Exp. Squared	(0.006)	(0.0080)	(0.0068)	(0.0083)
Total Exp. Squared	-0.0023	-0.0030	-0.0019	-0.0026
Experience in Industry	(0.0004)	(0.0004)	(0.0004)	(0.0005)
Experience in Industry	-	-	0.0349	0.0306
Exp. in Industry Squared	-	-	(0.0068)	(0.0092)
Exp. in Industry Squared	-	-	-0.0025	-0.0026
R-Squared	0.7960	0.7840	0.7971	0.7848
Number of Observations	11,779			

SERVICE WORKERS

Independent Variable	(1)	(2)	(3)	(4)
Tenure	0.0420	0.0311	-0.0032	0.0047
Tenure Squared	(0.0080)	(0.0115)	(0.0121)	(0.0183)
Tenure Squared	-0.0041	-0.0041	-0.0007	-0.0020
Total Experience	(0.0008)	(0.0009)	(0.0013)	(0.0017)
Total Experience	0.0502	0.0682	0.0340	0.0606
Total Exp. Squared	(0.0114)	(0.0149)	(0.0116)	(0.0154)
Total Exp. Squared	-0.0001	-0.0001	0.0004	0.0001
Experience in Industry	(0.0007)	(0.0008)	(0.0008)	(0.0008)
Experience in Industry	-	-	0.0654	0.0353
Exp. in Industry Squared	-	-	(0.0121)	(0.0177)
Exp. in Industry Squared	-	-	-0.0040	-0.0023
R-Squared	0.6878	0.6767	0.6921	0.6788
Number of Observations	3,969			

Notes-Unknown covariates are the same as those in table 3. Standard errors in parentheses.

TABLE 6

EARNINGS FUNCTIONS ESTIMATES-PSID: INDUSTRY VS TENURE EFFECT
 (Dependent Variable: Log of Real Hourly Labor Income (\$1979))

Independent Variable	(1)	(2)	(3)	(4)
Tenure	0.0142	0.0002	0.0017	0.0020
Tenure Squared	(0.0012)	(0.0023)	(0.0020)	(0.0019)
Experience in Current Industry(3-digit)	-	0.0214	-	-
Experience in Industry Squared(3-digit)	-	(0.0024)	-	-
Experience in Current Industry(2-digit)	-	-	0.0229	-
Experience in Industry Squared(2-digit)	-	-	(0.0022)	-
Experience in Current Industry(1-digit)	-	-	-	0.0247
Experience in Industry Squared(1-digit)	-	-	-	(0.0021)
Potential Experience	0.0210	0.0182	0.0173	0.0170
Potential Experience Squared	(0.0017)	(0.0017)	(0.0017)	(0.0017)
Industry Dummies	YES	YES	YES	YES
Industry Dummies Squared	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Occup. Dummies	YES	YES	YES	YES
Occup. Dummies Squared	(0.0004)	(0.0004)	(0.0004)	(0.0004)
R-Squared	0.6679	0.6790	0.6808	0.6751

Notes-Other covariates include education in years, race, sex, regional, union coverage, marital status, occupation, industry and year dummies. Standard errors are shown in parentheses (rounded to 0.0001 when smaller). Sample size is 15,480.

TABLE 7

EARNINGS FUNCTIONS ESTIMATES-PSID: INDUSTRY VS TENURE EFFECT
 (Dependent Variable: Log of Real Hourly Labor Income (\$1979))

Independent Variable	(1)	(2)	(3)	(4)
Tenure	0.0191	0.0051	0.0037	0.0041
Tenure Squared	(0.0018)	(0.0031)	(0.0028)	(0.0027)
Tenure Squared	-0.0003	-0.0001	0.0001	0.0001
Experience in Current Industry(3-digit)	-	0.0202	-	-
Experience in Current Industry(3-digit) Squared	-	(0.0032)	-	-
Experience in Current Industry(3-digit)	-	-0.0003	-	-
Experience in Current Industry(2-digit)	-	-	0.0247	-
Experience in Current Industry(2-digit) Squared	-	-	(0.0029)	-
Experience in Current Industry(2-digit)	-	-	-0.0004	-
Experience in Current Industry(1-digit)	-	-	-	0.0258
Experience in Current Industry(1-digit) Squared	-	-	-	(0.0029)
Experience in Industry	-	-	-	-0.0004
Experience in Industry Squared(1-digit)	-	-	-	(0.0001)
Potential Experience	0.0189	0.0164	0.0154	0.0155
Potential Experience Squared	(0.0018)	(0.0018)	(0.0018)	(0.0018)
Potential Experience Squared	-0.0004	-0.0004	-0.0004	-0.0004
Industry Dummies	YES	YES	YES	YES
Occup. Dummies	YES	YES	YES	YES
R-Squared	0.6405	0.6382	0.6396	0.6451

Notes-Unshown covariates are the same as those in table 6. Standard errors in parentheses. Sample size is 15,480.

TABLE 8

INDUSTRY EFFECT (1-digit) VS TENURE EFFECT BY OCCUPATIONS-PSID
 (Dependent Variable: log of Real Hourly Labor Income (\$1979))

PROFESSIONALS, TECHNICAL WORKERS, MANAGERS AND ADMINISTRATORS

Independent Variable	(1)	(2)	(3)	(4)
Tenure	0.0097	0.0126	0.0055	0.0061
Tenure Squared	(0.0021)	(0.0028)	(0.0029)	(0.0039)
Potential Experience	(0.0001)	(0.0001)	(0.0001)	(0.0002)
Potential Experience Squared	0.0340	0.0321	0.0315	0.0288
Potent. Exp. Squared	(0.0033)	(0.0035)	(0.0034)	(0.0037)
Experience in Industry	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Experience in Industry Squared	-	-	0.0118	0.0148
Exp. in Industry Squared	-	-	0.0001	-0.0001
R-Squared	0.6960	0.6640	0.7002	0.6666
Number of Observations	5,228			

CLERICAL AND UNSKILLED WORKERS

Independent Variable	(1)	(2)	(3)	(4)
Tenure	0.0172	0.0152	-0.0074	-0.0265
Tenure Squared	(0.0048)	(0.0061)	(0.0074)	(0.0094)
Potential Experience	(0.0002)	(0.0002)	(0.0002)	(0.0003)
Potential Experience Squared	0.0081	0.0036	0.0056	-0.0007
Potent. Exp. Squared	(0.0045)	(0.0048)	(0.0046)	(0.0049)
Experience in Industry	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Experience in Industry Squared	-	-	0.0391	0.0568
Exp. in Industry Squared	-	-	-0.0009	-0.0016
R-Squared	0.6632	0.6225	0.6785	0.6287
Number of Observations	1,382			

TABLE 8-continued

CRAFTSMEN AND KINDRED WORKERS, OPERATIVES AND LABORERS

Independent Variable	(1)	(2)	(3)	(4)
Tenure	0.0160	0.0179	-0.0026	-0.0019
Tenure Squared	(0.0017)	(0.0025)	0.0004	(0.0041)
Potential Experience	0.0124	0.0117	0.0090	0.0091
Potent. Exp. Squared	(0.0001)	(0.0001)	(0.0001)	(0.0002)
Experience in Industry	(0.0001)	(0.0001)	0.0297	0.0298
Exp. in Industry Squared	-	-	(0.0031)	(0.0042)
R-Squared	0.6541	0.6300	0.6658	0.6341
Number of Observations	7,200			

SERVICE WORKERS

Independent Variable	(1)	(2)	(3)	(4)
Tenure	0.0370	0.067	0.0267	0.0605
Tenure Squared	(0.0072)	(0.0091)	(0.0114)	(0.0184)
Potential Experience	(0.0003)	(0.0004)	(0.0005)	(0.0006)
Potent. Exp. Squared	-0.0004	-0.0092	-0.0079	-0.0140
Experience in Industry	(0.0001)	(0.0002)	(0.0001)	(0.0002)
Exp. in Industry Squared	-	-	0.0249	0.0194
R-Squared	0.4629	0.4529	0.4839	0.4584
Number of Observations	1,610			

Notes-All regressions include industry and occupation dummies. Other unshown covariates are the same as those in table 6. Standard errors in parentheses.

Matching, Human Capital, and the Covariance Structure of Earnings

Quatrième étude

Flinn (1986) provides a statistical test of the restrictions imposed by the theory on the covariance structure of the log of earnings. To account for the presence of unobserved job-match heterogeneity, Flinn assumes that the residual from the log-earnings function contains a match-specific component, fixed within a job but different across all possible matches. Although the restrictions cannot be rejected by the data, it should be noted that his sample is very small (the sample is made of 248 individuals and is further divided into four distinct subsamples, the smallest of which has 36 workers) and short (three years). Therefore, as a first step in the analysis, it seems appropriate to assess the validity of the covariance structure as implied (in its simplest form) by the matching theory with a longer and larger sample. Next, I study the evolution of the variance of the job-match component across jobs and within jobs. Depending upon the nature of the matching process that one has in mind, we get different predictions concerning the evolution of that component. More particularly, it makes a difference whether jobs are considered as pure experience goods or pure inspection goods. In the former case, nothing can be learned about the characteristics of the job without actually gaining some experience in it. Consequently, we should first observe an increase in the

Two competing explanations for the existence of a positive return to tenure are human capital and matching. In short, both theories yield the prediction that the conditional mean of wages should rise with tenure. In this paper, I intend to compare and distinguish these two theories empirically by focusing on the implications in terms of the covariance structure. To do so, I use the residuals from a log-earnings equation using data from the National Longitudinal Survey of Youth (NLSY) for the period spanning the years 1979-1991. This data set contains detailed employment histories for young people who were aged 14 to 21 at the end of 1978 and who are making their transition to the labor market. Results are presented for the entire sample as well as for subsamples representing different schooling groups. The econometric methodology to be used is borrowed and adapted to unbalanced data from the method of moments estimator of Gallant and Jorgenson (1979).

I. INTRODUCTION

variance of the job-match component as all worker-firm matches start with the same prior information and workers are paid the same wage. Hence, in theory, the variance at the start should be zero. With each new pieces of information, both parties get a clearer idea as to the quality of their match and the variance increases up to the point where the selection process makes ill matched workers leave their job, thereby truncating the bottom part of the distribution. Then, assuming that the distribution is unbounded, we should see a decline in the variance of the match quality distribution followed by an increase as we reach the right-hand tail. On the other hand, if jobs are pure inspection goods, then everything can be learned merely by contacting a potential employer, in which case we find ourselves with a typical search model (e.g. Burdett (1978)). Both the variance across jobs and within jobs should decline from the start up to the point where, again, the upper tail of the match quality distribution is reached, in which case the variance starts increasing.

Since the predictions of the matching model are derived under the assumption of no acquisition of human capital, which means that workers are assumed to be sampling from a stable wage offer distribution, the results I get suggest that this assumption is likely to be violated. More precisely, it is difficult to find evidence of the patterns predicted by the theory, except for the evolution within jobs. The next step, then, is to focus our attention on the predictions of human capital theory, with the idea of trying to isolate a prediction that is not shared by the theory of job matching. With this goal in mind, I extend Hauske's (1980) approach by introducing the concept of job into his framework. Under certain specifications, Hauske was able to find evidence for a U-shape pattern of the time profile of variances of the log of earnings for a sample of young Swedes.¹ According to the theory of human capital, this pattern emerges because of a negative covariance between the individual-specific intercept and slope parameters. Those with a lower intercept invest relatively more in human capital by accepting a lower starting salary. Eventually, though, their earnings overtake those of workers who chose not to invest as much. Subsequent work on the theory of matching by

¹ See Mincer (1974), especially as it pertains to the overtaking concept.

The paper is organized as follows. Section II sketches the main results from the theory of matching with an emphasis on the predictions pertaining to the variance of the job-match component. Section III follows with a discussion on the theory of human capital and the covariance structure of earnings. Section IV focuses on the empirical implementation of the

All this is not to say that matching is not an important phenomenon. On the contrary, it is shown that even the simplest form of covariance structure implied by the theory of matching fits the data quite well. I would rather want to emphasize the complementarity between matching and human capital. This complementarity makes difficult the isolation of predictions specific to the theory of matching. But, when some control for the acquisition of human capital through on-the-job training is provided, predictions stemming from search/matching considerations are more easily highlighted.

By introducing jobs into the analysis, it is possible to isolate the key prediction from human capital theory that there should be a tradeoff between the job-specific intercept and the tenure slope. All things being equal, those who start out with a lower salary invest more in human capital and consequently should have a steeper slope. In contrast, the pure theory of matching does not predict such a tradeoff within jobs. In fact, in his work on the complementarity between the quality of a match and firm-specific investment, Jovanovic (1979b) shows that a better match should, *ceteris paribus*, involve more investment. Therefore, the only way that the theory of matching can account for the tradeoff between the job-specific slope and the intercept is through human capital considerations.

Jovanovic (1984) led to the prediction of a U-shape pattern as well when we are studying the wage-experience profile of individuals. The basic reason stems from the fact that individuals would be prepared to leave a job in which pretty much all learning has been done for another which initially pays less but which offers the possibility of growth. Therefore to have a chance of potentially discriminating between the two theories, we need a framework which allows us to circumvent this identification problem.

models discussed in sections II and III, including discussions on the data and on the econometric methodology. Section V presents the results, and at the end of the section, the predictions stemming from matching/search considerations are reexamined in a context where some control is provided for the accumulation of human capital. Section VI concludes the paper.

II. THE THEORY OF MATCHING AND ITS EMPIRICAL IMPLICATIONS

Originating with Stigler's model of search, economists have developed tools for analyzing the way in which individual units such as firms or workers gather and process information. For example, McCall (1970) pioneered the development of the theory of sequential job search in which individuals draw (at a cost) from a wage distribution F and then decide to either accept the offer in hand or to reject it and sample once again. By applying the principle of optimality, each individual will have a decision rule which is characterized by a reservation offer, say w^* , such that any wage offer above this threshold will be accepted and, conversely, all offers below w^* will be rejected.

As emphasized by Rothschild (1973), one problem with this model is the existence of the wage distribution itself in equilibrium. If we assume that the population of workers is homogeneous, then no wages less than w^* would be observed and profit maximizing firms would see no point in offering wages above w^* . Therefore, the wage distribution would collapse to a trivial one where all the probability mass is concentrated at w^* .

The theory of matching offers an answer to this problem by allowing, in an equilibrium context, the existence of both a non-trivial wage distribution and an optimal search strategy by individuals. The classic references on the theory of matching are Jovanovic (1979a, 1979b, 1984). The model described below is a discrete-time version of the one contained in Jovanovic's 1984 contribution. Note that I will focus mainly on the information processing part of the model as the predictions we are interested in stem from this learning mechanism.

II.1 The Theory in a Nutshell

124

Let's assume that θ , the quality of a match between a worker and a firm, can be characterized as having been drawn from some known normal distribution. The worker and the firm come from homogeneous populations. The match parameter θ can be seen as representing the marginal productivity of the worker in the firm. The main theme recurring in the analysis is that both the firm and the worker will gain information on the "true" quality of the match first by merely contacting a potential employer and also by observing the evolution of the worker's output over time. Both sources of information are noise ridden and each party is assumed to use Bayes' law to update its beliefs on θ . To be more specific, the assumptions can be summarized as follows:

Assumptions:

(i) The prior distribution of θ is $N(\gamma, 1/\tau_0)$, where τ_0 represents the precision of the distribution. This distribution is stable over time. All workers and firms share this prior distribution.

(ii) The outcome of the initial screen is denoted as $m_0 = \theta + \epsilon_0$, with ϵ_0 being normally distributed with mean 0 and precision τ_m . When $\tau_m \rightarrow \infty$, all information about the quality of a potential match can be learned merely by contacting an employer. In other words, jobs are pure search or inspection goods. When $\tau_m = 0$, nothing can be learned from the match without gaining some experience in it. In that case, jobs are pure experience goods.

(iii) The output at each period is $x_t = \theta + \epsilon_t$, with $\epsilon_t \sim N(0, 1)$ and $E(\epsilon_t \epsilon_s) = 0$, for all $t \neq s$. At each period t , both parties observe the worker's output and both are equally well informed about the quality of the match. They are assumed to use Bayes' law to update their beliefs on the perceived quality of the match.

a) The wage paid at period t , conditional on the outcome of the screen and on the sequence of outputs is equal to:

PROPOSITION 1

The time sequence of events is thus the following: (i) a firm and a worker make contact; (ii) upon contact, they both observe a noisy signal m_0 which provides information on the true value of the match; (iii) the firm makes an offer to the worker on the basis of that information, and the worker accepts; (iv) at each subsequent period t , both parties update their belief on the quality of the match by making use (through Bayes' Law) of the information provided by the output x_t ; (v) the worker makes a stay/quit decision and draws another match parameter, after which the same sequence resumes. Given the assumptions above, we can summarize their implications in the following proposition:

(viii) Workers accept the first offer they receive.

(vii) Embedded in (vi) is the assumption that workers maximize the expected present value of their lifetime earnings.

(vi) Firms are assumed to pay workers their expected marginal product. This is the equilibrium contract of Iovanovic (1979a) derived under the assumption that firms bid for workers by offering them lifetime contracts which they are assumed to honor ex post.

(v) All draws from the prior distribution are independent from one another. The quality of the present match provides no information on the quality of potential matches. There is no recall (or if recalls are allowed, the quality of the new match is independent of the old one).

(iv) At each period, each worker receives an offer with probability λ .

$$E(\theta | m_0, x_1, x_2, \dots, x_t) = \frac{\int \theta \gamma^{m_0 + t} x_1^{t_1} x_2^{t_2} \dots x_t^{t_t} d\theta}{\int \gamma^{m_0 + t} x_1^{t_1} x_2^{t_2} \dots x_t^{t_t} d\theta} = w(t) \quad (2)$$

with conditional variance given by:

$$V(\theta | m_0, x_1, x_2, \dots, x_t) = \frac{\int \theta^2 \gamma^{m_0 + t} x_1^{t_1} x_2^{t_2} \dots x_t^{t_t} d\theta}{\int \gamma^{m_0 + t} x_1^{t_1} x_2^{t_2} \dots x_t^{t_t} d\theta} - S(t)^2 \quad (2)$$

b) The current wage conditional on the previous wage is:

$$E(w(t) | w(t-1) = w) = w; \quad V(w(t) | w(t-1) = w) = S(t)S(t-1). \quad (3)$$

c) The initial wage (w_0) offer distribution is characterized by:

$$w_0 \sim N\left(\gamma, \frac{\tau}{m} \tau_\theta(\tau_\theta + \tau^m)\right) \quad (4)$$

d) The unconditional distribution of $w(t)$ is normal with moments equal to:

$$E(w(t)) = \gamma; \quad V(w(t)) = \frac{\tau^{m+\tau_\theta}}{1} - \frac{\tau^{t+\tau^m+\tau_\theta}}{1}. \quad (5)$$

PROOF

a) Let $f(\cdot)$ be the posterior distribution, $g(\cdot)$ the likelihood function, and $h(\cdot)$ the prior distribution of θ . From Bayes' Law we know that:

$$f(\theta | m_0, x_1, x_2, \dots, x_t) \propto g(m_0, x_1, x_2, \dots, x_t) h(\theta) \quad (6)$$

(7) with $g(m_0, x_1, x_2, \dots, x_t | \theta) \propto \exp\left(-\frac{1}{2} \sum_{i=1}^{t-1} (x_i - \theta)^2\right) \frac{1}{\sqrt{2\pi}} (m_0 - \theta)^2$;

(8) $h(\theta) \propto \exp\left(-\frac{1}{2} (\theta - \lambda)^2\right)$

If we develop equation (6) given that we know $g(\cdot)$ and $h(\cdot)$, then we obtain after some algebra:

(9)
$$f(\theta | m_0, x_1, x_2, \dots, x_t) \propto \exp\left(-\frac{1}{2} (\theta - \lambda)^2\right) \frac{1}{\sqrt{2\pi}} (m_0 - \theta)^2 \left(\sum_{i=1}^{t-1} x_i\right)^t$$

which is the kernel of a normal distribution with conditional expectation given by

(10)
$$w(t) = E(\theta | m_0, x_1, x_2, \dots, x_t) = \frac{\sum_{i=1}^{t-1} x_i + m_0}{t}$$

with conditional variance equal to

(11)
$$S(t) = V(\theta | m_0, x_1, x_2, \dots, x_t) = \frac{1}{t}$$

b) Rearranging equation (10), we can express $w(t)$ as:

(12)
$$w(t) = [S(t)/S(t-1)]w(t-1) + [1-S(t)/S(t-1)](\theta + \epsilon^t)$$

Hence,

$$E(w(t)|w(t-1)=w) = [S(t)/S(t-1)]w + [1-S(t)/S(t-1)]E(\theta+\epsilon_t|w(t-1)=w) \quad (13)$$

For the conditional variance, we have:

$$E((w(t)-w)^2|w(t-1)=w) = E([S(t)/S(t-1)]^2 w^2 - [1-S(t)/S(t-1)](\theta+\epsilon_t)^2 | w(t-1)=w) \\ = (1-S(t)/S(t-1))^2 (S(t-1)+1) = S(t)^2 \frac{(t-1)^{m_0+\tau_\theta}}{t^{m_0+\tau_\theta}} = S(t)S(t-1) \quad (14)$$

c) From a) and from the fact that $m_0 - \gamma$ is normally distributed with mean 0 and variance

$1/\tau_\theta + 1/\tau_m$, we have:

$$w(0) = \gamma - \frac{t^{m_0-\gamma}}{t^m} - \frac{t^{m_0+\tau_\theta}}{t^m} (m_0-\gamma) = \gamma - \frac{t^{m_0-\gamma}}{t^m} = E(w(0)) = \gamma; \quad V(w(0)) = \frac{t^{m_0-\gamma}}{t^m} = \frac{t^{m_0+\tau_\theta}}{t^m} \quad (15)$$

d) Taking the unconditional expectation of $w(t)$ gives:

$$E(w(t)) = \gamma = \frac{E(x)^{t+\tau} E(m_0)^{t+\tau_\theta}}{E(x)^{t+\tau} E(m_0)^{t+\tau_\theta}} = \frac{E(x)^{t+\tau} E(m_0)^{t+\tau_\theta}}{E(x)^{t+\tau} E(m_0)^{t+\tau_\theta}} = \frac{E(x)^{t+\tau} E(m_0)^{t+\tau_\theta}}{E(x)^{t+\tau} E(m_0)^{t+\tau_\theta}} \quad (16)$$

To derive the predictions pertaining to the evolution of the variance, two extreme cases are

considered. The first case is when jobs are pure search goods and the second one when they are pure experience goods.

A. Jobs as pure search goods ($\tau_m \rightarrow \infty$).

In this environment, the decision rule for each worker at time t is to accept an outside offer w_t whenever $w_t > w_{t-1}$, where w_{t-1} is assumed to be the wage payed in the current match. Otherwise, the worker stays with the same employer. Note that we are assuming no accumulation of human capital. Therefore, the wage paid on a job is a constant throughout the employment relationship. Under these assumptions, the threshold above which workers quit their match rises at each period t . Hence, the wage distribution of job changers gradually moves over time toward the right-hand tail.

PROPOSITION 2

For each successive t , the moments of the truncated distribution are then given by:

$$\begin{aligned}
 E(w_t^j | w_t > w_{t-1}) &= \gamma + \left(\frac{1}{1 - \Phi(\tau_\theta)} \right) \frac{\phi((w_{t-1} - \gamma)\tau_\theta)}{\phi((w_{t-1} - \gamma)\tau_\theta)} = \gamma + M_1 \\
 V(w_t^j | w_t > w_{t-1}) &= \frac{1}{1 - \Phi(\tau_\theta)} - \frac{1}{M_1} = \frac{1}{M_1} - \frac{1}{M_1} = 0
 \end{aligned}
 \tag{17}$$

where $\Phi(\cdot)$ is the standard normal distribution function and $\phi(\cdot)$ is the corresponding density.

PROOF

From c) above, the initial ($t=0$) distribution of wages in the population coincides with the prior distribution. Thus w_0 is normal with mean γ and variance $1/\tau_\theta$. At each period t , the wage w_t has the same distribution as the initial wage only if w_t is greater w_{t-1} , i.e. the distribution is truncated. Defining $w_0^* = \gamma + (1/\tau_\theta)\epsilon$, where ϵ is a standard normal variate, the following set of results are obtained:

d)

$$(20) \quad E(w^t | w^t > w^{t-1}) = \frac{P(w^0 > w^{t-1})}{E(w^0 | w^0 > w^{t-1})} = \frac{\frac{1}{1-\Phi} \phi(w^{t-1} - \lambda) \tau_{\theta, 5}}{\frac{1}{1-\Phi} \phi(w^{t-1} - \lambda) \tau_{\theta, 5}} + \lambda =$$

e)

$$(19) \quad E(w^t) = E(w^0 | w^0 > w^{t-1}) = E(\lambda + \frac{\tau_{\theta, 5}}{1} \epsilon | \epsilon > (w^{t-1} - \lambda) \tau_{\theta, 5}) + \frac{\tau_{\theta, 5}}{1} \phi(w^{t-1} - \lambda) \tau_{\theta, 5}$$

b)

$$(18) \quad P(w^t > w^{t-1}) = P(w^0 > w^{t-1}) = P(\lambda + \frac{\tau_{\theta, 5}}{1} \epsilon > w^{t-1}) = P(\epsilon > (w^{t-1} - \lambda) \tau_{\theta, 5}) = 1 - \Phi(w^{t-1} - \lambda) \tau_{\theta, 5}$$

a)

$$E(w'_2) = E(w_2 | w_0 > w_{r-1}) = E(\gamma + \frac{1}{2} \epsilon | \epsilon > (w_{r-1} - \gamma) \tau_{\theta.5}) + \gamma^2 (1 - \Phi((w_{r-1} - \gamma) \tau_{\theta.5})) + \frac{1}{2} (1 - \Phi((w_{r-1} - \gamma) \tau_{\theta.5})) \quad (21)$$

$$V(w'_2 | w'_2 > w_{r-1}) = E(w_2^2 | w_2 > w_{r-1}) - (E(w_2 | w_2 > w_{r-1}))^2 = \frac{E(w_0^2 | w_0 > w_{r-1})}{E(w_0 > w_{r-1})} - \frac{P(w_0 > w_{r-1})}{(E(w_0 | w_0 > w_{r-1}))^2} = \frac{1}{2} \frac{\tau_{\theta}}{\phi((\gamma - w_{r-1}) \tau_{\theta.5})} \frac{1}{\phi((\gamma - w_{r-1}) \tau_{\theta.5})} - \frac{1}{2} \frac{\tau_{\theta}}{\phi((\gamma - w_{r-1}) \tau_{\theta.5})} = V(w_0^2 | w_0 > w_{r-1}) = \frac{1}{I} - M_1(M_1 - (w_{r-1} - \gamma) \tau_{\theta}) \quad (22)$$

Taking the derivative of $V(\cdot)$ with respect to w_{r-1} gives:

$$\frac{dV(w'_2 | w'_2 > w_{r-1})}{dw_{r-1}} = -2M_1 \frac{dM_1}{dw_{r-1}} + M_1 + (w_{r-1} - \gamma) \frac{dM_1}{dw_{r-1}} \quad (23)$$

This expression can be negative or positive. First note that dM_1/dw_{r-1} is strictly positive.² Moreover, there exists a critical value w^* such that for all w_{r-1} below this threshold, $dM_1/dw_{r-1} > 1/2$ and for all values of w_{r-1} at least as large as w^* , then $dM_1/dw_{r-1} < 1/2$.³ By substituting $dV/dw_{r-1} > 0$ for all $w_{r-1} < w^*$. Thus, if $w_{r-1} \leq \gamma$, we know for

² $dM_1/dw_{r-1} = \phi(w_{r-1} - \gamma) / (1 - \Phi(w_{r-1} - \gamma)) - (\phi(w_{r-1} - \gamma)) / (1 - \Phi(w_{r-1} - \gamma))$, which is strictly positive.

³ dM_1/dw_{r-1} is a continuous function of w_{r-1} . It can be shown that the limit of dM_1/dw_{r-1} is zero for $w_{r-1} \rightarrow -\infty$ while it tends to infinity with $w_{r-1} \rightarrow \infty$. Thus, there exists a threshold w^* with $dM_1/dw_{r-1} = 1/2$.

Here the decision to quit an employer depends on the information that can be extracted from the sequence of outputs. Of course, workers know that they can obtain a starting wage equal to γ for any draw made from the prior distribution. But, contrary to the search model, this is the only offer they will receive over time from the outside as nothing is revealed about the quality of the match short of working in it. Therefore, in this environment, on the expected present value from leaving, say V , is a constant and the decision to quit hinges on the belief workers form about the quality of the match. From (d) above, we know the unconditional distribution of the wage received by a sample of workers. This would be the eventual distribution of wages across these workers if all of them were to stay with the same employer for their entire career. Over time, however, some workers will quit as the information they get from the sequence of output indicates with more and more precision that it is preferable to accept a wage equal to γ in a new job which offers an expected present value of V . It can be shown that the optimal decision rule for the individual is to stay with the same employer if w_t given by equation (1) is above a critical value w_t^* which rises over time and to quit otherwise.⁴ Consequently, the moments of the truncated distribution are

B. Jobs as pure experience goods ($\tau_m = 0$).

The same result holds for the evolution within jobs. In a search context, individuals stay with the same employer only if the wage they receive at the start of the employment relationship (and which stays the same throughout the duration of the match) is at least as great as any offer they receive over time. Thus, as $t \rightarrow \infty$, the only people still with the same employer must have drawn from the upper portion of the distribution, where the variance is larger.

sure that equation (23) is negative. Conversely, for all $w_{t-1} > w^*$, $dV(\cdot)/dw_{t-1} > (w_{t-1} - \gamma)$. With $w_{t-1} \geq \gamma$, then $dV(\cdot)/dw_{t-1} > 0$. Consequently, we should first observe a decline in the variance from job to job followed by an increase as we move towards the right-hand tail.

given by:

$$\begin{aligned}
 E(w^t | w^t > w^*) &= \lambda + \sigma^t \frac{1 - \Phi((w^* - \lambda) / \sigma^t)}{\phi((w^* - \lambda) / \sigma^t)} = \lambda + M_1 \\
 V(w^t | w^t > w^*) &= \sigma^t \left(M_1 - M_1^2 (M_1 - (w^* - \lambda)) \right)
 \end{aligned}
 \tag{24}$$

where σ^t is the unconditional variance given in equation (5).⁵ As in the case where jobs are pure search goods, the derivative of $V(\cdot)$ could be calculated to get the evolution of the variance within jobs. However, the presence of σ^t greatly complicates the task of determining the sign of that derivative. The reason is that σ^t is not constant and thus additional terms appear in the analog of equation (23). The only solid conjectures that can be advanced are the following: (i) from $t=0$ to $t=1$, there is an increase in the variance (the variance is zero at $t=0$). After that, it may increase or decrease depending upon the speed at which the learning process evolves. If all learning is done rapidly, then it is as though the unconditional distribution given by (d) above stabilizes (the variance stops increasing). Thus, a selection process identical to the one in case A above would impart a decreasing-increasing pattern to the variance within jobs.

The evolution of the variance across jobs is simple: with a totally uninformative screen, all jobs are ex ante identical (draws are independent) and thus the variance should be stable from job to job.

To recapitulate, if jobs are pure search or inspection goods, the variance within jobs and across jobs will first decline and eventually will increase as we reach the right-hand tail

⁴ See Sargent (1985) for a proof in a finite lifetime context.

⁵ Note that the conditional mean indicates that if we were to estimate an earnings equation with ordinary least-squares, we would find a positive return to tenure simply because those that happen to have better matches are still with the same employer, even though their wage does not necessarily rise with tenure.

with $t \rightarrow \infty$. If jobs are pure experience goods, the variance across jobs should be similar while the evolution within jobs is characterized by an increase from $t=0$ to $t=1$. For $t > 1$, the evolution is undetermined. If the learning process is completed fairly rapidly, then we would expect the selection process to impart a decreasing/increasing pattern like in the case where jobs are pure search goods.

III. THE THEORY OF HUMAN CAPITAL AND THE COVARIANCE STRUCTURE.

The question of examining the empirical implications of the theory of human capital is, of course, not new. The human capital earnings function has been used extensively since its development by Mincer (1974). However, the implications in terms of the covariance structure have rarely been looked upon. Hause (1980) is a notable exception. This section will first summarize these implications before turning to the empirical section.

Mincer's overtaking concept is at the heart of the analysis in terms of the covariance structure of earnings. Assuming that two individuals have identical attributes (total labor market experience, schooling, etc.), then, according to human capital theory, their earnings profiles should differ only if their rates of post-schooling investment are not the same.⁶ The worker who invests more should have lower initial earnings than the other worker. However, as she accumulates human capital, her earnings should increase at a faster rate and eventually surpass the other worker's earnings. Generalizing to a sample of relatively homogeneous workers followed over time, we should then observe a declining profile of the variance of the log of earnings in the years leading to the overtaking point, followed by an increase in variance afterwards. Therefore, we should have a negative correlation between the slope of the earnings profile and its intercept.

⁶ Of course, we also assume away the presence of uncertainty or informational asymmetries.

Note however that the observed time profile of variances might not necessarily be U-shaped were individuals to differ in some unobserved dimension which is correlated with the rate of investment, as Mincer points out in his analysis of residuals.⁷ More precisely, assuming that individuals who are intrinsically more able also invest more in on-the-job training, then we might have a positive correlation between the slope and the intercept of the earnings profile if the correlation between unobserved "ability" and the rate of investment is sufficiently strong.

These empirical implications of the theory of human capital have been tested by Hause (1980) with a six-year sample of young Swedish white collars who were born the same year and had similar levels of schooling. He allows each individual to have his own level and slope (i.e. experience slope) parameters. In other words, his residual structure of earnings is of the form:⁸

$$w_{it} = \alpha_{1i} + \alpha_{2i} \text{Experience}_{it} + u_{it} \quad (25)$$

The slope and intercept parameters are assumed to be independent of the residual term u_{it} and also independently distributed across individuals. Consequently, taking the expectation over individuals of the cross product of w_{it} and w_{is} gives us:

$$E(w_{it} w_{is}) = \phi_{11} + \phi_{22} \text{Exp}_{it} \text{Exp}_{is} + \phi_{12} (\text{Exp}_{it} + \text{Exp}_{is}) + \sigma^2 \quad (26)$$

$$\phi_{11} = \text{var}(\alpha_{1i}); \phi_{22} = \text{var}(\alpha_{2i}); \phi_{12} = \text{cov}(\alpha_{1i}, \alpha_{2i})$$

Relating equation(26) to the discussion above, the overtaking year should be associated with the year at which we observe the minimum variance. Thus, minimizing equation(26) with respect to the level of experience t_0 (i.e. time of overtaking) gives us $t_0 = \phi_{12} / \phi_{22}$. Because of unobserved ability, t_0 will represent a lower bound of the true overtaking year. With six years of data, the empirical covariance matrix contains 21 distinct elements

⁷ For the full derivation of these assertions, see Mincer (1974), chapter 6.

⁸ Actually, Hause directly fitted the wage observations instead of using the residuals from a fully specified earnings function.

¹⁰ Six distinct residual (uit's) variance parameters each following a distinct AR(1) process, which produces 11 parameters (in addition to the other three) for the 21 elements to be fitted.

⁹ Hause estimated the parameters of his model with quasi-maximum likelihood methods.

The predictions derived above are examined using unbalanced data from the National Longitudinal Survey of Youth (NLSY) for the period spanning the years 1979-1991. This data set contains the full employment history of young Americans from the moment they make their (full time) transition from the school to the labor market. The workers who were

IV.1 The data

IV. EMPIRICAL IMPLEMENTATION

To conclude on Hause's paper, it should be mentioned that the time profile of the variances of the log of earnings which he tries to fit is not U-shaped. Instead, the variances shows a monotonic decrease from year one to year six. But this is one pattern which learning theories like the theory of matching predict, especially in the formulation where all learning about the job characteristics is done rapidly. Therefore, to discriminate between the two theories, it is not sufficient to look at the residual earnings structure by focusing only on the variance of the experience profiles of workers. What is done in the next section is to extend Hause's approach to the tenure profiles.

which Hause then fits with the structure implied by human capital theory. He does find a negative correlation between the slope and intercept parameters but whether this correlation is significant or not depends on the residual error structure that he imposes. Under the specification that produces a statistically significant result, he estimates t_0 to be in the neighborhood of five years.⁹ Note that unless he imposes a very loose structure on the u_t ¹⁰ term, the model is rejected by the data and there is no significant tradeoff between the slope and the intercept.

considered as having entered the labor market on a full-time basis were (i) those whose primary activity was either working full-time, on a temporary lay-off or looking actively for a job, (ii) those who did not return to school on a full-time basis within six years and (iii) those who had worked at least half the year since the last interview and who were working at least 20 hours per week. Individuals excluded from the sample are those younger than 18, those that had been in the military at any time, the self-employed, the ones whose jobs were part of a government program and the ones working without pay, those who were in the farming business and also all public sector employees. These restrictions leave us with a sample consisting of 29,020 observations (5,649 workers). Summary statistics are reported in table 1.

IV.2 Log-Earnings Equation

Let the wage of person i at time t be determined according to the following equation:

$$w_{ijt} = E(w_{ijt} | \text{observables}) + \epsilon_{ijt} \quad (27)$$

where w_{ijt} is the log of the real hourly wage of worker i in job j . The first step is then to estimate equation (27) with ordinary least-squares to obtain the residuals. After that, the models that I estimate are models for the expectation of the cross-products of the residuals, $E(\epsilon_{ijt} \epsilon_{i'ks})$.¹¹ I try to determine to what extent the predictions from the two theories are supported by the data. It is worth reiterating that my primary interest is to examine the implications of the selection process by which workers are sorted into jobs. Therefore, in estimating equation (27), I make no attempt at trying to correct for selection biases.

¹¹ The matrix of observables includes total experience, total experience squared, tenure, tenure squared, experience in the industry, experience in the industry squared, industry and occupation dummies, year dummies, race, sex, marital status, union membership, SMSA, urban/rural, health, unemployment rate in corresponding region, an intercept, four region and three education dummies. The vector of estimated residuals is then expressed in deviations from annual means.

IV.3 Econometric Models of Matching

As a first step, a test of the matching model in its simplest form is presented. More precisely, I test whether the covariance structure of the error term of a standard human capital earnings equation satisfies the restrictions imposed by the job matching theory. The simplest way to account for the process of matching is to include an error component θ_{jt} as part of the total error term ϵ_{ijt} . It represents the unobserved (to the econometrician) quality of the match which affects the wage of individual i . This component is assumed to be fixed within matches although there is a whole distribution of match productivities. Assuming that the error term also contains an individual-specific component to reflect time-invariant unobserved individual characteristics, then it can be written as:

$$\epsilon_{ijt} = \alpha_i + \theta_{jt} + \eta_{ijt} \quad (28)$$

where η_{ijt} is a white noise error term. Note that the three terms are assumed to be independently distributed. Taking the expectation of the cross-products gives:

MODEL 1

$$E(\epsilon_{2t}^2) = \sigma_{\alpha}^2 + \sigma_{\theta}^2 + \sigma_{\eta}^2$$

$$E(\epsilon_{1t}^2 | I_t) = \sigma_{\alpha}^2 + I_t' \sigma_{\theta}^2$$

(29)

where I_t is an indicator variable equal to 1 if worker i holds the same job at periods t and s . To test the predictions of the theory of matching pertaining to the evolution of the variance of the job-match component across jobs, model 1 is expanded by allowing a different job-match component for each job held by the worker up to job five and over. Let job_k be an indicator variable equal to 1 if worker i is in job k ($k=1, \dots, 5$) and equal to 0 otherwise, then the model to be estimated is the following:

MODEL 2

$$\begin{aligned}
 E(\epsilon_{2}^{it} | job_{it}) &= \sigma_{2}^{\alpha} + \sum_{s+}^{k-1} (job_{it}^{s} \sigma_{2}^{\theta_s}) + \sigma_{2}^{\eta} \\
 E(\epsilon_{1s}^{it} | I^{s} job_{it}^{k}) &= \sigma_{2}^{\alpha} + I^{s} \sum_{s+}^{k-1} (job_{it}^{k} \sigma_{2}^{\theta_k})
 \end{aligned}
 \tag{30}$$

In the same spirit, to study the evolution of the variance of the match component within jobs, model 1 is generalized with the use of dummy variables for different levels of tenure. More precisely, let $dummies_s$ be equal to 1 if the tenure level of worker i is included in the interval corresponding to s and equal to 0 otherwise, where $s=(1$ if $tenure < 1$ month, 2 if 1 month $\leq tenure < 2$ years, 3 if 2 years $\leq tenure < 4$ months, 4 if 4 months $\leq tenure < 1$ year, 5 if 1 year $\leq tenure < 2$ years, 6 if 2 years $\leq tenure < 3$ years, ..., 15 if $tenure \leq 12$ years). By including dummies for low levels of tenure, the hope is to be able to capture any increase (if there is any) in the variance of the job-match component, which is the prediction of the matching model where at least some learning about the job occurs in the course of the employment relationship. If it is impossible to isolate an increasing pattern with this specification, then it may be that most of the on-the-job learning process is completed very quickly. Or, it may also be the case that some other factor such as the accumulation of human capital is playing a role. The model for the within-job evolution is then:

MODEL 3

$$\begin{aligned}
 E(\epsilon_{2}^{it} | dummies_{it}) &= \sigma_{2}^{\alpha} + \sum_{15}^{s+1} (dummies_{it} \sigma_{2}^{\theta_s}) + \sigma_{2}^{\eta} \\
 E(\epsilon_{1s}^{it} | I^{s} dummies_{it}) &= \sigma_{2}^{\alpha} + I^{s} \sum_{15}^{s+1} (dummies_{it} \sigma_{2}^{\theta_s})
 \end{aligned}
 \tag{31}$$

IV.4 Econometric Model of Human Capital

Generalizing Hausess's model to the tenure profile, let the error term be specified as:

MODEL 4

Now remember that I am using the residuals produced by estimating equation (27). That equation already contains quadratic functions of both tenure and experience. Therefore, how can the presence of these two explanatory variables in the error structure be justified? It turns out that assuming a random coefficient model is perfectly in tune with the spirit of Mincer's derivation of the human capital function.¹² In deriving the earnings function, Mincer first includes a subscript *i* representing each individual with the idea that the rate of investment as well as the return on that investment are individual-specific parameters.¹³ Since these parameters are not directly observable, their corresponding *i* subscript is then eliminated and we are left with estimating average rates of returns. By focusing on the second moments, it is possible to take into account the randomness of the tenure and experience slope parameters. Also, given that training occurs within jobs, it seems more natural to model it as such. In view of the foregoing, the model to be estimated has the following basic structure:

$$\epsilon_{ijt} = \alpha_1 + b_{1i} Experience_{ijt} + \theta_{ij} + b_{12} Tenure_{ijt} + u_{ijt} \quad (32)$$

¹² See Mincer (1974), chapter 5, section 5.2.

¹³ "If information were available on all variables and parameters for each individual *i*, the [earnings] equation would represent a complete accounting (...) of the human capital characteristics entering into the formation of earnings." Mincer (1974), page 90.

$$w_{ij} = \beta_1 X_{ij} + \beta_2 T_{ij} + \epsilon_{ij}$$

$$\epsilon_{ij} = \alpha_1 + b_{1i} X_{ij} + \theta_{ij} + b_{12} T_{ij} + \eta_{ij}$$

$$E(\epsilon_{ij} | X_{ij}, T_{ij}) = E(\alpha_1) = E(\theta_{ij}) = E(b_{1i}) = E(b_{12}) = E(\eta_{ij}) = 0;$$

$$E(\epsilon_{ij}^2 | X_{ij}, T_{ij}) = \sigma_2^2 + \sigma_{b_1}^2 X_{ij}^2 + \sigma_{\theta}^2 + \sigma_{b_2}^2 T_{ij}^2 + \sigma_{\eta}^2$$

$$E(\epsilon_{ij} \epsilon_{i'j'} | X_{ij}, X_{i'j'}, T_{ij}, T_{i'j'}) = \sigma_2^2 + \sigma_{b_1}^2 X_{ij} X_{i'j'} + \sigma_{\theta}^2 (T_{ij} T_{i'j'}) + \sigma_{b_2}^2 (T_{ij} T_{i'j'}) + \sigma_{\eta}^2$$

where T and X are respectively the tenure and experience levels and I_1 is the same indicator variable as in the previous section. Note the following implicit assumptions: a) α , θ and η are independent from one another, as previously, b) b_1 and b_2 are also independently distributed, and c) experience and tenure are not correlated with the random parameters. Therefore, I am assuming away any selection effects that would occur were tenure to be positively correlated with the random coefficient b_2 . This orthogonality assumption is likely to be a strong one as we would expect workers with a high b_2 to stay longer on their job. Also, I am assuming that the observable individual characteristics used in estimating equation (33) are independent of the unobservables.

The same type of tradeoff than the one analyzed by Hause is applicable to the tenure profile. Namely, according to human capital theory, we should observe a negative correlation between job-specific slope and intercept parameters. In other words, workers who have relatively more on-the-job training should pay for it through lower initial earnings.¹⁴ The same sort of counter-effect which made Hause's estimate of Mincer's overtake point a lower bound is present here. More precisely, Jovanovic (1979b) has shown that the quality of the match and the level of investment in firm-specific capital are complementary. Consequently,

¹⁴ Note that we are not making any statement about the degree of firm-specificity of the human capital acquired in the course of OJT.

according to this line of thought, other things being equal, we should observe a positive correlation between the job-specific slope and intercept parameters. Thus, the U-shape pattern of the variance of the log of earnings within jobs is not a prediction that is shared by the two theories.

Another prediction that is implied by human capital theory is that training should occur early in one's career (e.g. Ben-Porath (1967)). Therefore, the U-shape pattern should be more evident in the first few jobs than later on. On the other hand, the quality of the match should improve as one moves from one job to another. Hence, it is likely that a worker entering the labor market and a firm who find themselves in a bad match would be reluctant to invest in firm-specific skills. It is even conceivable that the worker would be less willing to invest in general human capital. It takes time to complete training and the fact that she is more likely to leave before completing the program than in a job where she would be better matched could provide a rationale for delaying investments in general skills. If these job-matching considerations are at play, then the correlation between the job-specific intercept and slope parameters should tend to be weaker in the first job than in the second. To verify these predictions, equation (33) is expanded by fitting five different quadratic functions to allow examination of the evolution of the parameters from job to job.

IV.5 Estimation Methodology

To estimate and test the restrictions imposed by the models, I make use of the methodology proposed by Gallant and Jorgenson (1979) in the context of a system of nonlinear implicit equations. The basic idea is to compare the weighted sum of squares of an unrestricted model with that of the (restrictive) model I wish to estimate. This methodology is closely related to the minimum distance method proposed by Chamberlain (1982, 1984) and

adapted by Abowd and Card (1989) to study the covariance structure of earnings (see the appendix for further details). The conditional moments equations are stacked up into a system of equations which, as an example, would be of the following form for model 1:

$$E(m_i | I_i) = f(\beta, I_i) \quad (34)$$

$$m_i = [\epsilon_{i1}, \epsilon_{i2}, \dots, \epsilon_{i12}, \epsilon_{i13}, \dots, \epsilon_{i13}, \dots, \epsilon_{i13}]$$

where T_1 is last period in which worker i is in the sample, β is the vector of parameters to estimate and $f(\cdot)$ is the mapping representing the model. Were all 5 649 individuals present in the sample from 1979 to 1991, each model above would represent a system of 91 equations (13 second moments and 78 cross-products), for a grand total of 514 059 observations. The sample being unbalanced, I end up with 116 715 observations. The objective is then to minimize the following function:

$$S(\beta) = \sum_N^{t=1} (m_t - f(\beta, I_t))' V^{-1} (m_t - f(\beta, I_t)) \quad (35)$$

Where V^{-1} is computed with the cross-products of the residuals from the following unrestricted model:

$$E(\epsilon_{it}^2) = cov I_{it}^2 \quad (36)$$

$$E(\epsilon_{it}^s \epsilon_{it}^{ts} | I_{it}^s) = cov I_{it}^s + I_{it}^s cov 2_{it}^s$$

This unrestricted model contains 91 different cov1 parameters and 78 distinct cov2 parameters. Note that if only the cov1's were estimated, each of these parameters would be equal to the corresponding sample moment. But for the matching models that I want to estimate to be nested in a more general model which depends on each individual-specific I_{it} , the cov2's have to be estimated as well. Let $S(\beta)_1$ be the value of the objective function for the unrestricted model and $S(\beta)_2$ be defined likewise for the restricted version of equation

(34). For more general (i.e. non-linear) models, Gallant and Jorgenson (1979) show that $T_0 = S(\beta)_2 - S(\beta)_1$ is asymptotically distributed as a chi-square with $r-s$ degrees of freedom where $r-s$ is equal to the difference in the number of parameters in the two models.

V. RESULTS

V.1 Matching Models

Results for the entire sample are reported in tables 2, 3 and 4. To get an idea of the impact of including a job-match component in the error structure, columns 1 and 2 of table 2 show the results from estimating model 1 without the match component ("model 0"). The first conclusion to be drawn is the absence of statistical support for the model which provides control for individual heterogeneity only and in which η_i is i.i.d. (column 1). The distance statistic of 262 is a surprising value coming from a χ^2 (167). When the job-match component is added (column 2), the fit is substantially improved. In fact, as in Finn (1986), the structure cannot be rejected. It might seem surprising that the data cannot reject such a simple structure, especially given some recent attempts (e.g. Abowd and Card (1989), Farber and Gibbons (1994)) at trying to fit parsimonious structures to the error term. However, it should be remembered that the estimated log-wage equation from which the residuals are extracted contains all the explanatory variables which typically appear in such an equation. Consequently, all observable time-varying individual or job-match characteristics are controlled for, which makes it considerably easier to fit a restrictive structure on the error term. On the other hand, Abowd and Card to take an example, control only for potential experience in the labor market.¹⁵

To examine the evolution of the variance of the job-match component from job to job,

¹⁵ If I try to fit the same structure controlling only for the experience level, then the model is rejected by the data.

Turning now to the within-job evolution of the variance of the job-match component (model 3), we want to see if the contribution of that component is monotonically declining, which would suggest that the screening process reveals a lot of information, or if it is first increasing and then decreasing, a pattern which is indicative of an imperfect screen and of some learning taking place early while gaining experience in the job. Note that if all learning takes place quite early in the employment relationship (say less than one month after the start of the job), then it will be impossible to pick up any increase in the contribution of the variance of the job-match component. We will only be able to see the impact of the selection

I need to account in some way for the presence of aggregate effects to avoid as much as possible these effects being picked up by the coefficients. The simplest way to provide such a control is by assuming that the residual error term η is independently but not identically distributed from period to period. Therefore, I estimate 13 period specific variances of η . The results for the whole sample are reported in table 3, column 1. We can see that no clear pattern emerges although the overall tendency would be positive were it not for the component attached to job 4. Remember that in the version of the theory of matching where jobs are pure experience goods, successive and independent sampling from the job-match distribution implied that the variance of that distribution should be stable from job to job, whereas if the jobs were pure inspection goods, then the variance should be declining as we follow a cohort of workers over time and across different jobs. The overall pattern here is non-trivial. One of the reasons stems no doubt from using the entire sample which consists of workers who are heterogeneous in terms of schooling and also in terms of time of entry into the labor market. Also, because workers with more schooling enter the sample later, they are less likely to be represented in jobs 4 and 5. Consequently, to alleviate to a certain extent these sample composition problems, I present the results by levels of schooling in columns 2, 3 and 4. Here again, we can see that no clear pattern emerges. The fact that the variance increases is not consistent with the matching theory, whether jobs are considered as search goods or as pure experience goods.

¹⁶ Results obtained by Farber (1994) with the NLSY on the shape of the hazard function seem to indicate that job attributes are known quite early after the beginning of the employment relationship.

The model implied by equation (33) is estimated in the same fashion as the "pure" matching model of section II. More precisely, all moment conditions are stacked up into a system and estimated with generalized least-squares where the weighting matrix is computed from the residuals of an unrestricted model. The results are shown in tables 5 and 6 for the whole sample. For comparison purposes with Hause (1980), I have also estimated the structure with only a linear trend in the experience variable (see equation 25), both with and without the term representing the variance of the job-match component (columns 1 and 2). The predicted tradeoff between the experience slope and the individual intercept comes out quite strongly when σ^2 is absent from the equation. However, once we include it, results are much less convincing in favor of human capital. This illustrates the point made earlier about the need to devise an estimating framework which allows us to isolate the key prediction of the theory of human capital as it pertains to the overtaking concept. Turning now to the estimation of model 4, note the negative covariance between the tenure slope and the job-specific intercept. Depending on the specification, the average overtaking point within jobs

V.2 Human Capital Model.

process by which this contribution decreases with tenure.¹⁶ The results are reported in table 4, column 1 for the whole sample and in columns 2, 3 and 4 for the schooling groups. The estimated coefficients, even for the whole sample, are indicative of a declining pattern. This is particularly true for the more educated workers in the sample. Thus, it seems that the matching process which best corresponds to these results is the one where jobs are better characterized as inspection goods rather than as experience goods. Note however that for workers with less than a high school diploma, the variance at 4 months of tenure is slightly larger than at 1 month, which is quite different than for the other workers. In fact, it does not decrease in the first three years of the employment relationship.

is between three years and a little over four years. The estimated variance of the tenure slope imply a standard error of close to 3 percent in the return to tenure. This finding provides strong suggestion that there is no such thing as a return to tenure, but that there is a range of returns.¹⁷ Table 6 shows the results if the model is fitted with five different quadratic functions, one for each job. We can see that the lower bound estimate of the overtaking point reaches a peak in job two with about 7.5 years. The variance of the tenure profile also increases from job 1 to job 2. Note that with the number of degrees of freedom afforded by the size of the unrestricted model, all models tested here cannot be rejected.

The same caveats as those mentioned in section II apply here regarding the lack of homogeneity of the overall sample and the sample composition problems as we move into jobs 4 and 5. Nevertheless, the robustness of the fitted U-shape pattern to different specifications suggests a substantial role for on-the-job investment in human capital.

Tables 7 and 8 show the results by education category. An interesting result is the absence of a U-shape pattern for workers with less than high school education (column 1). Previous studies (e.g. Lynch (1992)) linking the probability of receiving formal training programs to various socio-economic background variables found a positive relationship between having a high school diploma and the occurrence of training. The results here go in the same direction (table 8, column 1). Also, the process described above which tends to make the tradeoff stronger on the second job than in the first job seems to be absent here. The results suggest that for workers with less than high school education, search activity does not lead them to higher quality jobs which entail some on-the-job-training. They seem to move from jobs which do not require a lot of training to other similar jobs. Another result from table 8, column 1 that supports this conclusion is the absence of any statistically significant variance in the tenure slopes for these less educated individuals, which is not the case for the other subsamples.

¹⁷ See Abowd, Kramarz and Margolis (1994) for similar conclusions with different data.

V.3 Matching with Control for Human Capital Accumulation

148

Coming back to the results of section II on matching, it was hypothesized that the relative difficulty of getting results that were consistent with the theory might be due to human capital considerations. A direct way of verifying that hypothesis is to enrich models 2 and 3 by including the linear trend in experience in the specifications. Assuming that this linear trend captures all factor reflecting the accumulation of human capital over the course of one's career, it should then be easier to isolate the predictions of the theory of matching. Results are reported in table 9 and 10.

Although the behavior of the variance across jobs does not change markedly, the declining pattern of the within job evolution of the job-match component variance is still further highlighted when the linear experience trend is added to the specification (see figures 1, 2, 3 and 4). For workers with less than a high school diploma, note that the variance in the first job is significantly larger than in jobs 5+, although the behavior of the variance from job 2 to job 4 is erratic according to the theory.

VI. Conclusion.

In extending Hause's original approach by making use of the concept of jobs, I was able to obtain results favoring the theory of human capital. In addition, the estimated U-shape pattern of variances within jobs is a prediction that is exclusive to human capital theory. In other words, there is no identification problem here as far as competing theories are concerned.

Regarding the predictions of the theory of matching, it is likely that better control for aggregate effects would have helped in finding stronger evidence of the patterns predicted by the theory, especially regarding the evolution across jobs of the variance of the job-match component. A more complex modeling allowing for autocorrelation of the residual term

(which is considered here as being independently distributed across periods) could have captured some noise which, under the models adopted, may find its way into the coefficients. Also, it may be that testing these predicted patterns requires a balanced panel where a single cohort making its entry into the labor market is followed for many years. Unfortunately, a large random sample of young workers making their transition to the labor market and where no attrition is observed over a fifteen-year (say) period is something labor economists have grown used to dream about. Nevertheless, the results for the within-job evolution are quite consistent with the theory, especially for more educated workers. We should realize, though, that these results are also consistent with workers undergoing training programs early in the employment relationship. For workers with less than a high school diploma, the results are indicative of perhaps some learning taking place in the first few years. It might also be the case the predicted declining pattern is not observed at the beginning of the employment relationship simply because these less educated workers are not being trained by their employers.

The most sensible conclusion from the study is that, while matching plays an important role in explaining the covariance structure of earnings, it is apparent that randomness in the tenure coefficient stemming from differential amounts of on-the-job training is also a major contributor.

Table 1. Mean Sample Characteristics (Weighted)-NLSY

Real Hourly Wage (\$1979)	5.75
Hours Worked	41.7
Tenure	2.44
Experience	5.82
Years in School	12.44
Percentage Nonwhite	12.6
Percentage Married	44.7
Percentage Female	45.4
Age	25.1
Number of Observations	29,020
Number of Individuals	5,649
Number of Jobs	13,590

Table 2. Estimation of Covariance Structure Implied
by the Theory of Matching
(Standard Errors in Parentheses)

Parameter	Model 0	Model 1
Variance of unmeasured fixed worker ability σ^2_α	0.0446 (0.0010)	0.0315 (0.0017)
Variance of unmeasured fixed job-match quality σ^2_θ	-	0.0284 (0.0013)
Variance of residual white noise error term σ^2_n	0.0658 (0.0019)	0.0534 (0.0020)
χ^2 statistic (degrees of freedom)	262 (167)	138 (166)
Number of workers	5 649	5 649
Number of observations	116 715	116 715

Table 3. Evolution of the Variance of the Job-Match Component from Job 1 to Job 5 and Beyond (Standard Errors in Parentheses)

Parameter	Total Sample	Less than H.S. Education	High School Education	Above H.S. Education
Variance of unmeasured fixed worker ability σ^2_{α}	0.0317 (0.0012)	0.0234 (0.0028)	0.0236 (0.0014)	0.0396 (0.0024)
Variance of unmeasured fixed job-match quality σ^2_{θ} (job 1)	0.0255 (0.0022)	0.0245 (0.0043)	0.0221 (0.0021)	0.0296 (0.0035)
σ^2_{θ} (job 2)	0.0299 (0.0021)	0.0226 (0.0049)	0.0223 (0.0025)	0.0393 (0.0041)
σ^2_{θ} (job 3)	0.0303 (0.0031)	0.0270 (0.0068)	0.0359 (0.0038)	0.0198 (0.0059)
σ^2_{θ} (job 4)	0.0256 (0.0049)	0.0444 (0.0097)	0.0282 (0.0063)	0.0302 (0.0103)
σ^2_{θ} (job 5+)	0.0408 (0.0070)	0.0186 (0.014)	0.0523 (0.0097)	0.0429 (0.0129)
Residual term i.d. (13 terms)	Yes	Yes	Yes	Yes
χ^2 statistic (degrees of freedom)	82 (150)	41 (150)	69 (150)	66 (150)
Number of workers	5 649	1 515	3 010	1 702
Number of observations	116 715	20 910	58 539	32 204

Table 4. Evolution of the Variance of the Job-Match Component within Jobs

Parameter	Total Sample	Less than H.S. Education	High School Education	Above H.S. Education
Variance of unmeasured fixed worker ability	0.0316	0.0230	0.0232	0.0404
Variance of unmeasured fixed job-match quality	0.0440	0.0221	0.0489	0.0807
1 mo. <= Tenure < 4 mo.	0.0354	0.0249	0.0234	0.0429
1 yr. <= Tenure < 2 yrs.	0.0292	0.0280	0.0267	0.0326
2 yrs. <= Tenure < 3 yrs.	0.0312	0.0310	0.0250	0.0377
3 yrs. <= Tenure < 4 yrs.	0.0258	0.0107*	0.0255	0.0266
4 yrs. <= Tenure < 5 yrs.	0.0258	0.0307	0.0211	0.0340
5 yrs. <= Tenure < 6 yrs.	0.0254	0.0268	0.0206	0.0402
6 yrs. <= Tenure < 7 yrs.	0.0220	0.0321	0.0264	0.0110
7 yrs. <= Tenure < 8 yrs.	0.0227	0.0146*	0.0237	0.0301
8 yrs. <= Tenure < 9 yrs.	0.0251	0.0197*	0.0285	0.0270
9 yrs. <= Tenure < 10 yrs.	0.0360	0.0165*	0.0267	0.0256
10 yrs. <= Tenure < 11 yrs.	0.0193	0.0067*	0.0101*	0.0194*
11 yrs. <= Tenure < 12 yrs.	0.0396	0.1227	0.0228*	0.0139*
12 yrs. <= Tenure	0.0285	-0.0262*	0.0210*	-0.0040*
Residual term i.d. (13 terms)	yes	yes	yes	yes
χ^2 statistic (degrees of freedom)	80 (140)	38 (140)	71 (140)	63 (140)
Number of workers	5 649	1 515	3 010	1 702
Number of observations	116 715	20 910	58 539	32 204

Note - * Not significant at the 5% level.

Table 5. Covariance Structure of Residuals as Function of Tenure, Experience, and Unobserved Components (Entire Sample) (Standard Errors in Parentheses)

Parameter	Hause's model without match component	Hause's model with match component	Model 4
Variance of unmeasured fixed worker ability- σ^2_α	0.0494 (0.0027)	0.0289 (0.0029)	0.0235 (0.0031)
Variance of experience slope $\sigma^2_{b_1}$	0.0008 (0.0001)	0.0004 (0.0001)	0.0003 (0.0001)
Covariance of experience slope and worker-specific intercept- $\sigma(b_1, \alpha)$	-0.0026 (0.0004)	-0.0009 (0.0004)	-0.0001 (0.0005)
Variance of unobserved match quality (job-specific intercept) σ^2_θ	-	0.0258 (0.0014)	0.0372 (0.0025)
Variance of tenure slope $\sigma^2_{b_2}$	-	-	0.0005 (0.0001)
Covariance of tenure slope and job-specific intercept $\sigma(b_2, \theta)$	-	-	-0.0029 (0.0005)
Residual term i.d. (13 terms)	Yes	Yes	Yes
χ^2 statistic (degrees of freedom)	204 (257)	108 (334)	142 (514)
Number of workers	5 649	5 649	5 649
Number of observations	116 715	116 715	116 715

Table 6. Evolution of Parameters of Quadratic Functions from Job to Job (Entire Sample)

Parameters	Job 1 σ^2 σ^2_{b2} $\sigma(b_2, \theta)$	Job 2 σ^2 σ^2_{b2} $\sigma(b_2, \theta)$	Job 3 σ^2 σ^2_{b2} $\sigma(b_2, \theta)$	Job 4 σ^2 σ^2_{b2} $\sigma(b_2, \theta)$	Job 5+ σ^2 σ^2_{b2} $\sigma(b_2, \theta)$	Residual term i.d. (13 terms)	χ^2 statistic (degrees of freedom)	Number of workers	Number of observations
Model 4 Expanded	0.0322* 0.0007 -0.0024*	0.0431* 0.0009* -0.0042*	0.0350* 0.0015* -0.0030	0.0287* 0.0017 -0.0027	0.0500* 0.0033 -0.0058	Yes	133 (502)	5 649	116 715

Note. * Significant at the 5% level. Other parameters not shown are very similar to those estimated in Table 5.

Table 7. Covariance Structure of Residuals as Function of Tenure, Experience, and Unobserved Components (by Education Level)
(Standard Errors in Parentheses)

Parameter	Less than H.S. Education	High School Education	Above H.S. Education
Variance of unmeasured fixed worker ability- σ^2_α	0.0263 (0.0070)	0.0159 (0.0036)	0.0436 (0.0075)
Variance of experience slope- σ^2_{b1}	0.0008 (0.0002)	0.0003 (0.0001)	0.0004 (0.0001)
Covariance of experience slope and worker-specific intercept- $\sigma(b1, \alpha)$	-0.0020 (0.0011)	0.0001 (0.0005)	-0.0015 (0.0010)
Variance of unobserved match quality (job-specific intercept)- σ^2_θ	0.0244 (0.0054)	0.0330 (0.0029)	0.0414 (0.0048)
Variance of tenure slope σ^2_{b2}	-0.0003 (0.0003)	0.0006 (0.0002)	0.0004 (0.0002)
Covariance of tenure slope and job-specific intercept $\sigma(b2, \theta)$	0.0001 (0.0013)	-0.0027 (0.0006)	-0.0029 (0.0010)
Residual term i.d. (13 terms)	Yes	Yes	Yes
χ^2 statistic (degrees of freedom)	121 (514)	166 (514)	153 (514)
Number of workers	1 515	3 010	1 702
Number of observations	20 910	58 539	32 204

Table 8. Evolution of Parameters of Quadratic Functions from Job to Job (by Education Level)

Parameters	Job 1 σ^2_θ σ^2_{b2} $\sigma(b2, \theta)$	Job 2 σ^2_θ σ^2_{b2} $\sigma(b2, \theta)$	Job 3 σ^2_θ σ^2_{b2} $\sigma(b2, \theta)$	Job 4 σ^2_θ σ^2_{b2} $\sigma(b2, \theta)$	Job 5+ σ^2_θ σ^2_{b2} $\sigma(b2, \theta)$	Residual term i.d. (13 terms)	χ^2 statistic (degrees of freedom)	Number of workers	Number of observations
Above H.S. Education	0.0457* 0.0007* -0.0039*	0.0580* 0.0008	0.0147* 0.0014	0.0255* 0.0004	0.0371 -0.0001 0.0001	Yes	148 (502)	1 702	32 204
High School Education	0.0257* 0.0007* -0.0019*	0.0330* 0.0010* -0.0039*	0.0348* 0.0009	0.0340* 0.0031* -0.0005*	0.1242* 0.0123* -0.0338*	Yes	160 (502)	3 010	58 539
Less than H.S. Education	0.0205* -0.0002 0.0003	0.0083 -0.0013 0.0041	0.0254* -0.0008	0.0437* -0.0011 -0.0015	-0.0028 0.0023 0.0043	Yes	113 (502)	1 702	20 910

Note. * Significant at 5% level.

Table 9. Evolution of the Variance of the Job-Match Component from Job 1 to Job 5 and Beyond with Quadratic Function in Experience (Standard Errors in Parentheses)

Parameter	Total Sample	Less than H.S. Education	High School Education	Above H.S. Education
Variance of unmeasured fixed job-match quality	0.0260 (0.0018)	0.0238 (0.0045)	0.0222 (0.0021)	0.0281 (0.0036)
$\sigma^2_{\theta}(\text{job } 1)$	0.0270 (0.0022)	0.0198 (0.0050)	0.0203 (0.0025)	0.0366 (0.0042)
$\sigma^2_{\theta}(\text{job } 2)$	0.0249 (0.0032)	0.0220 (0.0069)	0.0324 (0.0038)	0.0159 (0.0060)
$\sigma^2_{\theta}(\text{job } 3)$	0.0177 (0.0050)	0.0371 (0.0098)	0.0222 (0.0064)	0.0254 (0.0104)
$\sigma^2_{\theta}(\text{job } 4)$	0.0306 (0.0072)	0.0088 (0.0141)	0.0452 (0.0098)	0.0361 (0.0130)
Residual term i.d. (13 terms)	Yes	Yes	Yes	Yes
χ^2 statistic (degrees of freedom)	107 (332)	121 (332)	104 (332)	111 (332)
Number of workers	5 649	1 515	3 010	1 702
Number of observations	116 715	20 910	58 539	32 204

Note. Parameters not shown include variance of worker ability, variance of experience slope, and covariance of worker-specific intercept and experience slope.

Table 10. Evolution of the Variance of the Job-Match Component within Jobs with Quadratic Function in Experience (Standard Errors in Parentheses)

Parameter	Total Sample	Less than H.S. Education	High School Education	Above H.S. Education
Variance of unmeasured fixed job-match quality	0.0434	0.0204	0.0486	0.0792
1 mo. <= Tenure < 4 mo.	0.0346	0.0230	0.0234	0.0419
4 mo. <= Tenure < 1 yr.	0.0292	0.0240	0.0306	0.0281
1 yr. <= Tenure < 2 yrs.	0.0289	0.0261	0.0270	0.0305
2 yrs. <= Tenure < 3 yrs.	0.0309	0.0289	0.0251	0.0354
3 yrs. <= Tenure < 4 yrs.	0.0252	0.0087*	0.0253	0.0249
4 yrs. <= Tenure < 5 yrs.	0.0241	0.0277	0.0198	0.0318
5 yrs. <= Tenure < 6 yrs.	0.0225	0.0227	0.0182	0.0375
6 yrs. <= Tenure < 7 yrs.	0.0180	0.0284	0.0228	0.0081*
7 yrs. <= Tenure < 8 yrs.	0.0167	0.0084*	0.0184	0.0263
8 yrs. <= Tenure < 9 yrs.	0.0163	0.0121*	0.0215	0.0227
9 yrs. <= Tenure < 10 yrs.	0.0261	0.0093*	0.0186	0.0200*
10 yrs. <= Tenure < 11 yrs.	0.0084	0.0036*	0.0008*	0.0129*
11 yrs. <= Tenure < 12 yrs.	0.0263	0.1035	0.0129*	0.0057*
12 yrs. <= Tenure	0.0095*	-0.0511*	0.0070*	-0.0223*
Residual term i.d. (13 terms)	yes	yes	yes	yes
χ^2 statistic (degrees of freedom)	102 (320)	72 (320)	103 (320)	106 (320)
Number of workers	5 649	1 515	3 010	1 702
Number of observations	116 715	20 910	58 539	32 204

Notes. *Not significant at 5% level. Parameters not shown include variance of worker ability, variance of experience slope, and covariance of worker-specific intercept and experience slope.

Figure 1. Within-Job Evolution of Variance
Total Sample-NLSY

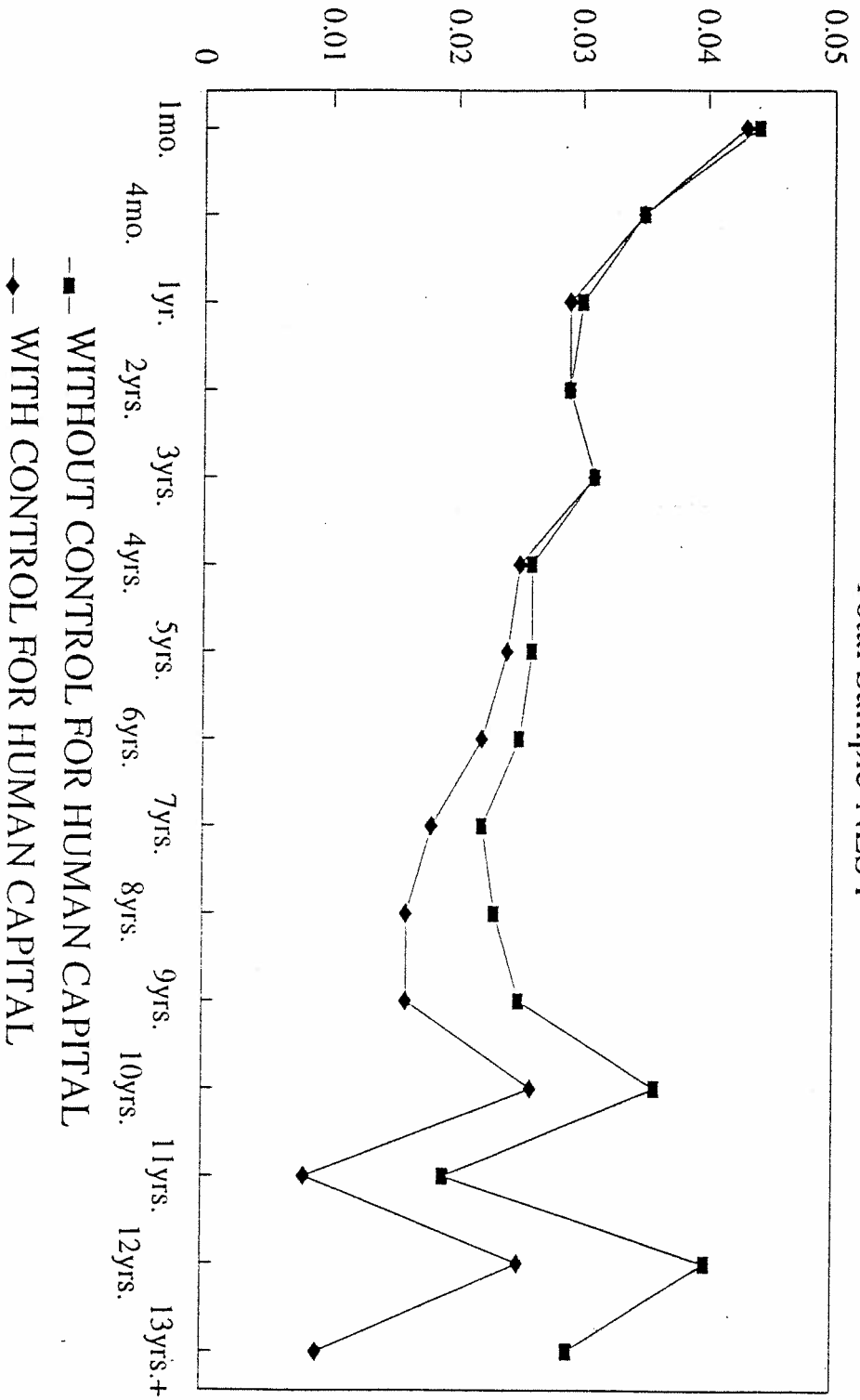


Figure 2. Within-Job Evolution of Variance
 Workers with Less than a High School Diploma

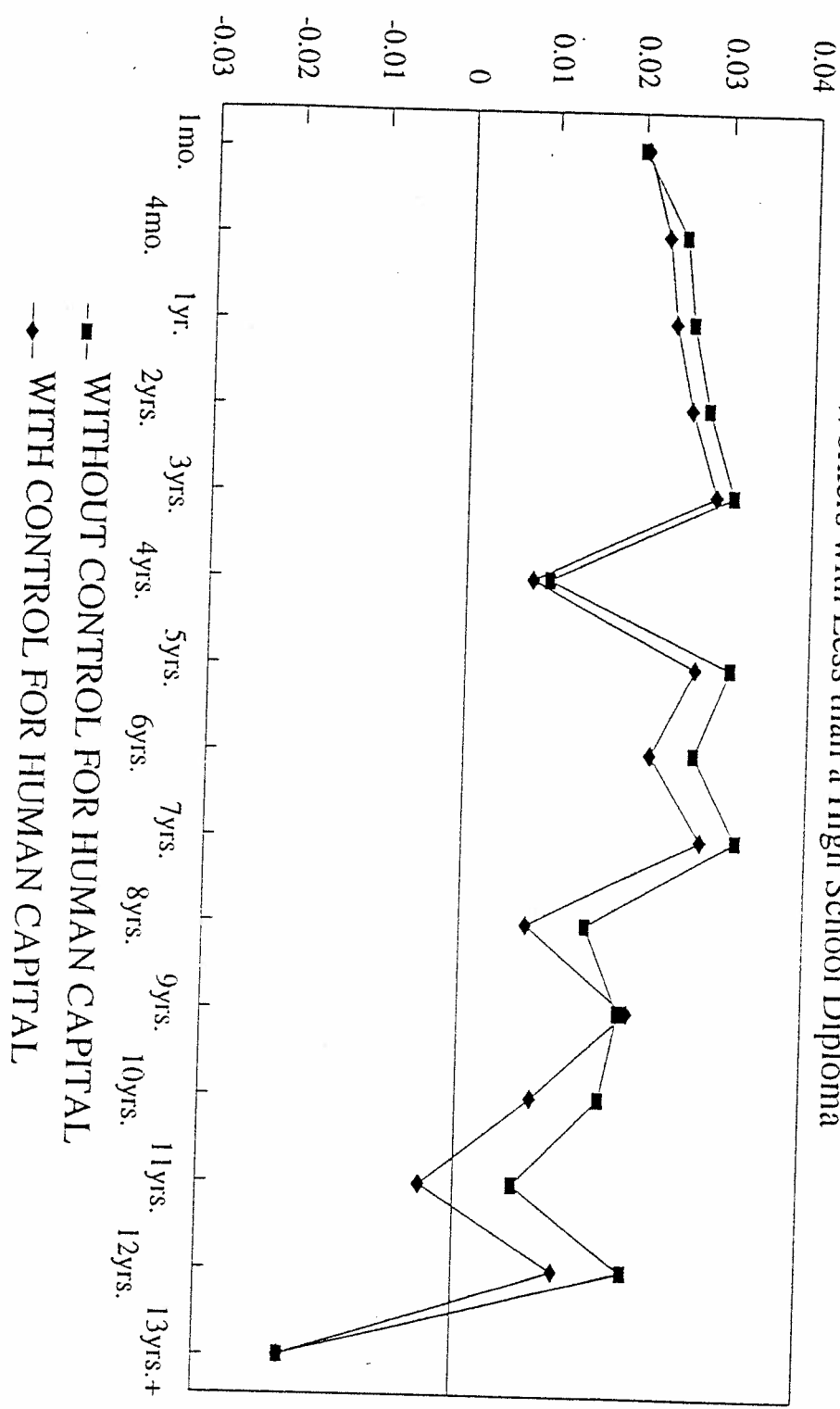


Figure 3. Within-Job Evolution of Variance
 Workers with a High School Diploma

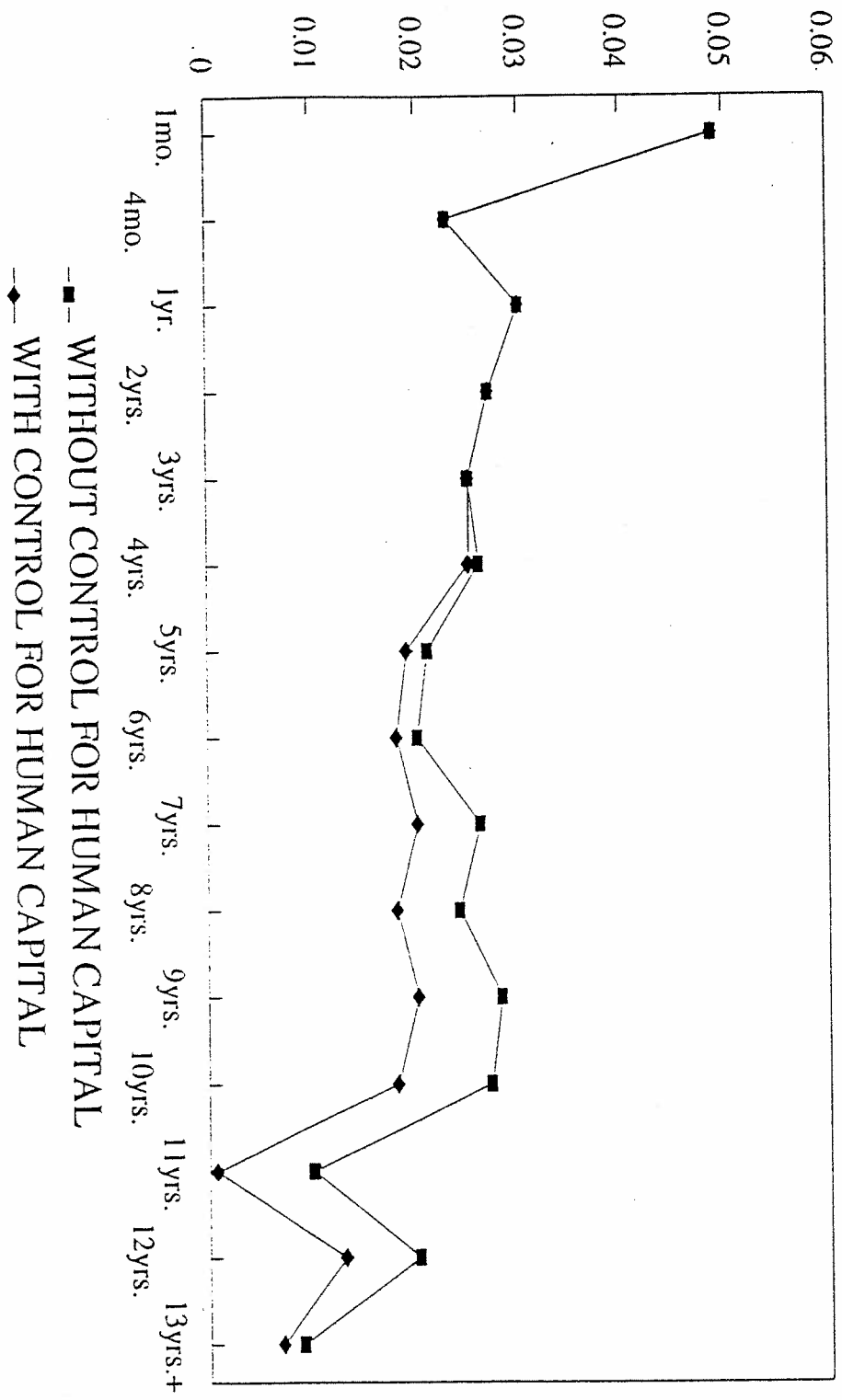
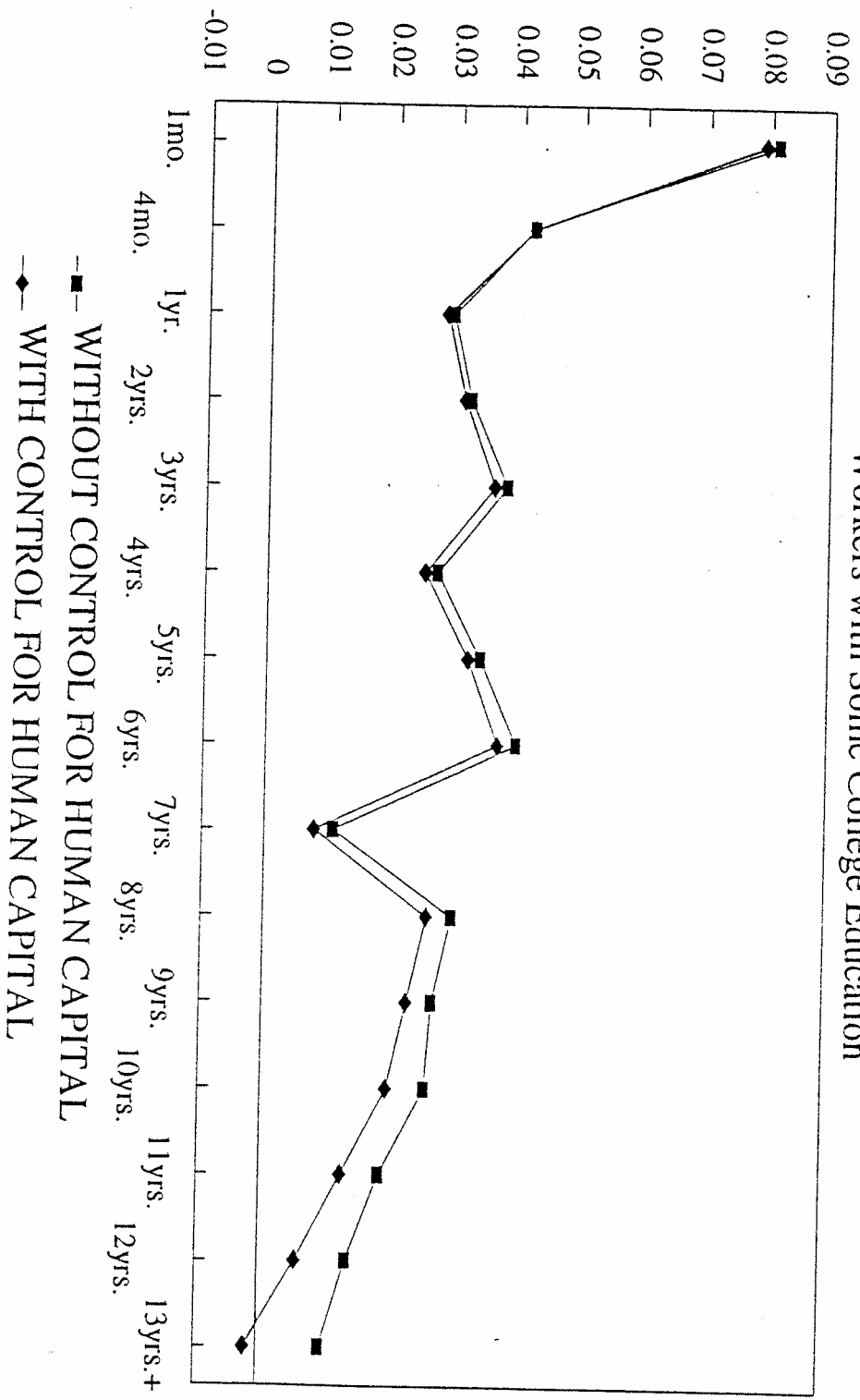


Figure 4. Within-Job Evolution of Variance
 Workers with Some College Education



Appendix

A.1 The Minimum Distance Procedure.

The following is a straightforward adaptation of Chamberlain (1982,1984) and especially Abowd and Card (1989). Let e_i ($i=1, \dots, N$) represent the $T \times 1$ vector of the residuals of individual i estimated in equation (27). It is assumed that e_i is independently and identically distributed across individuals. Therefore it follows that e_i , which represents individual i 's cross-product matrix of dimension $T \times T$ is also i.i.d. Let M be the estimated covariance matrix and m the $(T+1)/2$ column vector formed with the distinct elements of M . Fourth-moments are assumed to be finite. A typical element of m , $[m_s]$, represents the average of the individual contributions to this sample covariance over all workers at periods t and s , i.e.

$$m = \frac{1}{N} \sum_{i=1}^N m_i$$

$$m_i = \text{vec}(e_i e_i')$$

(37)

We wish to impose restrictions on m , i.e. we want m to depend on a lower-dimensional vector $f(\beta)$. More rigorously, let R^p be the p -dimensional Euclidean space and E be a compact subset of R^p that contains the true parameter vector β^0 where $p \leq (T+1)/2$. Then f is a continuous, one-to-one mapping of E into R^q with a continuous inverse where $q = (T+1)/2$. Also, f has continuous second partial derivatives in E , the rank of the matrix $[df(\beta)/d\beta']$ is p for $\beta^0 \in E$, and the covariance matrix M estimated with β is non-singular for $\beta \in E$. Under these regularity conditions, the assumption that the m_i 's are i.i.d. allow us to invoke the Central Limit Theorem (Lindeberg-Lévy version) which insures that the sample mean of m_i converges asymptotically to a normal distribution:

importance in small samples. On that subject, see Altonji and Segal (1994). acceptable or not. Note, however, that the choice of the weighting matrix is of some choose, we can then determine whether the restrictions imposed by the underlying model are which is known to converge in distribution to a $\chi^2(q-p)$. Given the level of the test that we weighting matrix. Finally, from this estimator we calculate the statistic given by equation (41), compute the generalized least squares estimator β using the covariance matrix of m as the try to fit with a parsimonious representation suggested by the model we want to test. We then (27). Then we form a column vector m with the $(T+1)/2$ distinct elements of M which we we compute the empirical covariance matrix M using the residuals estimated with equation zero is this statistic, the better is the fit of the underlying model. To summarize the procedure, has an asymptotic χ^2 distribution with $q-p$ degrees of freedom. Then, of course, the closer to

$$N(m - f(\beta))' V^{-1} (m - f(\beta)) \quad (41)$$

Under the hypothesis of a correct specification, the statistic

$$\beta = (X'V^{-1}X)^{-1}X'V^{-1}m \quad (40)$$

Where V is the sample analog of Ω . In the case where f is linear in β (e.g. $f(\beta) = X\beta$)

$$\beta = \underset{\beta \in \Omega}{\operatorname{argmin}} [m - f(\beta)]' V^{-1} (m - f(\beta)) \quad (39)$$

$$V = \frac{1}{N} \sum_{t=1}^{N+1} m_t m_t' - m m'$$

Now to test restrictions on m , Chamberlain shows that the following estimator is optimal:

$$\Omega = E(m_t m_t') - E(m_t)E(m_t')$$

$$\mu = E(m_t)$$

$$\sqrt{N}(m - \mu) \xrightarrow{d} N(0, \Omega) \quad (38)$$

A.2 Gallant and Jorgenson's Procedure

Instead of directly fitting sample moments as done in the previous paragraph, we fit each individual contributions to the corresponding sample second moments. Just to make things clear, in the previous paragraph we computed the sample average of the m_t , giving us m which we then try to fit with a lower-dimensional parameter vector through the f mapping. Here we stack all m_t into a $NT(T+1)/2$ column vector M and we fit M in the same manner, except that the f function can now be individual-specific if the model warrants it. Let's suppose, for comparison purposes, that the f mapping is not individual-specific. The objective is then to minimize the following function:

$$S(\beta) = \sum_{t=1}^{T+1} (m_t - f(\beta))' V^{-1} (m_t - f(\beta)) \quad (42)$$

Where V^{-1} is computed with the cross-products of the residuals from the following unrestricted model:

$$\begin{aligned} E(\epsilon_{1t}^2) &= m_{1t} \\ E(\epsilon_{2t}^2) &= m_{2t} \end{aligned} \quad (43)$$

Consequently, we can see that the V matrix is the same as in the optimal minimum distance estimator. Let $S(\beta)_1$ be the value of the objective function for the unrestricted model and $S(\beta)_2$ be defined likewise for the restricted version of, say, model 1 (equation (29)). For more general (i.e. non-linear) models, Gallant and Jorgenson (1979) show that $T_0 = S(\beta)_2 - S(\beta)_1$ is asymptotically distributed as a chi-square with r -s degrees of freedom where r -s is equal to the difference in the number of parameters in the two models.

Thus, it is clear that there is a close connection between the two estimation procedures. Namely, if the sample is balanced and the X matrix (the design matrix) is common to all individual, which is the case when we wish to test time series representations of the

error process, then the vector of estimated parameters is identical and the Chi-Square statistic is the same. Let β_{OMD} be the minimum distance estimator, β_{GL} the Gallant & Jorgenson estimator and $f(\beta) = X\beta$. Then,

$$\beta_{\text{OMD}} = (X'V^{-1}X)^{-1}X'V^{-1}m \quad = (X'V^{-1}X)^{-1}X'V^{-1}N^{-1}N^{-1}\sum_{t=1}^{T-1} m_t \quad (44)$$

whereas the formula for the estimator that makes direct use of each individual contributions is

$$\beta_{\text{GL}} = \left(\sum_{t=1}^{T-1} X_t'V_t^{-1}X_t \right)^{-1} \left(\sum_{t=1}^{T-1} X_t'V_t^{-1}m_t \right) \quad = \beta_{\text{OMD}} \quad (45)$$

However, the numerical equivalence of the two estimators does not hold in the case of unbalanced data. There will be a discrepancy between the two and this can be shown with a simple example. Let's assume that we have a sample of four individuals who are present in the data for a maximum of three periods. Without loss of generality, let V be the identity matrix. We wish to test the restriction that only a white noise error term can fit the covariance structure. Therefore, only the empirical variances at periods one, two and three are used in the computation of the coefficient. Let each m_t be given by $m_1=(s_{11}, s_{22}, s_{33})$, $m_2=(s_{22}, s_{33})$, $m_3=(s_{11}, s_{33})$ and $m_4=(s_{33})$. Then,

The estimator that makes direct use of the individual contributions to the sample moments is less susceptible to sampling error in that it gives relatively more weight to the covariance elements which are computed over a greater number of individuals. On the other hand, the minimum distance estimator places an equal weight on each sample moment; it does not matter whether some of them were computed over 3 individuals or 3000 individuals. In fact, even if the f function is not individual-specific, the design matrix X becomes individual-specific simply because all individuals are not present in the data at the same time. Furthermore, the Chi-Square statistic computed with the optimal minimum distance estimator fails to adjust for the true degrees of freedom which is not the case for the Gallant and Jorgenson procedure which directly accounts for the fact that the sample is unbalanced. Naturally, if the f mapping is individual-specific, which is the case for the models in this paper, then the optimal minimum distance estimator is not applicable.

$$\beta^{OMD} = (X'X)^{-1}X'm = \frac{1}{3} \left(\frac{2}{s_{11} + s_{13}} + \frac{2}{s_{22} + s_{23}} + \frac{2}{s_{33} + s_{34}} \right)$$

$$\beta^{GL} = \sum_4^{l-1} (X'_i X_i)^{-1} \sum_4^{l-1} X'_i m_i = \frac{8}{s_{11} + s_{13}} + \frac{8}{s_{22} + s_{23}} + \frac{8}{s_{33} + s_{34}}$$

(46)

Par contre, et ce résultat est à mon sens assez fondamental, il semble bien que le capital humain spécifique à la firme soit un facteur explicatif beaucoup moins important en ce qui concerne le salaire des travailleurs. Nous avons vu dans la troisième étude que le simple fait d'ajouter l'expérience dans l'industrie annulait à toutes fins utiles l'effet positif des années d'ancienneté. Souignons au passage la robustesse de ce résultat par rapport aux méthodes d'estimation ainsi que par rapport aux données utilisées. À la lumière de ce résultat, il faut en conclure que les estimés de Topel (1991) étaient le résultat d'une erreur de spécification. Il faut également conclure que, tout comme dans l'étude précédente, le processus de formation des salaires semble correspondre au modèle néo-classique concurrentiel. Par ailleurs, notons également que si le rendement moyen découlant de l'ancienneté dans la firme semble être nul,

Je crois qu'il faut tout d'abord souligner l'impact substantiel du processus d'acquisition du capital humain sur le profil de salaire des travailleurs. Que ce soit par le biais de mesures directes, comme le nombre de semaines passées en programme de formation, ou par les mesures indirectes que sont l'expérience sur le marché du travail et l'ancienneté dans la firme, il apparaît clair que l'impact positif de ces variables n'est pas seulement dû à des effets de sélection. Si c'était le cas, la méthode des variables instrumentales utilisée dans les études deux et trois aurait éliminé ce biais. Plus précisément, les résultats trouvés dans la deuxième étude contredisent en partie ceux de Lynch (1992) qui avait pourtant utilisé la même banque de données (quoique sur des périodes plus courtes). Le capital humain acquis avec les employeurs précédents par l'entremise de programmes de formation a un impact statistiquement et économiquement significatif. De plus, le fait que les individus paient pour ces programmes de formation par l'entremise de salaires de départ plus faibles et qu'ils ne semblent pas bénéficier de prime salariale liée à leur capital spécifique à la firme nous laisse croire en la pertinence du modèle concurrentiel.

Conclusion générale

il y a tout de même des variations selon la profession exercée par les individus.

Ces variations nous donnent à croire que les économistes font fausse route en partie lorsqu'ils imposent de façon arbitraire que l'hétérogénéité individuelle (par exemple) se résume à supposer que les individus n'ont que des ordonnées à l'origine différentes dans leurs profils de salaire, et non des pentes différentes. Il n'y a aucune raison de penser, a priori, que cela soit nécessairement le cas. Bien sûr, accepter cette hypothèse s'avère pratique et permet de contourner certaines difficultés, tout en allant à l'essentiel, si tant est que l'essentiel se résume à énoncer des résultats quant aux rendements moyens de l'ancienneté et de l'expérience. Cependant, il se trouve qu'une théorie concurrente comme la théorie du matching est parfaitement en mesure de justifier ces résultats et ce, pour des raisons indépendantes du processus d'acquisition du capital humain. À cet égard, l'apport principal de la quatrième étude de cette thèse est de permettre de faire la distinction entre ce qui relève du capital humain et ce qui relève du matching. En portant notre attention sur les moments d'ordre deux d'une équation de salaire, nous pouvons arriver à conclure avec un fort degré de confiance que le capital humain joue un rôle majeur. Par la même occasion, nous sommes en mesure d'estimer la variance dans le rendement lié à l'ancienneté dans la firme. Ce résultat met en lumière le fait qu'on ne peut pas à vrai dire parler d'un effet d'ancienneté, mais bien d'un éventail de rendements selon les pairages entre firmes et travailleurs. En même temps, nous pouvons également mettre en relief la pertinence du modèle de matching puisqu'une simple structure de covariance où l'on tient compte de la qualité du pairage entre la firme et l'individu ne peut pas être rejetée par les données.

- Abowd, J.M., D. Card, "On the Covariance Structure of Earnings and Hours Changes", *Econometrica*, 57 (1989): 411-445.
- Abowd, J., F. Kramarz, D. Margolis, "High-Wage Workers and High-Wage Firms", INSEE Working Paper, September 1994.
- Abraham, K. & H. Farber, "Job Duration, Seniority, and Earnings," *American Economic Review* 77 (June 1987): 278-97.
- Altonji, J.G., L.M. Segal, "Small Sample Bias in GMM Estimation of Covariance Structures", NBER Technical Working Paper No. 156, June 1994.
- Altonji, J., & R. Shakotko, "Do Wages Rise with Seniority?" *Review of Economic Studies* 54 (July 1987): 437-59.
- Altonji, J.G., J. Spletzer, "Worker Characteristics, Job Characteristics, and the Receipt of On-the-Job Training", *Industrial and Labor Relations Review*, vol. 45, No. 1 (October 1991): 58-79.
- Altonji, J.G., N. Williams, "The Effects of Labor Market Experience, Job Seniority, and Job Mobility on Wage Growth", NBER Working Paper No. -, 1993.
- Barton, J., M. Berger, D. Black, "Do Workers Pay for On-the-Job Training?" Center for Business and Economic Research, College of Business and Economics, University of Kentucky. Working paper no. E-169-93. Lexington, Kentucky, 1993.
- Barton, J., D. Black, Loewenstein, M., "Job Matching and On-the-Job Training", *Journal of Labor Economics*, 1989, vol. 7, no. 1.
- Bartel, A., "Training, Wage Growth, and Job Performance: Evidence from a Company Database", NBER Working Paper No. 4027, 1992.
- Bartel, A.P., "Wages, Nonwage Job Characteristics, and Labor Mobility", *Industrial and Labor Relations Review*, 35 (1982): 578-588.
- Bartel, A.P., G.J. Borjas, "Wage Growth and Job Turnover: An Empirical Analysis", in S. Rosen ed., *Studies in Labor Markets*, University of Chicago, 1981: 65-84.
- Becker, G.S., "Investment in Human Capital: A Theoretical Analysis", *Journal of Political*

BIBLIOGRAPHIE GÉNÉRALE

- Becker, G.S., *Human Capital*, The University of Chicago Press, 1975.
- Becker, G.S., B.R. Chiswick, "Education and the distribution of Earnings", *American Economic Review, Proceedings*, 56, 1966: 358-369.
- Ben-Porath, Y., "The Production of Human Capital and the Life-Cycle of Earnings", *Journal of Political Economy*, August 1967: 352-365.
- Bishop, J.H., "The Impact of Previous Training on Productivity and Wages", in Lisa M. Lynch ed., *Training and the Private Sector*, NBER Comparative Labor Markets Series, The University of Chicago Press, 1994: 161-199.
- Blanchflower, D.G., L.M. Lynch, "Training at Work: A Comparison of U.S. and British Youths", in Lisa M. Lynch ed., *Training and the Private Sector*, NBER Comparative Labor Markets Series, The University of Chicago Press, 1994: 233-260.
- Borjas, G.J., "Job Mobility and Earnings over the Life Cycle", *Industrial and Labor Relations Review*, 34 (1981): 365-376.
- Brown, J., "Why Do Wages Increase with Tenure? On-the-Job Training and Life-Cycle Wage Growth Observed within Firms." *American Economic Review* 79 (December 1989): 971-91.
- Brownstone, D., R. Valetta, "Modeling Measurement Error Bias in Cross-Section and Longitudinal Wage Equations", University of California (Irvine) Working Paper, 1993.
- Burdett, K., "A Theory of Employee Job Search and Quit Rates", *American Economic Review*, 68, 1978: 212-220.
- Carmichael, L., "Firm-Specific Capital and Promotion Ladders", *Bell Journal of Economics* 14, 1983: 251-258.
- Chamberlain, G., "Multivariate Regression Models for Panel Data", *Journal of Econometrics*, 18, 1982: 5-46.
- Chamberlain, G., "Panel Data", *Handbook of Econometrics*, volume 2, chap. 22, 1986.
- Duncan, G.J., S. Hoffman, "On-the-Job Training and Earnings Differences by Race and Sex", *The Review of Economics and Statistics*, 61, November 1979: 594-603.

Farber, H.S., R. Gibbons, "Learning and Wage Dynamics", NBER Working Paper No. 3764, 1991.

Finnie, R., "Tenure, Experience, and Men's and Women's Wages: Panel Estimates from the National Longitudinal Survey of Youth", Groupe de Recherche en Politique Economique, cahier de recherche 9305, Université Laval, 1993.

Finn, C.J., "Wages and Job Mobility of Young Workers", *Journal of Political Economy*, 94, 1986: S88-S110.

Gallant, R., D.W. Jorgenson, "Statistical Inference for a System of Simultaneous, Non-Linear, Implicit Equations in the Context of Instrumental Variable Estimation", *Journal of Econometrics*, 11 (1979): 275-302.

Hashimoto, M., "Firm-Specific Human Capital as a Shared Investment", *American Economic Review* vol.71, no. 3, June 1981: 475-482.

Hause J.C., "The Fine Structure of Earnings and the On-the-Job Training Hypothesis", *Econometrica*, 48, 1980: 1013-1029.

Heckman, J.J., "A Life-Cycle Model of Earnings, Learning, and Consumption", *Journal of Political Economy*, 84, 1976: S11-S44.

Hutchens, R.M., "Seniority, Wages and Productivity: A Turbulent Decade", *Journal of Economic Perspective*, vol.3, no.4, 1989: 49-64.

Intriligator, M.D., *Mathematical Optimization and Economic Theory*, Prentice-Hall Inc., Englewood Cliffs, N.J., 1971.

Jovanovic, B., "Job Matching and the Theory of Turnover", *Journal of Political Economy* 87 (1979a): 972-90.

Jovanovic, B., "Firm-specific Capital and Turnover" *Journal of Political Economy* 87 (1979b): 1246-60.

Jovanovic, B., "Matching, Turnover, and Unemployment" *Journal of Political Economy* 92 (1984): 108-122.

Kiefer, N.M., "Economic Duration Data and Hazard Functions", *Journal of Economic Literature*, vol. XXVI (June 1988): 646-679.

Lancaster, T., *The Econometric Analysis of Transition Data*, Econometric Society Monographs No. 17, Cambridge University Press, 1990.

- Lazear, E.P., "Why Is There Mandatory Retirement?", *Journal of Political Economy* 87, December 1979: 1261-84.
- Levine, D.I., "Worth Waiting for? Delayed Compensation, Training, and Turnover in the United States and Japan. *Journal of Labor Economics*, vol. 11, no.4, October 1993
- Lillard, L.A., H.W. Tan, "Private Sector Training: Who Gets it and What Are its Effects?", *Research in Labor Economics*, Volume 13, 1992: 1-62.
- Lynch, L., "The Role of Off-the-Job vs. On-the-Job Training for the Mobility of Women Workers." *A.E.R. Papers and Proceedings* May 1991: 151-56.
- Lynch, L., "Private-Sector Training and the Earnings of Young Workers." *American Economic Review* 82 (March 1992): 299-312.
- MacCurdy, T., "An Intertemporal Model of Portfolio Choice and Human Capital Accumulation under Uncertainty with Extensions Incorporating Taxes, Consumer Durables, Imperfection in Capital Market, and Nonseparable Preferences", Working Paper No. E-80-4, Domestic Studies Program, Hoover Institute, Stanford University, 1981.
- Marshall, R.C., G.A. Zarkin, "The Effect of Job Tenure on Wage Offers", *Journal of Labor Economics*, 1987, vol. 5: 301-324.
- McCall, J.J., "Economics of Information and Job Search", *Quarterly Journal of Economics*, 84, 1970: 113-126.
- Medoff, J., K. Abraham, "Experience, Performance, and Earnings", *Quarterly Journal of Economics* 95, December 1980: 703-36.
- Medoff, J., K. Abraham, "Are Those Paid More Really More Productive?", *The Journal of Human Resources*, 41, 1981: 186-216.
- Meyer, B.D., "Unemployment Insurance and Unemployment Spells", *Econometrica*, vol. 58, No.4 (July, 1990): 757-782.
- Mincer, J., "Investment in Human Capital and Personal Income Distribution", *Journal of Political Economy*, 67, 1958: 281-302.
- Mincer, J. " On-the-Job Training: Costs, Returns, and Some Implications", *Journal of Political Economy*, Supplement, October 1962: 50-79.

- Mincer, J., "Job Training, Wage Growth, and Labor Turnover." Working Paper no. 2690. Cambridge, Mass.: National Bureau of Economic Research, August 1988.
- Mincer, J., "Job Training: Costs, Returns, and Wage Profiles." Working Paper no. 3208. Cambridge, Mass.: National Bureau of Economic Research, December 1989.
- Mincer, J., Y. Higuchi, "Wage Structures and Labor Turnover in the U.S. and in Japan", NBER Working Paper No. 2306, 1987.
- Mincer, J., B. Jovanovic, "Labor Mobility and Wages", in S. Rosen, ed., *Studies in Labor Markets*, The University of Chicago Press, 1981.
- Mortensen, D.T., "Job Search and Labor Market Analysis", *Handbook of Labor Economics*, Volume 2, Chap. 15, 1986: 849-919.
- Neal, D., "Industry-Specific Human Capital: Evidence from Displaced Workers", University of Chicago Working Paper, 1993.
- Oi, W. Y., "Labor as a Quasi-Fixed Factor", *Journal of Political Economy*, 70 (1962): 538-555.
- Parsons, D. O., "Specific Human Capital: An Application to Quit Rates and Layoff Rates", *Journal of Political Economy*, 80, 1972: 1120-1143.
- Rothschild, M., "Models of Market Organization with Imperfect Information: A Survey", *Journal of Political Economy*, 81, 1973: 1283-1308.
- Sargent, T. J., *Dynamic Macroeconomic Theory*, Harvard University Press, 1987.
- Schultz, T. W., "Investment in Human Capital", *American Economic Review*, 51 (1961): 1-17.
- Stigler, G. J., "The Economics of Information", *Journal of Political Economy*, 69, 1961: 213-225.
- Topel, R., "Specific Capital, Mobility, and Wages: Wages Do Rise with Seniority." *Journal of Political Economy* 99 (February 1991): 145-76.
- Topel, R. H., M. P. Ward, "Job Mobility and the Careers of Young Men", *Quarterly Journal of Economics*, 1992: 439-479.
- Weiss, Y., "The Determination of Life-Cycle Earnings: A Survey", *Handbook of Labor*

Economics, Volume 1, Chap. 11, 1986: 603-640

Williams, N., "Reexamining the Wage, Tenure and Experience Relationship", *The Review of Economics and Statistics*, 1991: 512-517.

Williamson, O.E., M. Wachter and J.E. Harris, Understanding the Employment Relationship: The Analysis of Idiosyncratic Exchange, *Bell Journal of Economics* 6, 1975: 250-280.

Willis, R.J., "Wage Determinants: A Survey and Reinterpretation of Human Capital Earnings Functions", *Handbook of Labor Economics*, Volume 1, Chap. 10, 1986:525-602.