

Université de Montréal

Économie comportementale : Retrouve-t-on un effet d’ancrage dans la LNH ?

Par
Samuel F. PICHETTE

Département des sciences économiques
Faculté des arts et sciences

Mémoire présenté à la Faculté des arts et sciences
en vue de l’obtention du grade de maîtrise
en sciences économiques

Août 2016

©, Samuel F. PICHETTE, 2016

RÉSUMÉ

De par leur nature scientifique, les sciences économiques visent, entre autre, à observer, qualifier, ainsi que quantifier des phénomènes économiques afin de pouvoir en dégager diverses prévisions. Ce mémoire se penche sur ces prévisions et, plus particulièrement, sur les facteurs pouvant biaiser les prévisionnistes au niveau comportemental en référant à l'effet d'ancrage, un biais propre à l'économie comportementale – une sous-discipline des sciences économiques. Il sera donc question de comprendre, par une analyse selon la discipline que représente l'économie comportementale, ce qui peut les affecter, avec un accent mis sur l'effet d'ancrage plus précisément. L'idée générale de ce dernier est qu'un agent peut être biaisé inconsciemment par la simple connaissance d'une valeur précédente lorsqu'il est demandé de faire une estimation ultérieure. De cette façon, une analyse des salaires des joueurs de la Ligne Nationale de Hockey (NHL) selon leurs performances passées et leurs caractéristiques personnelles, de 2007 à 2016, a été réalisée dans ce travail afin d'en dégager de possibles effets d'ancrage. Il est alors possible de constater que les directeurs généraux des équipes de la ligue agissent généralement de façon sensible et rationnelle lorsque vient le temps d'octroyer des contrats à des joueurs mais, néanmoins, une anomalie persiste lorsqu'on porte attention au rang auquel un joueur a été repêché. Dans un tel contexte, il semble pertinent de se référer à l'économie comportementale afin d'expliquer pourquoi le rang au repêchage reste une variable significative huit ans après l'entrée d'un joueur dans la NHL et qu'elle se comporte à l'inverse de ce que prévoit la théorie à ce sujet.

Mots-clés : Économie comportementale – Effet d'ancrage – Biais d'ancrage – Ancrage – Biais – LNH – Économétrie – Prévion – Risque – Aversion au risque

ABSTRACT

Economic analysis, by its nature, involves observing, qualifying and quantifying economic data with the ultimate goal of making forecasts. In this masters thesis, I am interested in factors that could bias a forecaster's behavior – with special focus on phenomena, like the anchoring effect, that have been proposed in behavioral economics. At a fundamental level, the anchoring effect states that an agent's ability to accurately forecast may be affected by placing unwarranted emphasis on certain economic variables. To study this effect, I analyze how the salaries of National Hockey League (NHL) players are determined by the players' characteristics and past performance. From the results, it would appear that NHL general managers are generally sensible and rational when it comes to using historical data to make decisions about player salaries. However, there is a persistent anomaly regarding the draft position of a player. Although one would not expect the draft position to be very important after eight years of experience in the NHL, the analysis shows that it remains a significant determinant of player salary. Behavioral economics and more specifically, the anchoring effect, helps explain why this might be so.

Keywords : Behavioral economics – Anchoring effect – Anchoring Bias – Anchor – Bias – NHL – Econometrics – Forecast – Risk – Risk aversion

REMERCIEMENTS

J'aimerais tout d'abord remercier mon directeur pour ce mémoire, Dr. Sean Horan, pour m'avoir permis d'explorer un sujet qui m'intéresse, pour son ouverture d'esprit, pour ses recommandations ainsi que pour sa disponibilité tout au long de ce travail.

Je suis aussi très reconnaissant à l'égard de ma conjointe, Karolan, pour son aide en ce qui a trait à l'édition et à la révision de cet ouvrage, ainsi que pour son support inconditionnel au travers chaque étape que j'ai accomplie.

Enfin, je remercie ma famille et particulièrement mes parents, grâce à qui j'ai été capable de me rendre jusqu'ici, et qui sont toujours là lorsque j'ai besoin de leur support et de leurs encouragements.

TABLE DES MATIERES

INTRODUCTION	1
CHAPITRE 1 - REVUE DE LA LITTÉRATURE	4
1.1 – FONDEMENTS DE L'ÉCONOMIE COMPORTEMENTALE	4
1.2 – FONDEMENTS DE L'EFFET D'ANCRAGE	4
1.3 – ANCRAGE ET AJUSTEMENT	5
1.4 – ACCESSIBILITÉ SÉLECTIVE	6
1.5 – CHANGEMENT D'ATTITUDE	7
1.6 – FACTEURS AFFECTANT L'EFFET D'ANCRAGE	8
1.7 – PERTINENCE POUR CETTE RECHERCHE	9
CHAPITRE 2 - DESCRIPTION DES DONNÉES	10
CHAPITRE 3 - STATISTIQUE DESCRIPTIVE	14
3.1 – SOMMAIRE DES VARIABLES PRINCIPALES	14
3.2 – MODÈLE	16
CHAPITRE 4 - RÉSULTATS	18
4.1 – MODÈLE DE BASE	18
4.2 – PREMIÈRE VERSION DU MODÈLE	22
4.3 – DEUXIÈME VERSION DU MODÈLE	25
4.4 – TROISIÈME VERSION DU MODÈLE	28
4.5 – QUATRIÈME VERSION DU MODÈLE	31
4.6 – VERSION FINALE DU MODÈLE	35
4.7 – EXPLICATION DE L'ANOMALIE	38
CHAPITRE 5 - ROBUSTESSE DES RÉSULTATS	40
CHAPITRE 6 - DISCUSSION	45
6.1 – MODÈLE	45
6.2 – MÉTHODOLOGIE	45
CONCLUSION	47
ANNEXES	49
ANNEXE A	50
ANNEXE B	52
BIBLIOGRAPHIE	53

LISTE DES TABLEAUX

Tableau I : Notation principale _____	13
Tableau II : Variables des régressions _____	14
Tableau III : Régression chronologique du modèle de base _____	20
Tableau IV : Régression par performances du modèle de base _____	20
Tableau V : Régression chronologique du premier modèle _____	23
Tableau VI : Régression par performances du premier modèle _____	23
Tableau VII : Régression chronologique du deuxième modèle _____	26
Tableau VIII : Régression par performances du deuxième modèle _____	26
Tableau IX : Régression chronologique du troisième modèle _____	29
Tableau X : Régression par performances du troisième modèle _____	29
Tableau XI : Régression chronologique du quatrième modèle _____	32
Tableau XII : Régression par performances du quatrième modèle _____	32
Tableau XIII : Régression chronologique du modèle final _____	36
Tableau XIV : Régression par performances du modèle final _____	36
Tableau XV : Test de rationalité selon la série chronologique _____	41
Tableau XVI : Test de rationalité selon la série par meilleures saisons _____	41
Tableau XVII : Descriptifs de l'ensemble des variables utilisées _____	50
Tableau XVIII : Notation du logiciel Stata _____	52

INTRODUCTION

L'idée générale derrière ce mémoire est que l'humain est un être complexe dont les réflexions ne peuvent pas toujours être purement rationnelles. En effet, la psychologie a démontré qu'il y a plusieurs mécanismes qui opèrent au niveau cérébral pour chaque action qu'un individu pose, remettant en question l'hypothèse de rationalité parfaite des agents proposée par l'économie classique. Selon Thaler (2008),

If you look at economics textbooks, you will learn that homo economicus can think like Albert Einstein, store as much memory as IBM's Big Blue, and exercise the willpower of Mahatma Gandhi. Really. But the folks that we know are not like that. Real people have trouble with long division if they don't have a calculator, sometimes forget their spouse's birthday, and have a hangover on New Year's Day. They are not Homo economicus; they are homo sapiens.

À partir de cette réflexion, il apparaît important de comprendre à l'aide de l'économie comportementale les liens qui existent entre la psychologie et l'économie. En particulier, cette sous-discipline de l'économie générale s'intéresse aux limites de la rationalité des agents économiques qu'elle tente de comprendre avec cette perspective mixte des deux champs d'étude. Bien entendu, il existe, tout comme pour l'économie traditionnelle, une multitude de théories en économie comportementale tentant d'expliquer ces limites et défaillances du rationalisme des individus actifs économiquement. L'intérêt de ce mémoire et la fondation de son hypothèse principale porte sur l'une de ces multiples théories : l'effet d'ancrage. Celui-ci consiste principalement à comprendre comment la connaissance d'une donnée quelconque à une période temporelle antérieure peut affecter un individu dans son estimation d'une valeur à une période ultérieure. Afin d'illustrer ce type de biais, Kahneman et Tversky (1974) ont réalisé une expérience classique qui consiste à tourner, devant plusieurs groupes successifs, une roue de fortune numérotée de 1 à 100, mais truquée pour qu'elle n'arrête qu'à 10 ou 65. Une fois la roue arrêtée à l'un ou l'autre de ces nombres, chacun des individus du groupe observé se voit demandé de répondre à deux questions. La première est de dire si le pourcentage de nations africaines au sein des nations unies est supérieur ou inférieur au résultat affiché, et la deuxième est de donner leur meilleure estimation du pourcentage réel de nations africaines dans les nations unies. De façon rationnelle, le nombre affiché sur la roue ne

devrait pas influencer les estimations soumises par les participants de l'étude, car celui-ci est aléatoire et ne contient donc aucune information quant au réel pourcentage à estimer. Or, il s'avère que ce nombre a bel et bien un impact. En fait, ceux qui ont vu le « 10 » s'afficher ont soumis une estimation moyenne de 25%, tandis que ceux qui ont vu « 65 » ont plutôt soumis 45% comme estimation moyenne.¹ Alors que les résultats de Kahneman et Tversky montre qu'il est convenable de croire en l'existence de cet effet d'ancrage, ce mémoire cherche à questionner comment cela peut se refléter dans les prévisions économiques où il est coutume d'observer des séries de données passées pour estimer des valeurs pour le futur. Ainsi, ces données passées pourraient créer un effet d'ancrage au sein des prévisionnistes, biaisant ainsi leurs anticipations.

Pour explorer comment cela se traduit empiriquement, ce mémoire étudie les données des joueurs de la NHL occupant le poste d'attaquant et qui sont restés actifs au courant de la période entre 2007 et 2016 par une régression les caractéristiques des joueurs et leurs performances passées. Son intérêt, gravitant autour des trois variables que sont la production en points des huit dernières années d'un joueur, le risque associé à cette production et son rang auquel il a été sélectionné à l'année de son repêchage, repose sur l'hypothèse de base suivante :

H0 : Les directeurs généraux ont été, au cours des années 2007 à 2016, biaisés par au moins une des trois variables d'intérêt mentionnées ci-haut lors de l'octroi de contrats à des attaquants affectant irrationnellement le salaire de ceux-ci.

H1 : Les directeurs généraux sont sensibles et rationnels vis-à-vis l'information qui leur était disponible lors de l'octroi des contrats et salaires aux attaquants, qui à leur tour ont reçu un salaire juste sur la base de cette information.

(H0 est l'hypothèse de base ; H1 est l'hypothèse alternative)

Ce mémoire consistera donc, dans un premier temps, en une revue de la littérature portant sur l'économie comportementale avec une attention plus soutenue sur l'effet d'ancrage. Celle-ci étant au cœur et à l'origine de ce travail, cette partie couvrira l'état actuel des connaissances

¹ Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.

sur ce sujet et servira de base à la compréhension des origines du modèle statistique développé subséquemment. Entre autres, il sera question des développements historiques en la matière, expliquant les avancées de Slovic, Kahneman et Tversky, Chapman et Johnson, Blanckenship ainsi que de Thaler.

Dans un deuxième temps, suite à la revue de la littérature, suivra la description des données utilisées. Il sera expliqué comment la base de données a été construite, ce qu'elle contient et ses utilités pour la suite de ce travail de même que les explications nécessaires à la compréhension des données directes - provenant directement des sources de données - et indirectes - provenant de calculs dérivés des données directes.

Dans un troisième temps, suivra la partie statistique, soit le fondement des résultats obtenus. Cette partie décrira les variables utilisées, le choix de celles-ci, les modèles utilisés pour les régressions statistiques réalisées et l'argumentaire justifiant ces derniers. En effet, les résultats de ce mémoire reposent sur une série de modifications apportées, de façon incrémentale, à un modèle de base, d'où l'importance de comprendre les origines de ce dernier.

Dans un quatrième temps, en lien à la description statistique des régressions suivra l'explication des résultats obtenus. Il s'agit ici de résumer et interpréter les conclusions des régressions réalisées antérieurement afin d'en dégager un résultat général pour ce travail. Tel que mentionné précédemment, une approche incrémentale a été favorisée pour ce travail et, conséquemment, cette section portera sur les aboutissements de chacune de ces étapes.

Dans un cinquième temps, ce travail discutera de la robustesse des résultats et les facteurs pouvant affecter la fiabilité de ceux-ci. Il sera question de ce qui aurait pu être fait autrement, de ce qui semble avoir bien fonctionné et de suggestions pour un travail semblable futur.

Ensuite, une partie sera consacrée à une discussion générale sur le travail accompli. Il sera, entre autres, mention du modèle prôné et de ses possibles faiblesses ainsi que la méthodologie favorisée pour cette recherche.

Enfin, ce mémoire se terminera avec une conclusion résumant l'ensemble du travail réalisé de même que les résultats obtenus.

CHAPITRE 1 - REVUE DE LA LITTÉRATURE

1.1 – FONDEMENTS DE L'ÉCONOMIE COMPORTEMENTALE

« *Since my world picture approximates reality only crudely, I cannot aspire to optimize anything; at most, I can aim at satisficing. Searching for the best can only dissipate scarce cognitive resources; the best is the enemy of the good.* » - Herbert A. Simon

Bien que le domaine d'étude qu'est l'économie comportementale soit relativement nouveau, ses fondations remontent tout de même aux années 50 au moment où Herbert A. Simon (1957), scientifique social de renom, soumet l'idée de *bounded rationality* (rationalité limitée) dans son recueil « *Models of man; social and rational* » composé de ses précédentes recherches en la matière. Il y explique qu'il est quasi, sinon totalement, impossible que l'agent économique typique puisse être en parfaite connaissance de toute l'information pertinente, capable de calculs complexes et ayant un système de préférences clair et stable.² Bien qu'il n'ait pas parlé spécifiquement de l'économie comportementale dans cette recherche, il est possible de comprendre que son apport a servi de fondement à cette discipline. En effet, l'économie comportementale utilise des évidences psychologiques et biologiques pour comprendre les décisions économiques des agents lorsque celles-ci ne reflètent pas la théorie économique traditionnelle, où les choix de ces derniers sont issus de maximisations et optimisations rationnelles, par des explications biologiquement et psychologiquement plausibles.³

1.2 – FONDEMENTS DE L'EFFET D'ANCRAGE

L'une des voies empruntées par l'économie comportementale pour expliquer certaines décisions a priori irrationnelles consiste à comprendre comment l'information passée peut influencer les décisions présentes et futures des agents. Slovic (1967), un des pionniers en la matière, a introduit le concept d'ancrage pour comprendre l'aversion au risque d'un individu en situation de loterie en réalisant une expérience où des individus devaient tourner deux

² Simon, H. A. (1955). A behavioral model of rational choice. *The quarterly journal of economics*, p.99-118.

³ Camerer, C. F. (2014). Behavioral economics. *Current Biology*, vol.24(18), R867-R871.

roulettes. L'une d'elles avait une probabilité de gagner et un certain montant à gagner, l'autre avait une probabilité de perte et un certain montant, différent, à perdre et où pour les deux roulettes, si l'évènement de gagner ou perdre n'arrivait pas, la probabilité restante indiquait que l'individu ne recevait ou ne déboursait rien.⁴ Dans cette étude, il démontre que le déterminant le plus important d'un individu pour évaluer le risque d'une loterie est la probabilité de perte, celle-ci étant considérablement plus décisive que le montant à perdre, et que la probabilité de gagner ainsi que le montant à gagner sont relativement davantage négligés que la probabilité de perte. Par conséquent, il a été observé que l'échantillon de cette expérimentation était soumis à un effet d'ancrage et, donc, biaisé vers la probabilité de perte. Il faut toutefois noter que ce n'est pas Slovic qui a formellement introduit la notion spécifique d'« effet d'ancrage »; son travail a plutôt offert l'un des premiers résultats empiriques évidents que représente ce type de biais.

1.3 – ANCRAGE ET AJUSTEMENT

En effet, c'est Kahneman et Tversky (1974) qui, tel que présenté dans l'introduction de ce mémoire, introduisent non seulement l'effet d'ancrage comme tel mais aussi les mécanismes de son fonctionnement psychologique. Ils étendent ainsi l'empreinte de leur travaux jusqu'à comprendre comment l'effet d'ancrage biaise un individu vers une valeur donnée en introduisant la théorie « Anchoring and adjustment » (« ancrage et ajustement »). L'idée est qu'à partir d'un ancrage soumis, l'individu sera incapable de se distancier suffisamment de celui-ci lors de son estimation d'une valeur, et c'est à cet écart que réfère la notion de biais. Plus précisément, lorsqu'il lui est demandé de proposer une estimation à la suite de la soumission d'un ancrage, l'agent va graduellement s'éloigner de cette valeur initiale et ce même si elle n'a aucun lien avec la question posée telle que présenté en exemple dans l'introduction de ce travail. Ainsi, en s'éloignant de plus en plus de la valeur de départ, l'individu finira par atteindre une zone où il se retrouvera dans l'incertitude à savoir s'il doit encore, ou non, diverger davantage de l'ancrage. C'est à cet instant, dès qu'il atteint cette

⁴ Slovic, P. (1967). The relative influence of probabilities and payoffs upon perceived risk of a gamble. *Psychonomic Science*, vol.9(4), p.223-224.

zone, que l'individu soumettra son approximation. Par conséquent, le biais réside dans ce phénomène où l'agent typique demandé d'estimer une donnée, sous l'influence d'un ancrage, ne donnera pas sa meilleure estimation mais plutôt la valeur la plus éloignée de l'ancrage sans atteindre la zone d'incertitude.⁵ La raison pour laquelle une personne ne cherche pas à aller plus loin dans sa déviation de l'ancre initiale a été soumise par Kahneman et Tversky (1972) antérieurement à ces recherches sur l'effet d'ancrage et les autres types de biais psychologiques. L'idée proposée est que les mécanismes psychologiques derrière chaque action et réflexion peuvent se diviser en deux systèmes, numérotés 1 et 2. Le premier est le mode automatique, non réfléchi, intuitif, rapide ainsi que moins ardu psychologiquement et énergétiquement, derrière les actions communes et répétitives. Le deuxième, à l'inverse, est celui en charge de la réflexion avancée, des calculs exigeants, de l'analyse détaillée et il est aussi plus lent en plus de demander un effort plus grand. Pour ces raisons, un individu va toujours, autant que possible, essayer d'opter pour agir en fonction du Système 1 afin de minimiser son effort, celui-ci offrant souvent de bonnes estimations, et ne va ainsi pas se questionner davantage à savoir s'il devrait se distancer encore plus de l'ancre initial lorsqu'une estimation lui est requise.⁶ Cette façon de voir le biais d'ancrage est la plus répandue en la matière, et est à la base des recherches effectuées subséquemment.

1.4 – ACCESSIBILITÉ SÉLECTIVE

En 1999, Chapman et Johnson ont soumis une autre façon d'aborder l'effet d'ancrage. Celle-ci, dénommée « Selective accessibility » (« Accessibilité sélective »), consiste pour un individu à confirmer si l'ancrage est valide. Le biais provient ici du fait que l'agent va tendre à sélectionner des caractéristiques validant l'ancre initial au moment de faire un choix ou d'estimer une valeur quelconque. Un exemple populaire cité, dans leur travail, décrit un agent immobilier qui va afficher un prix très élevé pour une maison afin d'influencer les acheteurs potentiels à accepter de payer plus cher pour la propriété, ou un prix très bas et tenter de

⁵ Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.

⁶ Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.

recevoir une évaluation foncière plus basse.⁷ Il est possible de constater que cette approche est substantiellement différente de celle de Kahneman et Tversky, entre autres sur le fait que, contrairement à ces derniers, Chapman et Johnson posent la possibilité de contrer l'effet d'ancrage. En effet, ils proposent que, dans certains cas, ce type de biais peut être évité en questionnant l'estimation sur la base d'information intrinsèque. Par exemple, l'acheteur potentiel d'une maison pourrait se demander comment vaut une maison semblable à celle surévaluée par l'agent immobilier sur la base du nombre de pièces, du quartier, de son âge, etc., et en arriver à un prix juste en fonction de ces caractéristiques.

1.5 – CHANGEMENT D'ATTITUDE

Une autre approche à l'effet d'ancrage, « Attitude change » (« Changement d'attitude »), a été proposée par Wegener et al. en 2001 et 2010, puis élaborée davantage par Blankenship et al. en 2008. Tout d'abord, Blankenship démontre que, selon le niveau d'effort mental disponible pour la tâche, l'ancrage initial envoie plutôt un signal vers la vraie valeur à estimer ou dirige la collecte d'information de l'agent vers l'ancrage lui-même.⁸ Par exemple : un individu, cherchant à acheter une télévision dans un magasin, se voit offrir trois modèles, 1, 2 et 3, de deux compagnies différentes, A et B. La compagnie A produit les modèles 1 et 2, tandis que la compagnie B produit le modèle 3. Le modèle 2 est la version « de luxe » du modèle 1, possédant de meilleures spécifications techniques. Ainsi, selon la théorie du changement d'attitude, l'individu sera biaisé vers le modèle 2, sans considération budgétaire, car il sait que cette télévision est meilleure que le modèle 1, mais surtout qu'il n'a aucune raison de croire en une meilleure offre de la part de la compagnie B. Dans cet exemple, les modèles 1 et 3 servent d'ancrages, déviant la collecte d'information de l'agent afin qu'il choisisse le modèle 2, d'une part parce que le modèle 1 est connu pour être moins bon et, d'autre part, le modèle 3 n'est pas

⁷ Chapman, G. B., & Johnson, E. J. (1999). Anchoring, activation, and the construction of values. *Organizational behavior and human decision processes*, vol.79(2), p.115-153.

⁸ Blankenship, K. L., Wegener, D. T., Petty, R. E., Detweiler-Bedell, B., & Macy, C. L. (2008). Elaboration and consequences of anchored estimates: An attitudinal perspective on numerical anchoring. *Journal of Experimental Social Psychology*, vol.44(6), p.1465-1476.

connu pour être meilleur. Aussi, si l'individu a déjà une préférence pour la compagnie A, même si cette dernière n'offrait qu'un modèle, il pourrait tout de même sélectionner celle-ci s'il a des raisons de croire que celle de l'entreprise B ne peut pas être meilleure. Ensuite, Wegener s'est quant à lui davantage intéressé aux limites de l'ancrage comme tel pour en venir à proposer l'idée du changement d'attitude. Dans son travail, il démontre que si un individu n'a pas de raison pour croire en l'ancre - étant trop extrême, absurde ou parce que l'individu a une connaissance lui permettant de la juger impertinente, par exemple - celle-ci ne soumettra peu voire aucun biais à l'individu en question. Par conséquent, arrivant à des résultats contradictoires avec les recherches précédentes de Kahneman et Chapman, Wegener arrive à cette théorie de sélection adverse de l'information où l'agent va plutôt essayer de trouver des raisons pour invalider l'ancre initiale, en opposition à tenter de le valider. Toutefois, si cette valeur initiale est proposée par une source fiable, réputée ou pour quelconques raisons difficilement réfutables, cette recherche d'informations adverses sera d'autant plus ardue et, conséquemment, l'ancrage sera plus fort.⁹

1.6 – FACTEURS AFFECTANT L'EFFET D'ANCRAGE

Bien entendu, plusieurs facteurs peuvent jouer un rôle quant à l'importance, voire la présence ou non, de l'effet d'ancrage. L'humeur (Bodenhausen, G. V., Gabriel, S., & Lineberger, M. 2000)¹⁰, l'expérience, le niveau de connaissance et d'expertise (Wilson, T. D., Houston, C. E., Etling, K. M., & Brekke, N. 1996)¹¹, la motivation, les incitations et les avertissements (Epley, N., & Gilovich, T. 2005)¹², la personnalité (McElroy, T., & Dowd, K. 2007)¹³ et les habilités

⁹ Wegener, D. T., Petty, R. E., Detweiler-Bedell, B. T., & Jarvis, W. B. G. (2001). Implications of attitude change theories for numerical anchoring: Anchor plausibility and the limits of anchor effectiveness. *Journal of Experimental Social Psychology*, vol.37(1), p.62-69.

¹⁰ Bodenhausen, G. V., Gabriel, S., & Lineberger, M. (2000). Sadness and susceptibility to judgmental bias: The case of anchoring. *Psychological Science*, vol.11(4), p.320-323.

¹¹ Wilson, T. D., Houston, C. E., Etling, K. M., & Brekke, N. (1996). A new look at anchoring effects: basic anchoring and its antecedents. *Journal of Experimental Psychology: General*, vol.125(4), p.387-402

¹² Epley, N., & Gilovich, T. (2005). When effortful thinking influences judgmental anchoring: differential effects of forewarning and incentives on self-generated and externally provided anchors. *Journal of Behavioral Decision Making*, vol.18(3), p.199-212.

cognitives (Bergman, O., Ellingsen, T., Johannesson, M., & Svensson, C. 2010)¹⁴ ont été explorés par différents chercheurs pour montrer que ceux-ci peuvent faire varier le niveau du biais chez un individu. Par contre, si certaines recherches ont démontré des résultats significatifs, d'autres n'y sont pas parvenues. Il n'est donc pas totalement évident si ces variables sont pertinentes quant à la magnitude de l'effet d'ancrage.

1.7 – PERTINENCE POUR CETTE RECHERCHE

Malgré les multiples recherches et travaux effectués dans ce domaine, il reste que la définition la plus largement acceptée de l'effet d'ancrage reste celle de Kahneman et Tversky, posant celle-ci comme un processus d'ancrage et d'ajustement. De ce fait, Thaler (2008) résume le phénomène comme étant un agent qui part d'une valeur initiale, puis l'ajuste dans la direction qui lui semble la plus appropriée. De cette façon, l'effet d'ancrage résulte de cet ajustement généralement insuffisant.¹⁵ De plus, il faut comprendre qu'il est empiriquement impossible de distinguer, entre elles, les différentes définitions de l'effet d'ancrage. En effet, chacune offre une compréhension des mécanismes engendrant le biais et non une définition distincte pour ce dernier. Pour ces raisons, c'est à partir de cette définition de ce type de biais que ce mémoire s'est construit. Il s'agit d'une démonstration empirique de l'effet d'ancrage auquel pourraient être soumis les directeurs généraux des équipes de la Ligue Nationale de Hockey (NHL) lorsqu'ils attribuent des salaires, par l'entremise de contrats, aux joueurs occupant le poste d'attaquant entre 2007 et 2016 sur la base de leurs performances passées et de leurs caractéristiques.

¹³ McElroy, T., & Dowd, K. (2007). Susceptibility to anchoring effects: How openness-to-experience influences responses to anchoring cues. *Judgment and Decision Making*, vol.2(1), p.48-53

¹⁴ Bergman, O., Ellingsen, T., Johannesson, M., & Svensson, C. (2010). Anchoring and cognitive ability. *Economics Letters*, vol.107(1), p.66-68.

¹⁵ Thaler, R. H. & Sunstein, C. R. (2008). *Nudge*. Penguin group.

CHAPITRE 2 - DESCRIPTION DES DONNÉES

Les données utilisées dans ce travail^{16,17,18} regroupent l'ensemble des statistiques des attaquants de la NHL actifs tout au long de la période s'écoulant de la saison 2007-2008 à la saison 2015-2016, à l'exception de la saison 2012-2013, car celle-ci ne contient que 48 parties en raison d'une grève des joueurs pour les négociations de leur nouvelle convention collective.¹⁹ La raison derrière le choix de cette période de temps était d'obtenir un échantillon d'environ 150 joueurs, le nombre formel d'observations obtenues est 151, ayant joué à chacune de ces saisons pour permettre une comparaison statistique entre eux. Un élément important à relever est qu'un joueur fait partie de l'échantillon uniquement s'il a joué à toutes les saisons retenues, incluant celle réduite par les événements entourant la négociation de la convention collective à la saison 2012-2013, ainsi que la saison 2015-2016.

Les salaires utilisés pour établir des relations avec les performances passées sont ceux de 2016 seulement afin d'isoler les effets qui ne devraient pas être significatifs à long terme, tel que le rang au repêchage, et pour s'assurer de ne pas considérer des contrats pour recrues ou transitoires. De plus, ceux-ci sont représentés sous la forme de leur logarithme naturel (ln), afin d'éviter l'obtention de salaires négatifs et de se concentrer sur les variations des salaires entraînées par les variables indépendantes.

Pour les performances, il a semblé préférable de garder les données en niveau, plutôt que de les transformer en données par partie ou par minute, afin de capturer le plus possible les effets des différentes variables. En particulier, dans l'absence de variables sur le nombre de parties manquées par un joueur en raison de blessures ou de suspensions, il semble normal d'éviter un tel remaniement par match. De la même façon, la transformation des données par minute empêche aussi la capture d'une telle information en plus d'être influencée par le temps passé en pénalité et sur la glace. En plus, cela peut aussi biaiser les données, par exemple dans le cas d'un joueur de quatrième trio qui passe peu de temps sur la glace mais qui réussit à faire

¹⁶ www.nhl.com

¹⁷ www.hockeyreference.com

¹⁸ www.nhlnumbers.com

¹⁹ <http://globalnews.ca/news/287019/timeline-the-2012-2013-nhl-lockout/>

quelques buts et assistances. Ce joueur pourrait paraître avoir des statistiques d'un joueur de premier trio, à l'extrême, mais sans un tel salaire, car en pratique il lui est impossible de performer autant et avoir le même nombre de points dans une saison avec un temps de jeu semblable.

Aussi, il faut comprendre que dans la NHL, un point est inscrit par un joueur s'il marque un but ou obtient une assistance. Cette dernière est obtenue par le ou les dernier(s) joueur(s), excluant celui qui a marqué le but, ayant touché la rondelle, pour un maximum de deux joueurs.²⁰

À noter, les performances sont divisées en deux grandes catégories : les performances chronologiques et les performances ordonnées selon le nombre de points pour chaque saison, de la meilleure à la pire. L'idée est d'en arriver à un comparatif en ces deux séries de données pour en dégager de possibles biais si un coefficient, dans la régression linéaire, ne réagit pas conformément aux attentes de la théorie économique. Ainsi, il y a l'ensemble des statistiques, à l'exception des salaires tel que mentionné précédemment, pour chaque saison de 2007-2008 à 2015-2016, sauf celle de 2012-2013, pour chaque attaquant actif durant cette période. Ensuite, il y a ces mêmes données pour chaque joueur, mais celles-ci ont toutes été ordonnées par ordre de meilleure saison en termes de points, et par conséquent, bien qu'il s'agisse des mêmes données, cela résulte en un réarrangement complet pour chaque joueur de leurs performances obtenues.

Enfin, la base de données contient aussi d'autres informations, plus générales, sur chaque joueur, telles que l'équipe pour laquelle il joue présentement, à la saison 2015-2016, si celle-ci est canadienne ou américaine, sa position à l'attaque, son âge et son rang au repêchage. Cette dernière variable est d'autant plus importante, car une partie considérable de ce mémoire consistera à analyser son effet sur le présent impact sur le cap salarial pour une équipe des joueurs de la NHL. D'ailleurs, il faut aussi noter que le rang 288 a été attribué aux joueurs qui n'ont pas été repêchés, mais qui ont tout de même évolué dans la NHL au courant des saisons observées.

²⁰ National Hockey League (2015), *Official rules 2015-2016*, article no 78, Repéré à <http://www.nhl.com/nhl/en/v3/ext/rules/2015-2016-Interactive-rulebook.pdf>

L'annexe A contient l'ensemble des variables utilisées, leur signification et leur abréviations utilisées lors de l'analyse statistique.

Le Tableau I contient les principales variables utilisées dans les régressions pour arriver aux résultats obtenus dans ce travail. Ainsi, seuls les salaires (*SCH16*) ont été récupérés du site www.nhlnumbers.com alors que l'ensemble des autres données proviennent du www.hockeyreference.com. Le site www.nhl.com a servi de base pour comparer les données afin d'en établir la fiabilité.

Tableau I : Notation principale

Notation	Description de la variable
SCH16	Impact du salaire sur le cap salarial de son équipe
InSCH16	Logarithme naturel de SCH16
Age	Âge d'un joueur à la saison 2015-2016
Draft	Rang auquel un joueur a été repêché
CL	Durée du contrat signé par un joueur
stdPTS	Écart-type des points marqués par un joueur au cours des saisons 2007-2008 à 2014-2015
_cons	Constante incluse dans les régressions
PTS15	Nombre de points marqués par un joueur au cours de la saison 2014-2015
PTS14	Nombre de points marqués par un joueur au cours de la saison 2013-2014
PTS12	Nombre de points marqués par un joueur au cours de la saison 2011-2012
PTS11	Nombre de points marqués par un joueur au cours de la saison 2010-2011
PTS10	Nombre de points marqués par un joueur au cours de la saison 2009-2010
PTS09	Nombre de points marqués par un joueur au cours de la saison 2008-2009
PTS08	Nombre de points marqués par un joueur au cours de la saison 2007-2008
BS_PTS1	Meilleure saison en termes de points marqués d'un joueur
BS_PTS2	Deuxième meilleure saison en termes de points marqués d'un joueur
BS_PTS3	Troisième meilleure saison en termes de points marqués d'un joueur
BS_PTS4	Quatrième meilleure saison en termes de points marqués d'un joueur
BS_PTS5	Cinquième meilleure saison en termes de points marqués d'un joueur
BS_PTS6	Sixième meilleure saison en termes de points marqués d'un joueur
BS_PTS7	Septième meilleure saison en termes de points marqués d'un joueur

CHAPITRE 3 - STATISTIQUE DESCRIPTIVE

3.1 – SOMMAIRE DES VARIABLES PRINCIPALES

Le Tableau II représente les principales caractéristiques des variables utilisées dans les régressions linéaires.

Tableau II : Variables des régressions

Variable	Mean	Std. Dev.	Min	Max	Mdn
Age	31.69536	2.993311	26	39	4
Draft	43.23841	78.45744	1	288	31
CL	4.18543	2.784251	1	13	17
stdPTS	13.12202	5.737203	1.718249	34.16068	4
SCH16	4.021344	2.403316	0.026	10.5	12.75
PTS15	36.25828	20.24433	0	84	34
PTS14	39.07947	21.45958	1	104	39
PTS12	42.56291	21.9486	0	109	44
PTS11	44.03974	21.60336	2	104	45
PTS10	43.38411	24.58126	2	112	41
PTS09	42.70861	26.62895	0	113	41
PTS08	36.98013	27.34458	0	112	32
BS_PTS1	58.21192	24.02793	8	113	59
BS_PTS2	51.4702	22.71587	7	110	52
BS_PTS3	46.12583	21.36362	5	109	46
BS_PTS4	40.82781	19.9318	4	85	41
BS_PTS5	36.61589	19.33696	4	81	36
BS_PTS6	30.4702	19.01326	0	79	28
BS_PTS7	21.29139	17.41822	0	65	16

Dans ce tableau, la colonne « Variable » contient le nom abrégé de chaque variable, la colonne « Mean » désigne leur moyenne respective, « Std. Dev. » donne leur écart-type, « Min » ainsi que « Max » représentent, respectivement, le minimum et le maximum de chaque variable et « Mdn » montre leur médiane.

De cette façon, il est possible de constater que la moyenne d'âge des joueurs ayant évolué dans la ligue nationale depuis 2007 est d'environ 32 ans, le plus vieux ayant 39 ans alors que le plus jeune a 26 ans. Le joueur moyen a été repêché approximativement au 43^e rang, a signé un contrat d'environ quatre ans pour un salaire moyen de quatre millions de dollars (US) par année. De plus, la durée des contrats varie entre une et treize années, et l'écart-type des points pour chaque joueur va de 1.72 à 34.16 pour les sept saisons. En regroupant les meilleures saisons pour chacun des 151 joueurs observés, la moyenne des points marqués pour ces saisons est de plus ou moins 58 points, dont un minimum de sept points et un maximum de 113 points, alors que celle de la pire saison est de 21 points, avec un minimum de zéro et un maximum de 65 points. Le joueur le moins bien payé durant la période de temps observée a reçu un salaire de 0.026 M\$ (26 000\$) pour la saison 2015-2016, tandis que le joueur le mieux payé a reçu 10.5M\$. Parallèlement, la moyenne des points inscrits en 2015 est d'environ 36, tous se situant dans l'intervalle entre zéro et 84 inclusivement.

En général, la série chronologique des points inscrits par saison varie peu, sa moyenne pour les sept années de l'échantillon se situant à 40.72. Par contre, prise individuellement, chaque saison connaît un écart-type très élevé au niveau des points inscrits par les attaquants, ce qui apparaît normal étant donné les différents niveaux de performances et aptitudes de ces derniers.

3.2 – MODÈLE

Tel que mentionné précédemment, les données sont divisées en deux catégories, soient chronologiquement ou par meilleures saisons en terme de points. Conséquemment, les régressions sont aussi divisées de la sorte, mais conserve tout de même le même format. En effet, il s’agit de régressions linéaires en paramètres prenant la forme générale suivante :

$$Y = \beta_i X_i + \varepsilon_i$$

- Où
- Y représente le vecteur de la variable dépendante, soit le logarithme des salaires
 - β_i représente le vecteur des coefficients* pour chaque variable dans la régression i
 - X_i représente le vecteur des variables indépendantes, soit les performances et caractéristique de chaque joueur dans la régression
 - ε_i représente le terme d’erreur dans la régression
- $i = \begin{cases} 0 & \text{pour la régression par ordre chronologique} \\ 1 & \text{pour la régression par ordre de performances} \end{cases}$

*À noter, le vecteur β_i contient aussi le terme constant α_i

La présentation des résultats dans ce mémoire se fera de façon incrémentale. Il y aura plusieurs étapes où, à chacune d’entre elles, une modification au modèle de base sera apportée afin de pouvoir cerner les phénomènes observés dans les régressions. Ces changements dans le modèle prendront principalement la forme de l’ajout ou le retrait de variables indépendantes (du vecteur X_i), et viseront à tester des hypothèses et des résultats obtenus. Chaque étape comprendra une justification de la démarche entreprise, un tableau résumant les résultats obtenus, un sommaire et une interprétation de ceux-ci, ainsi qu’une description des hypothèses derrière le modèle suggéré.

Voici, dans l'ordre, les régressions réalisées et présentées subséquemment dans la section « Résultats » :

- 1- a) $\lnSCH16 = \beta_{15} PTS15 + \beta_{14} PTS14 + \beta_{12} PTS12 + \beta_{11} PTS11 + \beta_{10} PTS10 + \beta_{09} PTS09 + \beta_{08} PTS08 + \alpha_1 + \varepsilon_1$
- b) $\lnSCH16 = \beta_1 BS_PTS1 + \beta_2 BS_PTS2 + \beta_3 BS_PTS3 + \beta_4 BS_PTS4 + \beta_5 BS_PTS5 + \beta_6 BS_PTS6 + \beta_7 BS_PTS7 + \alpha_1 + \varepsilon_1$
- 2- a) $\lnSCH16 = \beta_D Draft + \beta_{15} PTS15 + \beta_{14} PTS14 + \beta_{12} PTS12 + \beta_{11} PTS11 + \beta_{10} PTS10 + \beta_{09} PTS09 + \beta_{08} PTS08 + \alpha_0 + \varepsilon_0$
- b) $\lnSCH16 = \beta_D Draft + \beta_1 BS_PTS1 + \beta_2 BS_PTS2 + \beta_3 BS_PTS3 + \beta_4 BS_PTS4 + \beta_5 BS_PTS5 + \beta_6 BS_PTS6 + \beta_7 BS_PTS7 + \alpha_1 + \varepsilon_1$
- 3- a) $\lnSCH16 = \beta_D Draft + \beta_{CL} CL + \beta_{15} PTS15 + \beta_{14} PTS14 + \beta_{12} PTS12 + \beta_{11} PTS11 + \beta_{10} PTS10 + \beta_{09} PTS09 + \beta_{08} PTS08 + \alpha_0 + \varepsilon_0$
- b) $\lnSCH16 = \beta_D Draft + \beta_{CL} CL + \beta_1 BS_PTS1 + \beta_2 BS_PTS2 + \beta_3 BS_PTS3 + \beta_4 BS_PTS4 + \beta_5 BS_PTS5 + \beta_6 BS_PTS6 + \beta_7 BS_PTS7 + \alpha_1 + \varepsilon_1$
- 4- a) $\lnSCH16 = \beta_D Draft + \beta_{std} stdPTS + \beta_{CL} CL + \beta_{15} PTS15 + \beta_{14} PTS14 + \beta_{12} PTS12 + \beta_{11} PTS11 + \beta_{10} PTS10 + \beta_{09} PTS09 + \beta_{08} PTS08 + \alpha_0 + \varepsilon_0$
- b) $\lnSCH16 = \beta_D Draft + \beta_{std} stdPTS + \beta_{CL} CL + \beta_1 BS_PTS1 + \beta_2 BS_PTS2 + \beta_3 BS_PTS3 + \beta_4 BS_PTS4 + \beta_5 BS_PTS5 + \beta_6 BS_PTS6 + \beta_7 BS_PTS7 + \alpha_1 + \varepsilon_1$
- 5- a) $\lnSCH16 = \beta_D Draft + \beta_A Age + \beta_{std} stdPTS + \beta_{CL} CL + \beta_{15} PTS15 + \beta_{14} PTS14 + \beta_{12} PTS12 + \beta_{11} PTS11 + \beta_{10} PTS10 + \beta_{09} PTS09 + \beta_{08} PTS08$
- b) $\lnSCH16 = \beta_D Draft + \beta_A Age + \beta_{std} stdPTS + \beta_{CL} CL + \beta_1 BS_PTS1 + \beta_2 BS_PTS2 + \beta_3 BS_PTS3 + \beta_4 BS_PTS4 + \beta_5 BS_PTS5 + \beta_6 BS_PTS6 + \beta_7 BS_PTS7$
- 6- a) $\lnSCH16 = \beta_D Draft + \beta_A Age + \beta_{CL} CL + \beta_{15} PTS15 + \beta_{14} PTS14 + \beta_{12} PTS12 + \beta_{11} PTS11 + \beta_{10} PTS10 + \beta_{09} PTS09 + \beta_{08} PTS08$
- b) $\lnSCH16 = \beta_D Draft + \beta_A Age + CL + \beta_1 BS_PTS1 + \beta_2 BS_PTS2 + \beta_3 BS_PTS3 + \beta_4 BS_PTS4 + \beta_5 BS_PTS5 + \beta_6 BS_PTS6 + \beta_7 BS_PTS7$

CHAPITRE 4 - RÉSULTATS

4.1 – MODÈLE DE BASE

Tout d’abord, il faut noter que les régressions ont été réalisées à l’aide du logiciel Stata. Par conséquent, la notation utilisée au sein des tableaux présentant les résultats sera celle du logiciel, et la description de celle-ci se retrouve à l’Annexe B. Chaque tableau des résultats des régressions sera accompagné par un sommaire des statistiques générales de la régression, qui comprendra le nombre d’observations, la statistique de Fisher, la *p-value* de la statistique de Fisher, le coefficient de corrélation (R^2) et la racine carrée de la MSE (moyenne des écarts au carré).

De plus, il est important de mentionner que toutes les régressions ont été réalisées avec des écart-types robustes, plus précisément ceux de *Huber-White*, afin de corriger partiellement pour la colinéarité entre les variables et d’inclure l’hétéroscédasticité de celles-ci, en plus de considérer une certaine non-normalité des données. En effet, certaines variables peuvent avoir un influence directe sur d’autres, tel que le rang auquel un joueur a été repêché et son nombre de points inscrits. En effet, il est très probable qu’un joueur repêché tôt marque plus de points qu’un joueur repêché tard, voire jamais, car la raison même pour laquelle il a été sélectionné plus haut est que les recruteurs dans la ligue voyaient en lui et ses performances antérieures un plus grand potentiel de production en points.

Ainsi, la première paire de régressions représente le modèle de base, c’est-à-dire sans l’ajout de variables au sein des régressions chronologique et par performance. Elle vise à tester comment les directeurs généraux de la Ligue se comportent vis-à-vis les performances uniquement, afin de voir d’entrée de jeu s’il y a une anomalie dans cette série. Elle se traduit en les deux régressions suivantes, qui réfèrent au modèle fondamental de ce travail :

$$\begin{aligned} \text{a) } \ln\text{SCH16} &= \beta_{15} \text{PTS15} + \beta_{14} \text{PTS14} + \beta_{12} \text{PTS12} + \beta_{11} \text{PTS11} + \beta_{10} \text{PTS10} + \beta_{09} \text{PTS09} \\ &+ \beta_{08} \text{PTS08} + \alpha_0 + \varepsilon_0 \end{aligned}$$

$$\begin{aligned} \text{b) } \ln\text{SCH16} &= \beta_1 \text{BS_PTS1} + \beta_2 \text{BS_PTS2} + \beta_3 \text{BS_PTS3} + \beta_4 \text{BS_PTS4} + \beta_5 \text{BS_PTS5} + \beta_6 \text{BS_PTS6} \\ &+ \beta_7 \text{BS_PTS7} + \alpha_1 + \varepsilon_1 \end{aligned}$$

L'idée est d'évaluer l'effet de la réorganisation de la série chronologique en celle par ordre de la meilleure à la pire performance pour chaque joueur. Entre autres, la théorie sur l'effet d'ancrage pourrait soumettre l'hypothèse que le coefficient de la meilleure saison, dans la deuxième régression, soit significatif et plus élevé que les autres, suggérant que les directeurs généraux des équipes surpayaient les joueurs selon leur meilleure saison. Il se pourrait aussi que les responsables des équipes accordent une importance anormalement grande aux récentes saisons, suggérant un biais causé par la proximité temporelle de l'information. De plus, tout au long de cette section, il est mentionné de la raison et de la sensibilité des directeurs généraux. Ce sujet est traité en détail à la section « Robustesse des résultats », où il y est expliqué empiriquement par un test de rationalité. Ainsi, le Tableau III montre les résultats obtenus par la régression de la série chronologique, et le Tableau IV ceux de la série par meilleures performances, tous les deux accompagnés par leur sommaire statistique respectif.

Tableau III : Régression chronologique du modèle de base

InSCH16	Coef.	Std. Err.	T	P>t	[95% Conf. Interval]	
PTS15	0.0087294	0.0042204	2.07	0.04	0.000387	0.0170718
PTS14	0.0097446	0.0038078	2.56	0.012	0.0022177	0.0172715
PTS12	0.0139227	0.0035445	3.93	0	0.0069163	0.020929
PTS11	0.0028905	0.004323	0.67	0.505	-0.0056548	0.0114358
PTS10	0.0016229	0.0028408	0.57	0.569	-0.0039925	0.0072383
PTS09	0.0000932	0.0054074	0.02	0.986	-0.0105957	0.010782
PTS08	-0.0018875	0.0040438	-0.47	0.641	-0.0098808	0.0061059
_cons	-0.3099713	0.1374564	-2.26	0.026	-0.5816803	-0.0382623

Number of obs 151
F(7, 143) 32.36
Prob > F 0
R-squared 0.5265
Root MSE 0.65043

Tableau IV : Régression par performances du modèle de base

InSCH16	Coef.	Std. Err.	t	P>t	[95% Conf. Interval]	
BS_PTS1	-0.0112095	0.0138319	-0.81	0.419	-0.0385509	0.0161319
BS_PTS2	0.0086371	0.0104147	0.83	0.408	-0.0119494	0.0292237
BS_PTS3	-0.0154865	0.0165963	-0.93	0.352	-0.0482923	0.0173194
BS_PTS4	0.0491623	0.0305649	1.61	0.11	-0.0112551	0.1095796
BS_PTS5	0.0079328	0.0142575	0.56	0.579	-0.0202499	0.0361155
BS_PTS6	-0.0011138	0.0085389	-0.13	0.896	-0.0179926	0.015765
BS_PTS7	-0.0061239	0.0037495	-1.63	0.105	-0.0135355	0.0012877
_cons	-0.0992043	0.1635852	-0.61	0.545	-0.4225618	0.2241532

Number of obs 151
F(7, 143) 31.28
Prob > F 0
R-squared 0.501
Root MSE 0.66771

À partir de la première régression (Tableau 3.1a), les trois dernières saisons semblent être les plus importantes aux yeux des directeurs généraux. Ceci porte à croire que ces derniers sont raisonnables, car il semble juste d'affirmer que cet échantillon temporel permet de cerner de manière sensée le niveau de performance d'un joueur ainsi que son potentiel à court terme. En effet, si le nombre de points marqués diminue pour chaque saison jusqu'à la plus récente, il semble correct de croire que ce joueur est sur la pente descendante et que sa production va continuer de décroître, et l'inverse est aussi vrai. Aussi, il est possible que les années 2007 à 2010 correspondaient aux saisons recrues d'un joueur, et que celui-ci bénéficiait alors d'un temps de jeu réduit étant donné son manque d'expérience chez les professionnels, ce qui pouvait avoir un impact sur ses performances car moins de temps sur la glace entraîne probablement moins de possibilités de marquer un points. Il est à noter, aussi, que le coefficient de la constante est négatif et relativement élevé. Ceci suggère qu'un certain nombre de points minimal est nécessaire pour un joueur afin d'éviter une réduction de son salaire.

Puis, en regardant la seconde régression, la quatrième saison la plus productive de même que la pire saison sont les deux seules dont les directeurs semblent se préoccuper, avec une emphase marquée sur la quatrième. Ceci semble aussi normal et raisonnable puisque cette dernière correspond aussi à la médiane de la série, soit un estimateur dont il est possible de défendre la justesse étant donné qu'elle peut être perçue comme la production espérée d'un joueur pour la prochaine saison. Aussi, la considération de la pire saison semble être justifiable, car s'il s'agit de la saison recrue où un joueur ayant marqué plus de points peut être identifié à un grand potentiel; sinon cela peut être perçu comme une production minimale. Néanmoins, l'utilisation de cette dernière variable semble plus appropriée à titre de complément d'information, d'où son importance moindre au niveau de son coefficient.

4.2 – PREMIÈRE VERSION DU MODÈLE

À la suite de ces résultats, il semble impossible d'exclure un manque de variables dans le modèle. Ainsi, d'après les conclusions de la deuxième régression, il paraît approprié d'inclure la variable *Draft*, référant à la position auquel un joueur a été repêché, d'autant plus qu'il s'agit d'un point d'intérêt important pour ce mémoire. En effet, cette variable peut être perçue comme un indicateur du potentiel d'un joueur pouvant aider à la compréhension de ces saisons recrues où celui-ci devrait typiquement moins bien performer.

De plus, l'une des hypothèses principales de cette recherche est qu'un joueur repêché tôt, principalement en première ronde, est plus susceptible de recevoir une prime pour sa position au repêchage en comparaison à un joueur avec des performances similaires repêché plus tard, sur la simple base que ce premier a un potentiel plus grand. La confirmation de cette hypothèse laisserait croire à la présence d'un effet d'ancrage au sein des directeurs généraux lorsque leur vient le temps d'octroyer un contrat à un attaquant de la Ligue Nationale de Hockey. Advenant la concrétisation de ce cas, il faudrait s'attendre à un coefficient négatif pour cette variable, car un joueur repêché plus tard devrait recevoir un moins grand salaire dû à son potentiel relativement moindre.

Les deux régressions en lien avec l'ajout de cette variable sont :

$$\text{a) } \ln\text{SCH16} = \text{Draft} + \beta_{15} \text{ PTS15} + \beta_{14} \text{ PTS14} + \beta_{12} \text{ PTS12} + \beta_{11} \text{ PTS11} + \beta_{10} \text{ PTS10} \\ + \beta_{09} \text{ PTS09} + \beta_{08} \text{ PTS08} + \alpha_0 + \varepsilon_0$$

$$\text{b) } \ln\text{SCH16} = \text{Draft} + \beta_1 \text{ BS_PTS1} + \beta_2 \text{ BS_PTS2} + \beta_3 \text{ BS_PTS3} + \beta_4 \text{ BS_PTS4} + \beta_5 \text{ BS_PTS5} \\ + \beta_6 \text{ BS_PTS6} + \beta_7 \text{ BS_PTS7} + \alpha_1 + \varepsilon_1$$

En théorie, la variable *Draft* ne devrait pas être significative, car après huit ans les joueurs ont pu faire leurs preuves et les directeurs généraux possèdent donc un bon échantillon de performances pour mesurer la contribution possible d'un joueur à leur équipe. De cette façon, les résultats sont présentés dans les tableaux V et VI, accompagnés de leur sommaire statistique.

Tableau V : Régression chronologique du premier modèle

InSCH16	Coef.	Std. Err.	t	P>t	[95% Conf. Interval]	
Draft	0.0008709	0.0005553	1.57	0.119	-0.0002268	0.0019687
PTS15	0.0091651	0.0042966	2.13	0.035	0.0006716	0.0176586
PTS14	0.0101621	0.0038259	2.66	0.009	0.002599	0.0177252
PTS12	0.013371	0.0034389	3.89	0	0.0065729	0.0201692
PTS11	0.0030895	0.0042843	0.72	0.472	-0.0053797	0.0115587
PTS10	0.0016298	0.0028279	0.58	0.565	-0.0039605	0.00722
PTS09	0.0002896	0.005392	0.05	0.957	-0.0103695	0.0109486
PTS08	-0.0016449	0.0040589	-0.41	0.686	-0.0096685	0.0063788
_cons	-0.3826861	0.1589276	-2.41	0.017	-0.6968559	-0.0685163

Number of obs 151
F(8, 142) 29.42
Prob > F 0
R-squared 0.5314
Root MSE 0.6493

Tableau VI : Régression par performances du premier modèle

InSCH16	Coef.	Std. Err.	t	P>t	[95% Conf. Interval]	
Draft	0.0014242	0.0007041	2.02	0.045	0.0000324	0.002816
BS_PTS1	-0.0136013	0.0140841	-0.97	0.336	-0.0414428	0.0142402
BS_PTS2	0.0108967	0.0104891	1.04	0.301	-0.0098384	0.0316317
BS_PTS3	-0.0157563	0.0165345	-0.95	0.342	-0.0484419	0.0169294
BS_PTS4	0.0514851	0.0304169	1.69	0.093	-0.0086435	0.1116136
BS_PTS5	0.008951	0.0136615	0.66	0.513	-0.0180552	0.0359572
BS_PTS6	-0.002794	0.0086122	-0.32	0.746	-0.0198186	0.0142307
BS_PTS7	-0.0056103	0.0037726	-1.49	0.139	-0.013068	0.0018474
_cons	-0.2172643	0.1780461	-1.22	0.224	-0.5692278	0.1346991

Number of obs 151
F(8, 142) 27.05
Prob > F 0
R-squared 0.5142
Root MSE 0.66116

Étonnamment, les résultats montrent que la variable est non négligeable pour la première régression et très significative pour la deuxième mais, surtout, possède un coefficient positif. Cette anomalie est au cœur de ce mémoire, et les prochaines étapes auront pour but de faire disparaître cette variable, soit de la rendre non significative comme le suggère la théorie à cet effet, tout en continuant l'exploration des autres variables d'intérêt au sein de l'hypothèse de base. Effectivement, dans ce modèle, un joueur qui a marqué plus de points dans les trois dernières saisons et qui a été repêché plus tard est plus susceptible de recevoir un salaire plus élevé qu'un joueur aux performances semblables repêché plus tôt. D'autre part, l'inclusion du rang au repêchage n'a pas substantiellement altéré les coefficients des variables significatives précédemment, la principale variation étant l'accroissement de la négativité du coefficient de la constante.

Pour ce qui est de la deuxième régression, il est important de constater l'ampleur de la significativité de la variable *Draft*. Bien que son coefficient soit relativement inférieure, l'ajout de cette variable suggère, a priori, que les directeurs généraux garde toujours un œil sur cette variable lors de l'évaluation salariale d'un joueur, même après les huit années d'expérience de ce derniers. De plus, parmi les effets de l'addition du rang au repêchage, il est possible de constater une augmentation de la significativité de *BS_PTS4*, alors que pour *BS_PTS7* c'est une diminution qui est observée. Ceci porte à croire que la variable correspondant à la pire saison d'un joueur capturerait des variations qui ne lui étaient pas propre, et qui sont mieux représentées par *Draft*. Tout comme c'était le cas dans le modèle de base, la deuxième régression présente une constante non significative.

4.3 – DEUXIÈME VERSION DU MODÈLE

Il se dégage de ce nouveau modèle que le rang au repêchage semble important pour les directeurs des équipes dans la NHL, mais puisque cette variable présente un coefficient positif, à l’opposé des anticipations, il semble s’avérer nécessaire de pousser les investigations plus loin. En effet, il se pourrait que la variable ajoutée reflète des variations appartenant à d’autres variables possiblement omises. La prochaine étape adresse ce problème en incluant, avec *Draft*, la durée des contrats (*CL*).

Ce prochain modèle inclut donc cette nouvelle variable *CL*, et est représenté comme :

$$\text{a) } \ln\text{SCH16} = \text{Draft} + \text{CL} + \beta_{15} \text{PTS15} + \beta_{14} \text{PTS14} + \beta_{12} \text{PTS12} + \beta_{11} \text{PTS11} + \beta_{10} \text{PTS10} \\ + \beta_{09} \text{PTS09} + \beta_{08} \text{PTS08} + \alpha_0 + \varepsilon_0$$

$$\text{b) } \ln\text{SCH16} = \text{Draft} + \text{CL} + \beta_1 \text{BS_PTS1} + \beta_2 \text{BS_PTS2} + \beta_3 \text{BS_PTS3} + \beta_4 \text{BS_PTS4} + \beta_5 \text{BS_PTS5} \\ + \beta_6 \text{BS_PTS6} + \beta_7 \text{BS_PTS7} + \alpha_1 + \varepsilon_1$$

L’idée ici est que les joueurs choisis hâtivement au repêchage sont plus susceptibles de s’engager dans un contrat à long terme avec les équipes, un facteur qui peut influencer leur salaire. Conséquemment, les attentes généralement acceptées en la matière suggèrent qu’un joueur devrait accepter un salaire moins élevé en échange de la stabilité qu’apporte un contrat à plus long terme. Ceci expliquerait pourquoi le coefficient de *Draft* est positif. Les tableaux VII et VIII, accompagnés par leur sommaire statistique, présentent les résultats obtenus.

Tableau VII : Régression chronologique du deuxième modèle

InSCH16	Coef.	Std. Err.	t	P>t	[95% Conf. Interval]	
Draft	0.0006197	0.0004469	1.39	0.168	-0.0002638	0.0015032
CL	0.104876	0.0240503	4.36	0	0.0573302	0.1524219
PTS15	0.0040809	0.0037295	1.09	0.276	-0.0032921	0.0114539
PTS14	0.0088921	0.0034749	2.56	0.012	0.0020225	0.0157617
PTS12	0.0117031	0.0030597	3.82	0	0.0056542	0.0177519
PTS11	0.0053304	0.0038587	1.38	0.169	-0.0022981	0.0129588
PTS10	0.0001196	0.0027968	0.04	0.966	-0.0054096	0.0056487
PTS09	-0.0044902	0.0058086	-0.77	0.441	-0.0159734	0.006993
PTS08	0.0018158	0.0043219	0.42	0.675	-0.0067283	0.0103599
_cons	-0.4628135	0.1563813	-2.96	0.004	-0.7719685	-0.1536584

Number of obs 151
F(9, 141) 26.03
Prob > F 0
R-squared 0.5916
Root MSE 0.60832

Tableau VIII : Régression par performances du deuxième modèle

InSCH16	Coef.	Std. Err.	t	P>t	[95% Conf. Interval]	
Draft	0.0011712	0.0006052	1.94	0.055	-0.0000252	0.0023676
CL	0.1097454	0.0217556	5.04	0	0.0667361	0.1527547
BS_PTS1	-0.016023	0.013277	-1.21	0.23	-0.0422708	0.0102248
BS_PTS2	0.0222996	0.0094743	2.35	0.02	0.0035696	0.0410296
BS_PTS3	-0.028408	0.0164206	-1.73	0.086	-0.0608704	0.0040543
BS_PTS4	0.0496242	0.0285768	1.74	0.085	-0.0068701	0.1061185
BS_PTS5	0.0036526	0.0123566	0.3	0.768	-0.0207755	0.0280807
BS_PTS6	-0.0044366	0.0079293	-0.56	0.577	-0.0201123	0.0112391
BS_PTS7	-0.0011036	0.004253	-0.26	0.796	-0.0095114	0.0073042
_cons	-0.303945	0.1649429	-1.84	0.067	-0.6300258	0.0221359

Number of obs 151
F(9, 141) 22.69
Prob > F 0
R-squared 0.5851
Root MSE 0.61312

Maintenant, avec l'ajout de la durée des contrats, les résultats changent de façon plus substantielle. Tout d'abord, il est important de constater que pour les deux régressions, la variable *CL* est remarquablement significative et son coefficient élevé en comparaison à ceux des autres variables. Par contre, ce dernier est positif, ce qui va à l'encontre de la croyance populaire et largement répandue mentionnée plus haut. En effet, selon les résultats observés, un joueur qui signe à plus long terme devrait recevoir un salaire plus important. Ceci pourrait être expliqué de la façon suivante : Premièrement, les joueurs choisis plus haut au repêchage sont vraisemblablement plus susceptibles de s'engager dans un contrat à long terme et exiger un salaire plus élevé pour leur plus grand potentiel. Aussi, ils peuvent demander une rémunération supplémentaire de leur équipe d'accueil pour les priver de pouvoir tester le marché des joueurs autonomes après un contrat plus court, ce qui pourrait leur permettre de signer pour un salaire plus élevé.

En ce qui concerne la première régression, les points récoltés en 2015 ne sont plus significatifs et la variable *Draft* l'est relativement moins aussi. Si cette dernière ne devrait logiquement pas être significative après huit années d'expérience dans la NHL, il apparaît moins clair pourquoi les performances de 2015 ne le seraient pas. Une explication possible est que la production offensive de 2015 est possiblement trop près temporellement de 2016, et qu'une partie des joueurs était déjà sous contrat, impliquant conséquemment que les directeurs devaient regarder les années antérieures à 2015. Ceci est cohérent avec le gain important en significativité de la variable *PTS11*.

Pour ce qui a trait à la deuxième régression, deux principaux changements ont été observés. D'une part, il s'agit des variables *BS_PTS2* et *BS_PTS3* qui sont maintenant significatives. Dorénavant, la régression indique que les dirigeants des équipes regardent les deuxième, troisième et quatrième meilleures saisons d'un joueur pour évaluer son juste salaire, en plus de son rang au repêchage et la durée du contrat offert. De plus, les trois saisons suivent un schème particulier, avec un coefficient négatif pour *BS_PTS3* et une décroissance de l'importance absolue des coefficients de la quatrième à la deuxième meilleure production offensive. Ceci porte à croire que les directeurs sont sensibles au risque, et qu'ils ont une certaine préférence pour des performances constantes.

4.4 – TROISIÈME VERSION DU MODÈLE

Ainsi, deux principaux buts justifient le prochain modèle : Premièrement, il vise toujours à expliquer, et si possible éliminer, la significativité de la variable correspondant au rang repêché. Deuxièmement, il cherche à éclaircir les préférences vis-à-vis le risque que peuvent avoir les dirigeants des différentes franchises de la Ligue. De cette façon, en plus des variables précédemment ajoutées au modèle de base, le prochain modèle inclut l'écart-type des points inscrits par les joueurs entre les saisons 2007-2008 et 2014-2015 (*stdPTS*).

Il est représenté comme suit :

- a) $\lnSCH16 = \text{Draft} + \text{stdPTS} + \text{CL} + \beta_{15} \text{PTS15} + \beta_{14} \text{PTS14} + \beta_{12} \text{PTS12} + \beta_{11} \text{PTS11} + \beta_{10} \text{PTS10} + \beta_{09} \text{PTS09} + \beta_{08} \text{PTS08} + \alpha_0 + \varepsilon_0$
- b) $\lnSCH16 = \text{Draft} + \text{stdPTS} + \text{CL} + \beta_1 \text{BS_PTS1} + \beta_2 \text{BS_PTS2} + \beta_3 \text{BS_PTS3} + \beta_4 \text{BS_PTS4} + \beta_5 \text{BS_PTS5} + \beta_6 \text{BS_PTS6} + \beta_7 \text{BS_PTS7} + \alpha_1 + \varepsilon_1$

Ici, il n'y a pas vraiment de consensus à savoir comment la variable *stdPTS* devrait se comporter théoriquement. D'un côté, il est convenable d'énoncer qu'un joueur avec une performance plus constante peut être plus valorisé, car son apport à l'équipe est fiable et, par conséquent, plus aisément mesurable. Dans un tel cas, le coefficient devant la variable correspondant à l'écart-type des points devrait être négatif afin de refléter une pénalité pour une variabilité plus grande. D'un autre côté, un joueur connaissant plus de variations dans sa production offensive avec une moyenne semble à celle d'un joueur plus constant, mais avec un maximum atteint plus grand, peut aussi justifier un salaire plus élevé. Il peut être perçu comme un joueur avec un plus grand potentiel, ayant déjà prouvé l'étendue de son talent avec relativement plus de points à sa meilleure saison qu'un joueur, au talent comparable, plus fiable. De cette façon, le coefficient de *stdPTS* devrait être positif, comme quoi un joueur aux performances plus variables se voit offrir un salaire plus élevé, toute chose étant égale par ailleurs. Le tableau XI, de même que le tableau X, tous les deux suivis des statistiques générales des régressions, montrent les résultats obtenus.

Tableau IX : Régression chronologique du troisième modèle

InSCH16	Coef.	Std. Err.	t	P>t	[95% Conf. Interval]	
Draft	0.0006098	0.0004564	1.34	0.184	-0.0002925	0.0015122
stdPTS	-0.0090871	0.0135749	-0.67	0.504	-0.0359255	0.0177513
CL	0.105102	0.0243546	4.32	0	0.0569517	0.1532523
PTS15	0.0043443	0.0039915	1.09	0.278	-0.0035471	0.0122357
PTS14	0.0086239	0.0033295	2.59	0.011	0.0020413	0.0152065
PTS12	0.0125027	0.0036697	3.41	0.001	0.0052476	0.0197579
PTS11	0.0052592	0.003874	1.36	0.177	-0.0023999	0.0129182
PTS10	0.000965	0.0028641	0.34	0.737	-0.0046975	0.0066276
PTS09	-0.0039169	0.0051052	-0.77	0.444	-0.0140102	0.0061764
PTS08	0.0009788	0.0035848	0.27	0.785	-0.0061086	0.0080662
_cons	-0.4042751	0.1628809	-2.48	0.014	-0.7262994	-0.0822508

Number of obs 151
F(10, 140) 23.43
Prob > F 0
R-squared 0.5936
Root MSE 0.60898

Tableau X : Régression par performances du troisième modèle

InSCH16	Coef.	Std. Err.	t	P>t	[95% Conf. Interval]	
Draft	0.0011215	0.0005833	1.92	0.057	-0.0000318	0.0022748
stdPTS	-0.0649942	0.1274179	-0.51	0.611	-0.3169064	0.1869179
CL	0.1110241	0.0229248	4.84	0	0.0657005	0.1563476
BS_PTS1	0.0004745	0.0286986	0.02	0.987	-0.0562642	0.0572133
BS_PTS2	0.0307392	0.0221292	1.39	0.167	-0.0130114	0.0744899
BS_PTS3	-0.0237824	0.0169315	-1.4	0.162	-0.0572569	0.0096922
BS_PTS4	0.0484769	0.0278384	1.74	0.084	-0.0065612	0.1035149
BS_PTS5	0.0015895	0.0134017	0.12	0.906	-0.0249065	0.0280854
BS_PTS6	-0.0132233	0.019576	-0.68	0.5	-0.0519262	0.0254795
BS_PTS7	-0.0189252	0.0359653	-0.53	0.6	-0.0900306	0.0521802
_cons	-0.2928304	0.1621326	-1.81	0.073	-0.6133752	0.0277144

Number of obs 151
F(10, 140) 20.78
Prob > F 0
R-squared 0.586
Root MSE 0.6147

D'entrée de jeu, il apparaît évident que l'inclusion de l'écart-type de la production offensive des joueurs n'est pas pertinente dans ce modèle. En effet, celle-ci n'est pas significative, et de façon substantielle, autant pour la régression avec la série chronologique que pour celle de la série ordonnée par meilleures saisons. Néanmoins, ce résultat indique que les professionnels en charge des équipes sont indifférents au risque, plus précisément à la variation de la production offensive d'un joueur. Ceci revient à dire que les directeurs généraux dans la NHL sont indifférents vis-à-vis la signature d'un joueur constant par rapport à un joueur qui ne l'est pas, tous les deux étant identiques en ce qui a trait à leurs autres caractéristiques. Par ailleurs, il faut aussi noter que le coefficient de la durée des contrats est toujours significativement différent de zéro, preuve comme quoi son addition au modèle est vraisemblablement pertinente.

Pour la première régression, il s'en dégage que la durée du contrat et les saisons se terminant en 2012 et 2014 sont les principales variables importantes aux yeux des directeurs généraux. Ceux-ci, sur la base de ces résultats, semblent raisonnables et sensibles, d'autant plus que la significativité du rang au repêchage a été écartée.

Pour la deuxième régression, une telle conclusion est moins plausible. En effet, bien que la médiane semble être un bon estimateur de la contribution potentielle d'un joueur, et qu'il apparaît normal que la durée du contrat influence le salaire négocié, la significativité de la variable *Draft* persiste. Les directeurs dans la Ligue accordent toujours, selon ce modèle, une importance au rang auquel un joueur a été repêché, au moins huit ans après que celui-ci ait fait ses débuts dans la NHL.

4.5 – QUATRIÈME VERSION DU MODÈLE

De tels résultats suggèrent qu'une variable reste omise jusqu'à présent. Bien que l'écart-type des points inscrits ne soit pas significatif, il ne semble pas justifier de l'écarter pour l'instant, car il pourrait être pertinent avec l'ajout d'une autre variable. En effet, il est mentionné depuis le début de cette section de l'expérience d'un joueur chez les professionnels. Ainsi, l'inclusion de l'âge d'un joueur (*Age*) dans le modèle paraît tout à fait raisonnable. D'une part, cette variable reflète l'expérience d'un joueur et d'autre part, la maturité générale de celui-ci. De plus, elle adresse une problématique nouvelle aux régressions réalisées jusqu'ici : avec les résultats obtenus, plus un joueur est repêché tard, voire jamais, plus il est susceptible de gagner plus cher. Or, les joueurs non repêchés ou sélectionnés tardivement sont plus susceptibles, raisonnablement, de faire leur entrée dans la Ligue nationale plus tard, à un âge plus avancé. De plus, ce type d'attaquant représente un plus grand risque quant à leur niveau de performance, justifiant la rétention de la variable *stdPTS*.

De cette façon, cette nouvelle version du modèle étudié prend la forme ci-dessous :

- a) $\lnSCH16 = \text{Draft} + \text{Age} + \text{stdPTS} + \text{CL} + \beta_{15} \text{PTS15} + \beta_{14} \text{PTS14} + \beta_{12} \text{PTS12} + \beta_{11} \text{PTS11} + \beta_{10} \text{PTS10} + \beta_{09} \text{PTS09} + \beta_{08} \text{PTS08}$
- b) $\lnSCH16 = \text{Draft} + \text{Age} + \text{stdPTS} + \text{CL} + \beta_1 \text{BS_PTS1} + \beta_2 \text{BS_PTS2} + \beta_3 \text{BS_PTS3} + \beta_4 \text{BS_PTS4} + \beta_5 \text{BS_PTS5} + \beta_6 \text{BS_PTS6} + \beta_7 \text{BS_PTS7}$

Ainsi, il est possible d'argumenter en la faveur d'un coefficient négatif pour *Age*. En effet, bien qu'un joueur plus vieux puisse indiquer un niveau d'expérience plus élevé généralement, ceci n'est pas nécessairement le cas pour un joueur repêché tard ou jamais sélectionné tel que mentionné précédemment. Aussi, en plus d'expliquer cette possibilité, un coefficient négatif pourrait venir corriger pour le coefficient positif du rang au repêchage. Les tableaux XI et XII, et leur sommaire statistique, montrent les résultats observés à cet effet.

Tableau XI : Régression chronologique du quatrième modèle

InSCH16	Coef.	Std. Err.	t	P>t	[95% Conf. Interval]	
Draft	0.0006706	0.0004591	1.46	0.146	-0.0002371	0.0015784
Age	-0.0129535	0.0048994	-2.64	0.009	-0.0226398	-0.0032672
stdPTS	-0.0089636	0.0133573	-0.67	0.503	-0.0353717	0.0174445
CL	0.1020398	0.0237149	4.3	0	0.0551541	0.1489255
PTS15	0.0037867	0.0039647	0.96	0.341	-0.0040516	0.0116251
PTS14	0.0081771	0.003305	2.47	0.015	0.0016429	0.0147113
PTS12	0.0127066	0.0036974	3.44	0.001	0.0053967	0.0200165
PTS11	0.0056889	0.0038589	1.47	0.143	-0.0019405	0.0133182
PTS10	0.000829	0.0028226	0.29	0.769	-0.0047514	0.0064093
PTS09	-0.0039035	0.0051094	-0.76	0.446	-0.0140051	0.0061981
PTS08	0.0017865	0.0035683	0.5	0.617	-0.0052682	0.0088413

Number of obs 151
F(11, 140) 175.35
Prob > F 0
R-squared 0.8355
Root MSE 0.60783

Tableau XII : Régression par performances du quatrième modèle

InSCH16	Coef.	Std. Err.	t	P>t	[95% Conf. Interval]	
Draft	0.0012233	0.0005847	2.09	0.038	0.0000674	0.0023793
Age	-0.0104519	0.0049025	-2.13	0.035	-0.0201444	-0.0007594
stdPTS	-0.0695415	0.1280686	-0.54	0.588	-0.32274	0.1836571
CL	0.1075101	0.0223814	4.8	0	0.0632608	0.1517593
BS_PTS1	0.0025664	0.0289984	0.09	0.93	-0.0547649	0.0598977
BS_PTS2	0.0318226	0.0221879	1.43	0.154	-0.0120441	0.0756893
BS_PTS3	-0.0239903	0.0170917	-1.4	0.163	-0.0577814	0.0098009
BS_PTS4	0.047532	0.0279761	1.7	0.092	-0.0077781	0.1028422
BS_PTS5	0.0022274	0.0132936	0.17	0.867	-0.0240546	0.0285095
BS_PTS6	-0.0141423	0.0196628	-0.72	0.473	-0.0530168	0.0247322
BS_PTS7	-0.0198697	0.0361632	-0.55	0.584	-0.0913663	0.051627

Number of obs 151
F(11, 140) 149.58
Prob > F 0
R-squared 0.8333
Root MSE 0.61192

Tout d'abord, il faut mentionner que, pour ces régressions, le terme constant a été retiré. En effet, ce dernier n'était plus significatif suite à l'ajout de la variable *Age*, indiquant que celle-ci capturait en partie la même information en plus d'inclure d'autres variations présentes dans les données. De plus, l'âge des joueurs est un facteur très significatif dans ce modèle – sa *p-value* étant près de zéro –, et réagit conformément aux anticipations formulées précédemment. À remarquer, son inclusion a l'effet opposé à celui désiré sur la variable *Draft*. En effet, avec *Age* dans les deux régressions entraîne une augmentation de la significativité de la variable représentant la position au repêchage, en plus de laisser l'écart-type des points non significatif. Par contre, l'âge du joueur, par son coefficient plus grand en absolu, vient corriger l'incohérence obtenue antérieurement. Il est dorénavant observé qu'en théorie, un joueur repêché tard mais qui fait son entrée tôt dans la NHL est plus susceptible de gagner un salaire plus élevé. Bien qu'il existe des joueurs de ce genre, ceux-ci représentent plus des exceptions. En pratique, il est plus convenable que ce soient les joueurs repêchés hâtivement qui vont se voir offrir un contrat à un âge moins avancé. Ceci est un résultat intéressant de ce modèle, car il explique autant bien les exceptions que le cheminement courant d'un joueur en ce qui a trait à ces deux variables. Aussi, il est important de souligner que le coefficient de corrélation de chaque régression (R^2) est très élevé, indiquant que le modèle capture une grande partie des phénomènes qui ocurrent dans les données.

Les résultats de la première régression indiquent maintenant davantage que la durée des contrats, l'âge d'un joueur, et les saisons 2011-2012 et 2013-2014 sont les principaux facteurs influençant le salaire négocié à un joueur. Ceci semble raisonnable, car pour les performances passées, un directeur général ne veut pas nécessairement s'intéresser aux saisons trop anciennes étant donné la possible évolution d'un joueur, et la durée et l'âge de celui-ci ont été expliqués plus haut. De plus, bien qu'ils soient au-delà du seuil de 10% de significativité, le rang au repêchage ainsi que la performance en 2015 restent des variables non négligées pour une bonne partie des contrats octroyés par les directeurs généraux.

En ce qui a trait à la deuxième régression par ordre de meilleures saisons, la variable *Draft* reste toujours très importante pour expliquer les salaires des joueurs en 2016, tout comme *Age*, *CL*, et *BS_PTS4*. Tel qu'observé dans la première régression, l'écart-type reste non

significatif, suggérant qu'elle est tout simplement non pertinente aux yeux des directeurs des franchises. Ainsi, pour ce qui est des performances, la médiane reste le meilleur estimateur dans l'évaluation des salaires à verser, et elle conserve le deuxième plus grand coefficient parmi les variables significatives, démontrant son importance relative.

4.6 – VERSION FINALE DU MODÈLE

Avec tous ces résultats, il est possible de construire un modèle final pour ce mémoire pour comprendre les facteurs qui influencent les professionnels en charge des équipes dans l'évaluation du salaire qu'un joueur mérite. Celui-ci change du modèle précédent par le retrait de la variable *stdPTS*, dû à la persistance de sa non-significativité. Néanmoins, les autres variables relatives aux performances ont été conservées, malgré que certaines d'entre elles soient restées non significatives tout au long de ce travail, afin de conserver le modèle de base et permettre la comparaison de celles-ci.

Ce dernier modèle se traduit donc comme :

$$\text{a) } \ln\text{SCH16} = \text{Draft} + \text{Age} + \text{CL} + \beta_{15} \text{PTS15} + \beta_{14} \text{PTS14} + \beta_{12} \text{PTS12} + \beta_{11} \text{PTS11} + \beta_{10} \text{PTS10} \\ + \beta_{09} \text{PTS09} + \beta_{08} \text{PTS08}$$

$$\text{b) } \ln\text{SCH16} = \text{Draft} + \text{Age} + \text{CL} + \beta_1 \text{BS_PTS1} + \beta_2 \text{BS_PTS2} + \beta_3 \text{BS_PTS3} + \beta_4 \text{BS_PTS4} \\ + \beta_5 \text{BS_PTS5} + \beta_6 \text{BS_PTS6} + \beta_7 \text{BS_PTS7}$$

Les hypothèses derrière cette nouvelle version du modèle de base restent les mêmes qu'à l'étape précédente, avec comme seule exception le retrait de *stdPTS* sur la base de l'indifférence vis-à-vis la variabilité de la production offensive d'un joueur des directeurs de la Ligue. Il est ainsi anticipé qu'aucun changement majeur ne sera observé au sein des coefficients. Les tableaux XIII et XIV, suivis des statistiques générales des régressions réalisées, montrent les résultats obtenus.

Tableau XIII : Régression chronologique du modèle final

InSCH16	Coef.	Std. Err.	t	P>t	[95% Conf. Interval]	
Draft	0.0006831	0.0004504	1.52	0.132	-0.0002073	0.0015735
Age	-0.0147117	0.0047806	-3.08	0.003	-0.0241626	-0.0052608
CL	0.1013397	0.0230018	4.41	0	0.0558666	0.1468127
PTS15	0.0034485	0.0036451	0.95	0.346	-0.0037577	0.0106547
PTS14	0.0083736	0.0034117	2.45	0.015	0.0016288	0.0151183
PTS12	0.0119369	0.003135	3.81	0	0.0057391	0.0181346
PTS11	0.0057968	0.0038095	1.52	0.13	-0.0017343	0.0133279
PTS10	-0.0000264	0.0027436	-0.01	0.992	-0.0054503	0.0053975
PTS09	-0.004466	0.0057954	-0.77	0.442	-0.0159231	0.0069912
PTS08	0.0027206	0.0043652	0.62	0.534	-0.005909	0.0113503

Number of obs 151
F(10, 141) 186.63
Prob > F 0
R-squared 0.8347
Root MSE 0.60716

Tableau XIV : Régression par performances du modèle final

InSCH16	Coef.	Std. Err.	t	P>t	[95% Conf. Interval]	
Draft	0.0012725	0.0006062	2.1	0.038	0.0000741	0.002471
Age	-0.0107216	0.0049816	-2.15	0.033	-0.0205699	-0.0008734
CL	0.1059766	0.0210993	5.02	0	0.0642649	0.1476884
BS_PTS1	-0.0151335	0.0131715	-1.15	0.253	-0.0411726	0.0109057
BS_PTS2	0.0227804	0.0095281	2.39	0.018	0.0039441	0.0416167
BS_PTS3	-0.028924	0.0164181	-1.76	0.08	-0.0613816	0.0035335
BS_PTS4	0.0487191	0.0286599	1.7	0.091	-0.0079395	0.1053778
BS_PTS5	0.0044391	0.0122663	0.36	0.718	-0.0198105	0.0286886
BS_PTS6	-0.0046721	0.0078324	-0.6	0.552	-0.0201563	0.0108121
BS_PTS7	-0.0008014	0.0042285	-0.19	0.85	-0.009161	0.0075581

Number of obs 151
F(10, 141) 167.27
Prob > F 0
R-squared 0.8329
Root MSE 0.61045

Tout d'abord, il apparaît de plus en plus clair que la variable *Draft* influence les directeurs généraux, cette dernière ayant gagné en significativité par le retrait de *stdPTS* dans les deux régressions. À cet effet, la dernière partie de ce chapitre portera sur l'explication de sa significativité selon la perspective de l'économie comportementale, avec un accent mis sur l'effet d'ancrage. Par ailleurs, l'âge et la durée des contrats des joueurs sont toujours très significatifs tel qu'anticipé, et il ne faut pas passer sous silence l'ampleur des coefficients de corrélations de chaque régression, les deux avoisinant 83%.

La première régression, à partir de la série chronologique, affiche encore plus un intérêt centré sur les saisons 2011-2012 et 2013-2014, de même que sur celle de 2010-2011 mais de façon moins significative. À l'inverse, il semble que la saison 2014-2015 intéresse moins les directeurs généraux, possiblement à cause qu'elle est trop récente pour les contrats actifs en 2016. De plus, autant pour cette régression que la suivante, la constante n'a pas été prise en compte pour les mêmes raisons qu'à l'étape précédente.

Pour ce qui concerne la deuxième régression, les résultats ont relativement plus changé. En effet, il est possible de constater que la variable la plus significative est maintenant la deuxième meilleure saison. Ceci est étonnant, tout comme le gain en significativité de *BS_PTS3*, maintenant sous le seuil du 10% de significiance. C'est donc dire que selon ces nouveaux résultats, les directeurs généraux valorisent principalement, au niveau des performances, la médiane de la série ordonnée par meilleures saisons et les saisons en deuxième et troisième positions mais de façon relativement moindre à cause de leur coefficient moins élevé. En plus, les variables *Age* et *CL* sont restées très significatives, la première corrigeant encore pour la variable *Draft* et la deuxième étant encore la variable ayant le plus grand effet sur le salaire des joueurs.

4.7 – EXPLICATION DE L'ANOMALIE

Possible le résultat le plus intéressant de ce mémoire, la variable *Draft*, référant au rang qu'un joueur a été repêché, est restée significative à chaque modification du modèle de base suivant son introduction. De plus, lors des deux dernières étapes, les coefficients de corrélations des régressions étaient particulièrement élevés, indiquant que le modèle explique en grande partie les variations au sein des données.

Ainsi, une façon de comprendre cette anomalie réside dans l'effet d'ancrage, un biais proposé par l'économie comportementale. Tel que mentionné au début de ce travail, la perspective abordée de ce phénomène est celle de Kahneman et Tversky, soit l'ancrage et l'ajustement. Celle-ci suggère, en lien avec les résultats obtenus, que même après huit ans de données sur les performances des joueurs, les directeurs généraux valorisent toujours le rang auquel ceux-ci ont été repêchés alors qu'ils ne devraient pas théoriquement. Il faut comprendre que l'ancrage survient de façon inconsciente chez un individu. De cette façon, les dirigeants des équipes ont en tête la position au repêchage des joueurs qu'ils mettent sous contrats, influençant à leur insu le salaire qu'ils sont prêts à leur verser. Le biais ici peut être compris comme la simple considération de *Draft*, incitant les directeurs à systématiquement sous-payer un joueur repêché relativement plus tôt qu'un autre possédant les mêmes caractéristiques mais repêché plus tard. En effet, non seulement la significativité du rang au repêchage est anormale, mais son coefficient aussi. En théorie, le coefficient aurait dû, dans le cas déjà particulier où il est significatif, être négatif afin de récompenser un joueur sélectionné hâtivement. Néanmoins, le fait qu'il soit positif est cohérent avec la théorie de l'ancrage et l'ajustement, celle-ci proposant qu'en cas d'ancrage, un individu est incapable de différencier suffisamment son estimation de la valeur initiale suggérée.

Une autre façon d'analyser l'effet d'ancrage retrouvé dans les résultats est que le rang au repêchage crée des attentes chez les directeurs généraux. Par conséquent, un joueur repêché plus tard peut surprendre par ses performances et, ainsi, inciter les dirigeants des équipes à leur verser un salaire plus élevé. Cependant, un joueur repêché tôt est associé à des attentes élevées. Il est ainsi plus probable que ce dernier entraîne une déception engendrant une

pénalité salariale. Dans un tel scénario, l’ancrage réside en les performances attendues d’un choix hâtif au repêchage et est mieux compris par la théorie sur l’accessibilité sélective. En effet, les directeurs sont susceptibles d’attendre d’un choix de première ronde, par exemple, à ce qu’il performe comme des joueurs d’exception – tels que Sidney Crosby – et, ainsi, être déçus des résultats obtenus par celui-ci. Les responsables des équipes ne retiennent donc que l’information positive sur l’ancrage, niant d’une certaine façon les estimations plus sensées des performances anticipées d’un attaquant choisi au premier tour.

CHAPITRE 5 - ROBUSTESSE DES RÉSULTATS

La première considération pouvant être soulevée à l'égard des conclusions de ce mémoire ont trait au manque de variables de contrôle. Bien qu'il s'agisse d'un modèle en forme réduite, les régressions présentées ne représentent que la partie pertinente du travail réalisé. En effet, d'autres régressions incluant une partie sinon l'ensemble des variables de contrôle possibles, correspondant à l'ensemble des statistiques générales disponibles pour chaque joueur, ont été réalisées. Toutefois, aucune de ces variables non présentées ici ont montré des coefficients significatifs persistant au travers des différentes saisons de l'échantillon. Ceci étant dit, aucune de ces variables de contrôle incluait des statistiques avancées, telles que le Corsi, les données relatives au temps de possession de la rondelle ou en lien au nombre de jeux complétés par exemple.

De plus, il est plusieurs fois mention de la raison et de la sensibilité des directeurs généraux dans ce mémoire. Pour comprendre ces caractéristiques, un modèle a été construit sur la même base que le modèle final mais ayant les points inscrits en 2016 comme variable dépendante. Ainsi, les tableaux XV et XVI, accompagnés des sommaires statistiques de chaque régression, montrent les résultats obtenus :

Tableau XV : Test de rationalité selon la série chronologique

PTS16	Coef.	Std. Err.	t	P>t	[95% Conf. Interval]	
Draft	0.0091098	0.0094784	0.96	0.338	-0.0096295	0.027849
Age	-0.1282725	0.0695203	-1.85	0.067	-0.2657179	0.0091728
stdPTS	-0.4947766	0.2329079	-2.12	0.035	-0.955248	-0.0343052
CL	0.5168096	0.4822958	1.07	0.286	-0.436715	1.470334
PTS15	0.600406	0.0878993	6.83	0	0.4266243	0.7741876
PTS14	0.3742199	0.078566	4.76	0	0.2188907	0.5295491
PTS12	-0.0648461	0.0631124	-1.03	0.306	-0.1896228	0.0599305
PTS11	0.064251	0.0682849	0.94	0.348	-0.070752	0.199254
PTS10	0.1715818	0.0894616	1.92	0.057	-0.0052886	0.3484522
PTS09	0.0184918	0.0700629	0.26	0.792	-0.1200264	0.1570099
PTS08	-0.1536412	0.06241	-2.46	0.015	-0.2770291	-0.0302534

Number of obs 151
F(11, 140) 143.76
Prob > F 0
R-squared 0.9147
Root MSE 11.661

Tableau XVI : Test de rationalité selon la série par meilleures saisons

PTS16	Coef.	Std. Err.	t	P>t	[95% Conf. Interval]	
Draft	0.0121761	0.0151859	0.8	0.424	-0.0178472	0.0421995
Age	-0.1921697	0.0838255	-2.29	0.023	-0.3578972	-0.0264423
stdPTS	3.148591	2.873166	1.1	0.275	-2.531813	8.828995
CL	1.286488	0.6226536	2.07	0.041	0.0554687	2.517508
BS_PTS1	-0.8184733	0.7906931	-1.04	0.302	-2.381716	0.7447695
BS_PTS2	-0.7065254	0.5105	-1.38	0.169	-1.715811	0.3027604
BS_PTS3	-0.3224086	0.3641476	-0.89	0.377	-1.042348	0.3975308
BS_PTS4	1.396515	0.3748452	3.73	0	0.6554261	2.137604
BS_PTS5	0.0159855	0.4560127	0.04	0.972	-0.885576	0.917547
BS_PTS6	0.3375467	0.4525825	0.75	0.457	-0.5572332	1.232327
BS_PTS7	0.9656097	0.8238778	1.17	0.243	-0.6632409	2.59446

Number of obs 151
F(11, 140) 74.9
Prob > F 0
R-squared 0.8667
Root MSE 14.574

En comparant ces résultats avec ceux obtenus pour le quatrième modèle – incluant *Draft*, *Age*, *stdPTS* et *CL* –, il est ainsi possible de constater que, pour la régression chronologique, le rang au repêchage n'est pas significatif pour la production offensive (*PTS16*) mais l'est relativement plus pour les salaires en 2016 en comparaison aux autres variables. Ceci renvoie l'information comme quoi le rang repêché influence le montant versé à un joueur alors que cette variable n'est pas déterminante de son potentiel en points. À noter, une disparité de la significativité est observée pour les variables *stdPTS* et *CL*. En ce qui concerne l'écart-type des points, il est possible de comprendre cette incohérence, d'une part, par les préférences des directeurs en ce qui a trait au risque et, d'autre part, parce que l'écart-type des points passés peut s'avérer pertinent pour estimer la production offensive future. Pour ce qui est de la durée des contrats, sa pertinence pour les salaires a été expliquée dans la section « Résultats ». De plus, il paraît justifié de douter que *CL* ait une influence sur le nombre de points qu'un joueur marquera à une saison ultérieure car, bien qu'il soit possible d'associer de plus longs contrats à des attaquants de plus haut calibre, une telle déduction ne semble pas substantiellement fiable. Par ailleurs, une conclusion similaire peut être tirée de la régression selon la série des meilleures saisons. En effet, celle-ci montre aussi un coefficient non significatif devant la variable *Draft* tandis que la régression comparable du quatrième modèle affiche un coefficient significatif. Toutefois, l'ensemble des autres variables ajoutées au modèle de base sont cohérentes quant à leur significativité dans les deux régressions. Il semble donc que les directeurs généraux soient relativement plus sensibles et rationnels lorsque les régressions chronologiques, car l'importance qu'ils accordent au rang au repêchage est moindre – au niveau de la significativité des coefficients – et parce que l'incohérence entre les autres résultats peut être justifiée. De plus, il est à considérer que les mêmes variables relatives aux saisons des joueurs dans les deux régressions des deux modèles sont significatives, à l'exception des régressions chronologiques où ces 2015 est significative dans le test de la rationalité des dirigeants sans l'être dans le quatrième modèle, et de 2012 qui est significative dans le quatrième modèle sans l'être dans l'autre. Il est possible de comprendre ceci simplement par le fait que plusieurs joueurs dans la série de données avaient déjà un contrat avant 2016.

En lien avec ce test de rationalité vient un questionnement sur la robustesse des résultats : il se pourrait que le nombre de points inscrits par saison ne soit pas un bon estimateur – il en existe peut-être un meilleur – ou qu’il ne soit pas le seul dont les responsables des équipes tiennent compte. À cet égard, une régression a été réalisée pour mettre en lumière quelles variables devaient être considérées, tel que mentionné précédemment, mais il se peut que le problème persiste.

Ensuite, tel que mentionné précédemment, toutes les régressions ont été réalisées avec l’option *robust* de Stata. Celle-ci permet d’inclure l’hétéroscédasticité dans le modèle, de corriger pour une certaine non-normalité des données et, autant que possible, pour une colinéarité dans les variables. Néanmoins, ce dernier aspect ne peut être négligé et, bien que cela ne semble pas être le cas a priori, il se peut que certaines variables soient corrélées entre elles. En particulier, de la colinéarité pourrait être présente entre les variables concernant les performances des joueurs, entraînant possiblement la non-significativité non justifiée de certains d’entre eux.

Aussi, l’échantillon sélectionné est de huit années, ce qui peut être insuffisant dans certains cas. Lors de sa sélection, un arbitrage a été fait entre le nombre total de saisons incluses et le nombre de joueurs observés. Considérant que chaque année supplémentaire entraînait le retrait d’un certain nombre de joueurs, la saison 2007-2008 a été décidée comme étant la limite. De plus, une justification additionnelle pour ce choix est que plusieurs statistiques n’étaient pas disponibles pour les saisons précédentes, car la NHL n’en détenait pas le compte. À ce moment, il n’était pas encore question de construire un modèle uniquement sur la base du nombre de points inscrits par saison pour chaque joueur, donc éliminer des variables telles que le nombre de tirs bloqués ou le nombre de mises en échec réalisées par un joueur semblait être une perte trop importante. Or, avec les résultats présentés dans ce mémoire, il pourrait être pertinent de refaire le travail avec un échantillon temporel plus grand. La principale raison réside en la présence d’exceptions dans les données, des joueurs aux talents et performances hors du commun comme Sidney Crosby ou Alex Ovechkin. Ces derniers ont signé leur contrat tôt dans leur carrière et pour une longue durée et, par conséquent, ont le même impact salarial sur une bonne partie sinon l’ensemble de l’échantillon.

Par la suite, il faut mentionner que l'omission possible de variables pertinentes peut remettre en cause la robustesse des résultats. De ce fait, il se peut, et c'est fort probable, qu'un joueur dans l'échantillon se soit engagé dans un contrat avant 2016. Ainsi, avec l'évolution de la valeur de l'argent, tel que par l'inflation, un joueur a possiblement reçu un montant juste pour ses performances au moment de la signature de son contrat mais qui est perçu comme insuffisant en 2016 selon ses performances. Idéalement, il aurait été préférable d'avoir comme variable dépendante le contrat et son montant total au moment où il est entré en vigueur. Malheureusement, pour ce faire, il aurait été nécessaire, entre autres, d'avoir accès à l'ensemble des clauses du contrat et d'avoir pour l'ensemble des joueurs le moment de sa signature, sa durée et son montant total, ce qui est impossible à cause de certaines parties contractuelles qui sont inaccessibles au public.

Enfin, il est possible qu'une omission de transformations des variables ait eu lieu. À cet effet, il a été testé si l'ajout du carré de l'âge avait été pertinent. Toutefois, au-delà de ne pas être significative, l'addition de l'âge au carré entraînait le modèle à réagir bizarrement, rendant presque l'ensemble des autres variables non significatif. Aussi, des transformations des variables représentant les performances ont été essayées, mais aucune ne bénéficiait au modèle. Entre autres, le carré ainsi que la racine carré des séries chronologiques et ordonnées par meilleures productions offensives ont été alternativement considérées afin de capturer une possible non-linéarité des données. Toutefois, bien qu'un gain au niveau du coefficient de corrélation ait été réalisé, et principalement avec l'inclusion des racines carrées, le modèle réagissait encore mal à ces additions et affichait des coefficients à la fois non significatifs et de signes incohérents.

CHAPITRE 6 - DISCUSSION

6.1 – MODÈLE

Premièrement, le modèle proposé est linéaire en paramètre, mais il se peut qu'une version non linéaire ait été plus appropriée pour expliquer les phénomènes présents dans les données. Un tel modèle non linéaire pourrait possiblement capter divers effets résultants du croisement de certaines variables et peut-être isoler davantage les effets de chacune d'entre elles. Néanmoins, étant donné l'ampleur des coefficients de corrélations des deux derniers modèles, il semble qu'un tel ajout aurait un effet marginal sur les résultats, bien que la deuxième régression avec la série ordonnée par performances semble plus susceptible à être affectée par un changement considérant la variation des résultats obtenus.

6.2 – MÉTHODOLOGIE

Deuxièmement, la méthodologie utilisée pour ce mémoire se résume en trois étapes générales : Tout d'abord, il y a eu, comme première phase, la formulation de l'hypothèse de base et l'exploration de la littérature à ce sujet. En effet, il a tout d'abord été question de l'idée générale que la connaissance d'une information passée peut potentiellement influencer les prédictions futures. Par la suite, une recherche sur le sujet a permis de mettre en lumière différentes théories à ce sujet tel que l'effet d'ancrage et la proposition de Kahneman et Tversky à ce sujet. Par la suite, après l'exploration de diverses séries de données possibles pour mesurer un tel phénomène, est arrivée la deuxième étape soit la collection des données. Ce mémoire n'a pas eu recours à une expérimentation ni à une quelconque forme de sondage pour l'extraction de ses données. Tel que mentionné dans la section « Description des données », l'ensemble des informations sur les attaquants de la NHL a été trouvé sur le web. Ainsi, la construction d'une base de données s'est avérée nécessaire, et elle a été réalisée à l'aide des logiciels *Microsoft Excel et Microsoft Access*. Certaines variables, telles que la série des meilleures saisons de chaque joueur, ont été entièrement construites à partir des données extraites avec des formules sur *Excel*. De plus, l'association des données pour chaque attaquant a été exécutée à l'aide d'*Access*. Finalement, dans un troisième temps, c'est avec le

logiciel *Stata* que l'analyse statistique de l'information récoltée a été accomplie. De ce fait, les résultats présentés, ainsi que leur fiabilité, découlent de la précision de ce programme et de ses paramètres par défaut. Tout au long de ce processus, plusieurs étapes ont nécessité l'approbation et l'appui du directeur de ce mémoire, dont l'ouverture d'esprit, la compréhension et la disponibilité furent essentiels à la concrétisation de ce travail.

CONCLUSION

Finalement, il se dégage de ce mémoire que l'effet d'ancrage peut subvenir lors de la réalisation d'une estimation. Ce biais, découlant de l'économie comportementale, est décrit par Kahneman et Tversky comme étant l'ajustement insuffisant à partir d'une valeur initiale qu'est l'ancrage, mais il existe d'autres définitions et compréhensions de ce phénomène. Néanmoins, le but de cette théorie réside dans l'explication de certains résultats allant à l'encontre des attentes rationnelles de l'économie traditionnelle, sans toutefois fournir un éclaircissement exhaustif.

Ainsi, c'est par une analyse des salaires selon les performances et caractéristiques des attaquants de la Ligue Nationale de Hockey actifs entre 2007 et 2016 que ce travail tente de démontrer empiriquement l'effet d'ancrage. À partir d'un modèle de base linéaire en paramètres, plusieurs modifications incrémentales ont été apportées afin de cerner un tel biais. Chaque étape avait pour objectif d'isoler et confirmer, par des régressions linéaires, différents effets observés au sein des deux séries de données considérées. La première correspondait à l'ensemble des saisons de l'échantillon ordonnées chronologiquement, tandis que la seconde les ordonnait de la meilleure saison en termes de points inscrits à la pire saison. Il en résulte que de façon générale, les directeurs généraux des différentes équipes se comportent de façon rationnelle lorsqu'ils doivent octroyer des contrats, et conséquemment des salaires, aux attaquants de la Ligue. Toutefois, un possible biais a été trouvé en ce qui a trait au rang auquel un joueur a été repêché. Il s'en découle que les directeurs généraux tendent à sous-payer les joueurs sélectionnés plus tôt en comparaison aux joueurs possédant des caractéristiques identiques mais sélectionnés plus tard.

Évidemment, certains enjeux pourraient remettre en cause la robustesse des résultats. Entre autres, il est possible que des variables pertinentes aient été omises ou que l'échantillon ne comprenne pas assez de saisons observées. De plus, il est possible qu'un modèle non-linéaire capte mieux les variations présentes dans les données. Néanmoins, étant donné qu'un très bon coefficient de corrélation a été obtenu, il semble que le modèle proposé soit fiable et pertinent,

d'autant plus qu'il explique bien plusieurs situations possibles par le comportement de ses coefficients.

Enfin, il serait pertinent de reprendre les recherches avec un échantillon plus grand et comprenant plusieurs contrats à différentes périodes de temps afin de valider les résultats obtenus.

ANNEXES

ANNEXE A

Tableau XVII : Descriptifs de l'ensemble des variables utilisées

Variable	Description
Player	Nom du joueur
Age	Âge du joueur (en années)
Tm	Équipe du joueur en 2016
Pos	Position du joueur (RW: aile droite ; LW: aile gauche ; C: centre)
CAN16	Variable binaire prenant la valeur 1 si l'équipe est au Canada et 0 sinon
SCH16	Impact salarial en 2016
GPxx	Nombre de parties jouées en 20xx
Gxx	Nombre de buts marqués en 20xx
Axx	Nombre d'assistances inscrites en 20xx
PTSxx	Nombre de points inscrits en 20xx
PMxx	Différentiel obtenu en 20xx
PIMxx	Nombre de minutes de pénalités reçu en 20xx
EVGxx	Nombre de buts marqués à forces égales en 20xx
PPGxx	Nombre de buts marqués en avantage numérique en 20xx
SHGxx	Nombre de buts marqués en désavantage numérique en 20xx
GWGxx	Nombre de buts gagnant marqués en 20xx
EVAxx	Nombre d'assistances inscrites à forces égales en 20xx
PPAxx	Nombre d'assistances inscrites en avantage numérique en 20xx
SHAxx	Nombre d'assistances inscrites en désavantage numérique en 20xx
SOGxx	Nombre de tirs au buts en 20xx
S%xx	Pourcentage de tirs au but cadrés en 20xx
TOIxx	Temps total passé sur la glace en 20xx (en minutes)
ATOIxx	Temps moyen passé sur la glace en 20xx (en minutes)
BLKxx	Nombre de tirs bloqués en 20xx
HITxx	Nombre de mises en échec servies en 20xx
FOWxx	Nombre de mises en jeu remportées en 20xx
FOLxx	Nombre de mises en jeu perdues en 20xx
FO%xx	Pourcentage de mises en jeu remportées en 20xx
BS_PTS[i]	ième meilleure saison en termes de points
[k]_BS_PTS[i]	Variable k de la ième meilleure saison en termes de points
[j]_min	Variable j par minute (j/TOI)
BS_PTS_min[i]	ième meilleure saison en termes de points par minute
[k]_BS_PTS_min[i]	Variable k de la ième meilleure saison en termes de points par minute
varPTS	Variance des points récoltés entre les saisons 2007-2008 et 2014-2015
stdPTS	Écart-type des points récoltés entre les saisons 2007-2008 et 2014-2015
mPTS	Moyenne des points récoltés entre les saisons 2007-2008 et 2014-2015

mPTS_var	Moyenne des points récoltés divisée par la variance de ceux-ci entre les saisons 2007-2008 et 2014-2015
SCH_SC	Proportion que représente le salaire d'un joueur sur le plafond salarial d'une équipe en 2016
Risk	Écart-type des points récoltés entre les saisons 2007-2008 et 2014-2015 pondéré par la proportion que représente le salaire d'un joueur sur le plafond salarial d'une équipe en 2016
CV	Coefficient de variation (correspondant à $\text{stdPTS}/\text{mPTS}$)
wCV	Coefficient de variation pondéré par la proportion que représente le salaire d'un joueur sur le plafond salarial d'une équipe en 2016
Draft	Rang auquel un joueur a été repêché
CL	Durée du contrat (en années)
DR	Ronde à laquelle un joueur a été repêché

ANNEXE B

Tableau XVIII : Notation du logiciel Stata

Notation	Description
lnSCH16	Variable dépendante du modèle
Coef.	Coefficient β_{ki} devant la variable k de la régression i
Std. Err.	Écart-type du coefficient
t	Statistique t de Student
P>t	<i>p-value</i> de la statistique t
[95% Conf. Interval]	Intervalle de confiance du coefficient pour un niveau de confiance de 95%
Number of obs	Nombre d'observations
F(a, b)	Statistique F de Fisher pour un nombre de contraintes a et un degré de liberté b
Prob > F	<i>p-value</i> de la statistique F
R-squared	Coefficient de corrélation de la régression
Root MSE	Racine de la moyenne des écarts au carré

BIBLIOGRAPHIE

- Bergman, O., Ellingsen, T., Johannesson, M., & Svensson, C. (2010). Anchoring and cognitive ability. *Economics Letters*, *vol.107*(1), p.66-68.
- Blankenship, K. L., Wegener, D. T., Petty, R. E., Detweiler-Bedell, B., & Macy, C. L. (2008). Elaboration and consequences of anchored estimates: An attitudinal perspective on numerical anchoring. *Journal of Experimental Social Psychology*, *vol.44*(6), p.1465-1476.
- Bodenhausen, G. V., Gabriel, S., & Lineberger, M. (2000). Sadness and susceptibility to judgmental bias: The case of anchoring. *Psychological Science*, *vol.11*(4), p.320-323.
- Camerer, C. F. (2014). Behavioral economics. *Current Biology*, *vol.24*(18), R867-R871.
- Chapman, G. B., & Johnson, E. J. (1999). Anchoring, activation, and the construction of values. *Organizational behavior and human decision processes*, *vol.79*(2), p.115-153.
- Epley, N., & Gilovich, T. (2005). When effortful thinking influences judgmental anchoring: differential effects of forewarning and incentives on self-generated and externally provided anchors. *Journal of Behavioral Decision Making*, *vol.18*(3), p.199-212.
- McElroy, T., & Dowd, K. (2007). Susceptibility to anchoring effects: How openness-to-experience influences responses to anchoring cues. *Judgment and Decision Making*, *vol.2*(1), p.48-53
- Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.
- National Hockey League (2015), *Official rules 2015-2016*, article no 78, Repéré à <http://www.nhl.com/nhl/en/v3/ext/rules/2015-2016-Interactive-rulebook.pdf>
- Simon, H. A. (1955). A behavioral model of rational choice. *The quarterly journal of economics*, p.99-118.
- Slovic, P. (1967). The relative influence of probabilities and payoffs upon perceived risk of a gamble. *Psychonomic Science*, *vol.9*(4), p.223-224.
- Thaler, R. H. & Sunstein, C. R. (2008). *Nudge*. Penguin group.
- Wegener, D. T., Petty, R. E., Detweiler-Bedell, B. T., & Jarvis, W. B. G. (2001). Implications of attitude change theories for numerical anchoring: Anchor plausibility and the limits of anchor effectiveness. *Journal of Experimental Social Psychology*, *vol.37*(1), p.62-69.

Wilson, T. D., Houston, C. E., Etling, K. M., & Brekke, N. (1996). A new look at anchoring effects: basic anchoring and its antecedents. *Journal of Experimental Psychology: General*, vol.125(4), p.387-402