



L'impact des investissements publics en sport sur les médailles olympiques : une analyse économétrique.

Mémoire

Paul Blais-Morisset

Maîtrise en économie
Maître ès arts (M.A.)

Québec, Canada

© Paul Blais-Morisset, 2014

Résumé

Chaque année, une part significative des dépenses gouvernementales est consacrée aux sports de niveau professionnel et amateur. Selon certains analystes, de tels investissements pourraient constituer un facteur important du succès d'un pays aux Jeux olympiques (JO). Il devient ainsi pertinent d'évaluer si ces dépenses influencent réellement le nombre de médailles obtenues par les athlètes d'une nation. L'originalité de cette recherche est à la fois d'introduire l'investissement public dans le sport comme facteur déterminant du nombre de médailles remportées par un pays et de tenir compte des caractéristiques non observables d'un pays pouvant influencer son résultat olympique. Dans notre recherche, il est proposé de traiter l'hétérogénéité des pays participants aux JO à l'aide de modèles captant les caractéristiques invariantes des nations. De plus, l'inclusion de la variable d'investissement en sport crée un problème de simultanéité. Dans notre recherche, il est proposé de pallier ce problème à l'aide d'une variable instrumentale. Des résultats ont été obtenus à l'aide de modèles linéaires et non linéaires. Selon les spécifications du modèle retenues, un pays aux caractéristiques moyennes qui veut obtenir une médaille supplémentaire aux JO d'été doit, *ceteris paribus*, accroître son investissement de 35,6 à 106,7 millions de dollars.

Table des matières

| | |
|----------------------------------------------------------------|------------|
| Résumé | iii |
| Table des matières | v |
| Avant-propos | ix |
| Introduction | 1 |
| 1 État de la littérature | 5 |
| 1.1 Contribution historique | 5 |
| 1.2 Un modèle de référence : Bernard et Busse (2004) | 6 |
| 1.3 Le coût de la médaille | 9 |
| 2 Méthodologie | 11 |
| 2.1 Les données | 11 |
| 2.2 Spécification du modèle | 12 |
| 3 Analyse économétrique | 15 |
| 3.1 Données censurées | 15 |
| 3.2 Hétérogénéité des nations participantes | 17 |
| 3.3 Faible taille de l'échantillon | 18 |
| 3.4 Jeu relatif | 19 |
| 3.5 Endogénéité de l'investissement | 20 |
| 4 Résultats et analyse | 23 |
| 4.1 Spécifications linéaires | 23 |
| 4.2 Spécification non linéaires | 24 |
| 4.3 Estimation du coût d'une médaille | 26 |
| Conclusion | 29 |
| Bibliographie | 31 |
| A Définition des variables | 35 |
| B Résultats des spécifications | 37 |

Citius, Altius, Fortius.

Pierre de Coubertin
Devise olympique

Avant-propos

Il me serait impossible d'amorcer ce mémoire sans remercier les personnes sans qui ce travail n'aurait jamais vu le jour.

Mes plus sincères remerciements vont à mon directeur, Monsieur Bernard Fortin qui m'a pris sous son aile dans ce projet qui sortait de ses sphères traditionnelles de recherches. Ses précieux conseils, ainsi que sa rigueur m'ont permis de pousser ce projet à un niveau que j'aurais espéré autrement. Je vous suis très reconnaissant pour la confiance et pour le support que vous m'avez offert.

Un très grand merci doit aussi aller à mon codirecteur Monsieur Vincent Boucher qui a su apporter une vision novatrice à ce projet. Son aide et ses explications ont influencé grandement ma compréhension et se répercutent indubitablement dans l'ensemble de ce mémoire. La patience dont vous avez fait preuve à mon égard m'est très chère.

Je remercie sincèrement Luc Bissonnette dont la porte a toujours été ouverte à mes nombreuses questions et qui a passé d'innombrables heures à éclaircir les concepts les plus sombres de l'économétrie. Au nom de tous les étudiants qui ont profité de tes généreuses explications, je te remercie.

À tous mes amis économistes et à ceux provenant des autres milieux moins nobles, une grande partie de ce mémoire provient de vos commentaires et de vos encouragements. Les longues heures passées à bavarder avec vous de ce sujet qui me tenait à cœur ont apporté à ce mémoire une petite touche qui vous est propre. Un merci tout spécial à Marie-Noëlle Robitaille, Maria Adelaila Lopera et Simon Lalancette pour vos encouragements et votre aide.

Merci à mon père François et ma mère Marie. Merci, Marie, de m'avoir accompagné à travers l'ensemble de mes projets. François, merci de m'avoir soutenu dans mon cheminement d'économiste. Tu m'as assurément transmis cette curiosité intellectuelle qui t'est propre. Finalement, je remercie ma très populeuse famille pour votre soutien durant ces deux années de formation.

Introduction

La majorité des États affectent des sommes substantielles au développement de leurs athlètes, notamment l'élite vouée à se rendre aux Jeux olympiques, vitrine par excellence du sport dans le monde. Par exemple, pour l'année budgétaire 2011-2012, c'est près de 212,7 millions de dollars qui ont été investis par le gouvernement canadien dans la formation de l'élite sportive¹. Ces investissements dans le sport d'élite proviennent en grande partie de fonds publics. Ils permettent en principe de former des athlètes de pointe en mesure de rayonner sur la scène internationale et d'augmenter la fierté et le positionnement des nations qu'ils représentent.

Les Jeux olympiques d'été (JO), une compétition multidisciplinaire quadriennale, est l'évènement sportif le plus suivi au monde. Lors des Jeux de 2008 tenus à Beijing, on estime que pas moins de 4,7 milliards de personnes ont suivi les différentes épreuves². Les Jeux offrent une excellente opportunité pour les nations participantes de recevoir un rendement sur les sommes injectées annuellement dans la formation de sportifs d'élite. Lors de ces compétitions, les athlètes de tous les pays tentent de faire la fierté de leur pays en remportant les honneurs dans l'une des 38 disciplines olympiques. Le nombre de pays et le nombre d'athlètes participant aux Olympiques n'ont cessé de progresser au fil du temps. Lors des premiers JO modernes en 1896, 241 sportifs masculins de 14 nationalités différentes avaient pris part à l'évènement. Une centaine d'années plus tard, aux JO de Londres en 2012, 10568 athlètes (5892 hommes et 4676 femmes) ont représenté les 204 pays qui se sont battus pour remporter les grands honneurs. Des normes différentes et adaptées aux épreuves et aux disciplines permettront aux athlètes de se qualifier aux Olympiques. De plus, les nations qui sont désavantagées au niveau de leurs richesses ou de leurs populations se verront généralement offrir des normes de qualifications moins sévères afin que ceux-ci puissent tout de même envoyer des athlètes aux JO. Conséquemment, le nombre d'Olympiens d'un pays va différer d'une olympiade à l'autre, en plus de varier entre les nations. Il est de coutume de publier un tableau des résultats des médailles acquises par chacun des pays lors d'une édition des Jeux olympiques. Ce tableau

1. Voir *Patrimoine Canadien* (2012). Ce montant provient du programme de soutien aux sports, du programme de soutien aux athlètes et du programme d'accueil d'évènement. Il exclut les investissements canadiens pour les infrastructures sportives.

2. Cette statistique provient de la compagnie d'analyse de marché *The Nielson Company* (2008).

n'a rien d'officiel, mais les résultats cumulatifs des nations participantes y sont représentés. Le pays terminant au sommet du tableau est, officieusement du moins, couronné vainqueur de cette édition des Olympiques.

L'influence du financement public de l'élite sportive sur le positionnement des pays aux JO est un enjeu financier et social d'une grande importance et qui fait l'objet de nombreux débats publics, notamment lorsqu'un pays se voit déclassé. Il s'agit d'une question pour laquelle l'analyse économique peut tenter d'apporter un éclairage, puisque les réponses à ces questions sont nombreuses et souvent assez mal documentées. Par exemple, pourquoi l'Éthiopie, pays pauvre, a-t-elle remporté plus de médailles d'or que le Canada aux Jeux olympiques de Londres? Pourquoi l'Inde remporte si peu de médailles malgré sa forte population (Krishna et Haglund, 2008)? Quelle est la corrélation entre le niveau d'investissement public et le nombre de médailles remportées aux JO? D'un point de vue économique, serait-il préférable de concentrer les investissements dans certaines disciplines afin de maximiser le nombre de médailles obtenues ou doit-on chercher plutôt à répartir ces sommes dans un grand nombre de disciplines comme le fait le Canada, mais dans une perspective de gains à plus long terme? Ces questions et bien d'autres méritent l'attention des économistes.

Dans le présent mémoire, nous cherchons à déterminer le niveau d'investissement moyen qu'un pays doit engager pour améliorer son positionnement global aux JO. Notre étude présente une généralisation du modèle de Bernard et Busse (2004). Leur modélisation sert souvent de référence pour tenter de prédire, ou mieux pour expliquer, les succès olympiques d'une nation relativement à une autre. Les principaux facteurs de succès pour un pays qui sont retenus par le modèle sont : le PIB par tête, sa population, le fait d'être l'hôte de l'évènement ou encore d'être un pays communiste ou postcommuniste. Nous nous démarquons de ces derniers par l'ajout du financement public en sport comme variable explicative du succès olympique. Cependant, un plus grand investissement public risque vraisemblablement d'augmenter le nombre de médailles, et les pays avec un nombre de médailles plus élevé sont également susceptibles d'encourager davantage le sport en augmentant son financement. Pour palier ce biais de simultanéité, il est d'usage de se servir d'une variable instrumentale (VI). L'investissement public en sport par habitant retardé d'une période est proposé comme instrument pour corriger ce problème d'endogénéité. De plus, une attention toute particulière est apportée quant au traitement de la culture sportive dans notre modélisation.

La suite de ce mémoire est organisée en quatre chapitres. Le premier présente les variables et les modélisations du succès olympique proposées dans la littérature. Le second décrit la base de données employée et discute du modèle employé. Le troisième chapitre discute des différents problèmes rencontrés suite à l'inclusion de l'investissement public dans le sport

comme variable explicative du succès olympique et présente les spécifications qui permettent de les surmonter. La quatrième section présente les résultats obtenus et concentre l'analyse sur les spécifications tenant compte de l'hétérogénéité des pays. Vient ensuite la conclusion qui rappelle les principaux résultats obtenus et qui présente les avenues futures de recherche.

Chapitre 1

État de la littérature

Bien que le rêve olympique commence dans la tête de chaque athlète, les nations considèrent généralement qu'il existe plusieurs externalités positives au fait d'être pourvues de sportifs d'élite. Une population en santé est généralement considérée comme l'objectif principal et « avoué » du financement sportif, mais ce n'est pas tout. Le développement d'une élite sportive est aussi recherché par les nations pour la « gloire » qu'elles en tirent sur le plan international et la « cohésion » qui peut en découler sur le plan interne. De telles raisons expliquent pourquoi les pays investissent autant dans le sport d'élite.

1.1 Contribution historique

Les déterminants sociaux et économiques qui expliquent les performances aux Olympiques sont étudiés réellement depuis les années 1970. Les premiers à chercher des explications à la réussite des pays aux JO sont des sociologues. Ball (1972) a identifié 21 variables permettant de différencier les pays dominants des autres pays participants aux Jeux olympiques. De l'ensemble de ces variables, le chercheur fait ressortir deux déterminants principaux du succès : le produit intérieur brut par tête et l'importance de la population. Il insiste aussi sur la pertinence de pousser la recherche sur le fait d'appartenir ou non à un régime communiste.

Levine (1974) a repris les résultats obtenus par Ball et tente de les expliquer et de les quantifier. Dans son article, il estime que le PIB par personne et la population expliquent à chacun d'eux près de 40 % des résultats olympiques. En fait, un plus fort PIB par habitant signifie une population plus riche disposant de plus de moyens pour investir dans les infrastructures coûteuses nécessaires à la pratique du sport. De plus, elle aura ensuite les moyens de les entretenir, d'encourager sa population à la pratique du sport et d'acquérir les entraîneurs et les infrastructures de pointes nécessaires pour que son équipe nationale puisse être concurrentielle sur la scène internationale. De plus, une population plus riche a aussi plus de temps et de ressources à octroyer au loisir, ce qui favorise la pratique sportive de la population et donc

l'émergence d'athlètes de haut niveau. Le bassin de population, pour sa part, accroît la possibilité pour un pays de voir émerger l'un des leurs ayant les caractéristiques physiques et les habiletés idéales à la pratique d'un sport donné. L'article de Levine termine en mentionnant qu'il serait pertinent de tester la signification statistique de l'avantage d'être l'hôte des Jeux, ce que feront ultérieurement d'autres auteurs.

Beaucoup d'articles spécialisés reprennent les facteurs du PIB et de la population comme composantes pour prédire les résultats des Jeux. Comme ces deux variables expliquent à elles seules une part énorme du succès et qu'elles reposent sur des données très accessibles, beaucoup de chercheurs dans le sport ne prennent en considération que ces deux facteurs dans leur modélisation¹. Par ailleurs, Levine mentionne qu'il serait tout de même préférable de se fier à des indices d'utilisation des ressources directes, comme le montant déboursé pour le sport et la proportion du produit intérieur brut affectée par le pays au soutien des athlètes. Ces informations n'étant pas accessibles à son époque, la mesure du PIB par habitant demeure la meilleure approximation dans les circonstances.

1.2 Un modèle de référence : Bernard et Busse (2004)

Les économistes ont généralement utilisé des méthodes économétriques afin de tenter d'identifier les déterminants de la réussite aux olympiades. Les chercheurs ont utilisé des modèles assez similaires, tout en jouant sur les facteurs explicatifs (ajout de variables, séparation d'une variable en sous-ensembles, *etc.*) ou en modifiant les méthodes de régression utilisées. Le modèle classique le plus repris dans la littérature est celui de Bernard et Busse (2004). Ce dernier utilise le modèle Tobit afin de prédire et expliquer le succès des nations aux JO. Dans ce mémoire, nous nous basons sur ce modèle et tentons de l'améliorer par l'ajout d'une variable explicative. Il devient donc pertinent de bien présenter le modèle de Bernard et Busse afin de bien montrer les forces de celui-ci tout en soulignant les faiblesses qui lui sont associées.

L'inspiration

Condon *et al.* (1999) ont publié l'une des premières études sur les déterminants socio-économiques avec l'objectif de prédire le succès d'une nation aux JO. Les auteurs utilisent un système de pondération accordant des valeurs aux huit premières positions afin d'agrèger leurs données. Pour atteindre leur but, ils modélisent les performances en utilisant des réseaux de neurones artificiels. Cette méthode de modélisation statistique est considérée comme non linéaire et

1. C'est notamment le cas des études examinant l'efficacité de productions des médailles par les pays, selon ces deux variables, à l'aide de la méthode du *Data Envelopment Analysis* (Li *et al.*, 2008; Lozano *et al.*, 2002; Wu *et al.*, 2010).

permet d'associer des intrants à un extrant par des méthodes d'apprentissage. La performance des résultats de leur modèle est ensuite comparée à celle obtenue par une simple régression linéaire. Sur les 24 réseaux de neurones créés par les chercheurs, seulement l'un d'entre eux a une moyenne des erreurs en absolu inférieure à la régression par moindres carrés ordinaires (MCO). Malgré la faiblesse des résultats obtenus par les chercheurs, les deux facteurs explicatifs du succès de Ball et Levine semblent toujours être cohérents dans l'explication du succès olympique.

Accueillir les Jeux à domicile influence de façon significative la performance globale du pays hôte. Clarke (2000) a déterminé que les pays voient tripler leur pourcentage d'acquisition de médailles lorsqu'ils reçoivent les JO. Deux explications peuvent aider à comprendre ce phénomène. Premièrement, il y a les avantages dont bénéficient directement les athlètes du pays hôte : moins de fatigue due au transport, connaissance approfondie du terrain où se déroule la compétition, motivation provenant des encouragements des amateurs locaux qui remplissent les gradins, etc. Deuxièmement, les pays qui reçoivent les jeux bénéficient aussi du privilège accordé par le CIO de pouvoir inscrire plus de concurrents aux épreuves en plus d'avoir une place acquise dans chacune des disciplines. Bien que les athlètes qui sont ajoutés à la délégation nationale ne soient habituellement pas les meilleurs de la délégation, il se peut toujours qu'un nombre minime d'entre eux excellent au point d'obtenir une médaille. L'effet d'être l'hôte des jeux vient inévitablement augmenter le financement du sport d'élite et cela est particulièrement le cas dans les années qui précèdent l'événement. En effet, les futurs pays hôte ont généralement tendance à dépenser davantage quelques années avant de recevoir les Jeux afin d'exceller plus qu'à l'habitude à l'édition qui est disputée à domicile. L'omission d'une variable binaire déterminant le pays hôte d'une édition des JO viendrait ainsi biaiser nos estimateurs si celle-ci est corrélée avec les variables explicatives retenues dans le modèle.

Johnson et Ali (2000) ont proposé deux modèles qui analysent dans un premier temps la décision de participation d'un pays aux Jeux olympiques en nombre d'athlètes et, dans une seconde modélisation, le nombre de médailles qu'un pays devrait obtenir selon son profil. Dans les deux cas, ils se servent de variables explicatives presque identiques. Les auteurs se démarquent en incluant à leur modélisation une dizaine de variables binaires, afin d'évaluer quels sont les régimes politiques influençant significativement les résultats des Olympiades d'été et d'hiver. Ils obtiennent que pour les Jeux d'été, seuls les pays communistes et postcommunistes retirent un avantage réel de leur statut politique. Tcha (2004) compare la fonction de production de médailles du Bloc de l'Est à celui de l'Ouest afin d'expliquer la facilité des pays soviétiques à récolter des médailles. Il montre notamment que la valeur plus élevée accordée aux médailles ainsi que l'organisation au niveau sportif des pays du Bloc de l'Est leur permettent de rem-

porter à moindres coûts des médailles olympiques². C'est pourquoi, si l'on souhaite obtenir un estimateur non biaisé représentant le réel impact de l'investissement dans le sport, il est primordial d'intégrer à notre modèle des variables binaires explicatives associées spécifiquement aux régimes soviétiques et postsoviétiques.

1.2.1 La contribution

L'article de Bernard et Busse (2004) s'est positionné comme un classique dans la littérature. Cette étude rassemble et explique les différents apports qui ont été réalisés par le passé. Les auteurs veulent examiner les déterminants du succès olympique en utilisant les variables qui avaient été explicitées. Pour ce faire, les chercheurs décident d'utiliser un modèle Tobit. L'avantage de ce modèle est qu'il gère facilement le problème de données censurées dans leur modèle. En fait, comme la moitié des nations participantes n'obtiennent aucune médaille au JO en respect pour leurs caractéristiques, il faut gérer ce groupe de pays afin de ne pas biaiser la courbe estimée. Pour prouver leur point, les deux auteurs prennent soin de comparer les résultats du modèle Tobit à ceux obtenus par les moindres carrés ordinaires. L'apport principal de l'article est de montrer que ce modèle parvient à de meilleurs résultats que la régression classique par MCO. Cet avantage du modèle Tobit vient sans aucun doute de sa capacité de gérer la censure provenant de l'absence de médaille pour une majorité de pays.

De plus, Bernard et Busse (2004) transforme la variable dépendante afin d'obtenir un ratio du nombre de médaille remportée sur l'ensemble des médailles qui étaient disponibles lors de cette même édition des JO. Cette spécification de la variable permet notamment de considérer la croissance dans le temps du nombre d'épreuve olympique et conséquemment du total de médailles décernées. Ainsi, un pays remporte une part des médailles totales et celle-ci peut se modifier selon son amélioration vis-à-vis les nations adverses.

Plusieurs auteurs ont repris le modèle de Bernard et Busse (2004) en adaptant certains éléments afin d'en augmenter la performance. Certains comme Andreff *et al.* (2008) vont tenter de capter l'effet de la culture sportive en incluant à la modélisation des variables séparant les pays selon leurs différences régionales. D'autres vont utiliser un modèle Tobit très similaire dans l'objectif d'expliquer et de prédire le succès des nations aux Jeux olympiques d'hiver (Andreff et Andreff, 2012; Johnson et Ali, 2004; Pfau, 2006). Enfin, Buts *et al.* (2011) convertissent ce modèle classique afin d'expliquer le succès des pays aux Jeux paralympiques d'été. Le modèle de Bernard et Busse (2004) s'est donc placé comme une référence au niveau de la modélisation du succès olympique pour des raisons d'efficacité et de simplicité.

2. Poupaux (2006) explore plus amplement sur l'avantage comparatif que possèdent ces pays.

1.2.2 Les limites du modèle

Bernard et Busse (2004) ne tiennent pas compte de l'investissement réel des pays mais seulement du PIB per capita. Selon nous, l'ajout d'une variable captant les politiques de soutien aux sportifs propres à chaque pays permettrait d'augmenter le pouvoir explicatif du modèle en plus d'être bien plus représentatif de l'effort déployé par les États pour décrocher davantage de médailles olympiques. Une justification naturelle de l'ajout de cette variable est qu'elle capte le réel effort réalisé par les pays afin d'exceller dans le sport. Il devient beaucoup plus intéressant de regarder l'effet du financement du sport que d'observer uniquement le PIB dans un modèle. Le PIB par habitant reflète le niveau de richesse du pays et un niveau de vie permettant la pratique de sport amateur et de calibre supérieur. Par contre, il est très difficile d'influencer directement sur la croissance du PIB afin d'augmenter son nombre de médailles. Les nations ont davantage de facilité à financer directement de nouvelles politiques favorisant la pratique sportive. L'investissement dans le sport capte, non seulement un niveau de richesse supérieur provenant de la plus grande capacité de financer le sport élite et amateur des pays les plus riches, mais il saisit aussi la volonté de chacun d'investir ou non dans le sport.

1.3 Le coût de la médaille

Johnson et Ali (2004) sont les premiers à présenter des estimations sur le coût d'une médaille olympique pour un pays. Les auteurs reprennent la modélisation du succès obtenu par les pays aux JO de Johnson et Ali (2000). Selon les chercheurs, la hausse du PIB par habitant moyen nécessaire à un pays pour obtenir une victoire supplémentaire reflète la valeur d'une médaille olympique. Le problème avec cette conclusion, c'est qu'elle exclut la possibilité pour un pays de se munir de politiques d'investissements concrètes et ciblées afin d'améliorer son classement. Ceci est, donc, fort peu utile pour les décideurs publics. Dans ce mémoire, nous voulons plutôt estimer le coût de la médaille résultant du financement direct de l'État dans le sport.

Une seule étude réfère à l'investissement direct dans le sport comme déterminant du nombre de médailles obtenues par les nations aux Olympiques. Flores *et al.* (2008) tentent d'améliorer le modèle classique par l'ajout du logarithme de l'investissement en sport par habitant. Cette addition permet d'estimer à près de 40 millions d'euros (52 millions USD) le financement public moyen qu'un pays doit réaliser pour obtenir une médaille d'or supplémentaire aux Jeux olympiques d'été. Cependant un problème d'endogénéité, dont nous discutons plus loin, n'a pas été traité par ces auteurs suite à l'inclusion d'une telle variable.

Chapitre 2

Méthodologie

2.1 Les données

À notre connaissance, il n'existe aucune base de données recensant directement le financement public accordé aux sports pour l'ensemble des pays. Cependant, les Nations Unies ont développé un système de classification des dépenses gouvernementales appelé *Classification of Functions of the Government* (COFOG). À la manière de Flores *et al.* (2008), la catégorie identifiant les investissements gouvernementaux en *loisir, culture et affaires religieuses*¹ est retenue comme variable d'approximation des dépenses gouvernementales allouées aux sports. Quatre institutions différentes publient de l'information conformément à ce système de classification et ont permis de créer notre variable d'investissement sportif. La source principale de données sur cette variable d'approximation provient du Fonds Monétaire International (FMI) qui publie annuellement le *Government Finance Statistics* (GFS). Ce document recense les données des différentes catégories du COFOG pour une majorité de pays, et ce, depuis 1990. Du reste, Eurostat, la division statistique des Nations-Unis, ainsi que par l'Organisation de Coopération et de Développement Économiques (OCDE) donnent également accès en ligne à cette variable d'approximation du financement gouvernemental du sport. Ces trois dernières sources ont permis d'augmenter l'échantillon de pays recensés par la base de données. Malgré cet effort pour obtenir une base de données complète, l'information n'est pas disponible pour l'ensemble des pays participants aux Jeux olympiques d'été. Nous émettons l'hypothèse que les pays absents de l'échantillon sont aléatoirement manquants. Cependant, il est possible que l'hypothèse soit fautive si les pays manquants sont majoritairement ceux qui ne remportent aucune médaille et conséquemment ceux dont les investissements publics en sport sont moins importants. Dans ce cas précis, notre échantillon contiendrait davantage de nations riches et médaillées que la réalité. Les paramètres estimés seraient alors biaisés à la hausse et l'investissement d'un gouvernement afin de remporter une médaille serait sous-estimé. Afin de compléter les données d'investissement manquantes dans la série temporelle d'un pays, la méthode d'in-

1. *Recreational, Cultural and Religious Affairs.*

terpolation cubique spline a été retenue. Les pays dont plus de 8% des données avaient été générées par cette méthode ont été retirés de la base de données. Au total, le financement public du sport de 53 pays et couvrant 6 éditions des Jeux olympiques d'été, soit ceux de 1992-1996-2000-2004-2008 et de 2012, permettent d'obtenir les 318 observations utilisées dans ce mémoire. La variable reflétant l'investissement public en sport est convertie en devise commune puis actualisée à l'aide de la *Penn World Table*². Le financement gouvernemental du sport est alors exprimé en dollars américains courants de 2012 à parité du pouvoir d'achat de chacune des devises. Finalement, la variable explicative est divisée par le nombre d'habitants du pays de manière à considérer un montant investi par individu pour la pratique du sport.

Le produit intérieur brut par habitant, ainsi que l'estimation de la population des pays retenus sont disponibles sur le site de la *Data World Bank*³. Le nombre de médailles remportées par les pays dans chacune des six éditions des JO d'été couvertes et l'information sur les pays hôtes sont disponibles sur le site du Comité International Olympique (CIO). Les variables indiquant si le régime politique est communiste ou postcommuniste ont été incluses sous la forme de variables dichotomiques. Toutefois, il faut noter que la Chine est le seul État communiste de notre échantillon.

2.2 Spécification du modèle

Bernard et Busse (2004) définissent la variable dépendante comme étant la part du nombre de médailles remportées par un pays sur l'ensemble des médailles qui étaient en jeu à cette édition des JO d'été. Les modèles de ce mémoire utiliseront ce ratio afin de considérer la croissance du nombre de médailles à travers les éditions des Olympiques. Contrairement à ces derniers, l'idée de spécifier des variables en proportion est étendue aux variables explicatives. Ainsi, la population, le produit intérieur brut par habitant et les sommes gouvernementales injectées dans le sport par tête sont tous trois spécifiés en proportion de la même variable au niveau mondial. Cette construction de variables permet de considérer l'effet des nations adverses sur son propre résultat olympique. Par exemple, l'impact de la population sur la performance olympique peut s'atténuer même si son bassin d'athlètes ne connaît aucun changement. À population constante pour un pays, la croissance du bassin d'athlète mondial devrait avoir un impact négatif sur le succès olympique de la nation. Cette idée de ratio est particulièrement intéressante avec notre variable d'investissement puisqu'elle représente en quelque sorte l'effort qu'un pays déploie pour se démarquer de ses adversaires sur la scène sportive internationale. De façon similaire à la population, cette nouvelle variable rend compte de l'impact d'une hausse ou d'une baisse l'investissement des autres pays sur son propre résultat olympique.

2. Disponible au : <http://www.rug.nl/research/ggdc/data/penn-world-table>.

3. <http://data.worldbank.org/>

Ainsi, lorsque les dépenses en sport au niveau mondial double, un pays qui double lui aussi son investissement ne devrait pas être en mesure d'obtenir davantage de médailles. En fait, la spécification de la variable permet d'obtenir un jeu entre les pays aboutissant à un équilibre de Nash. Quel montant une nation décidera-t-elle d'allouer aux activités sportives considérant ses a priori sur le financement qui sera réalisé par ses adversaires ?

Finalement, une transformation logarithmique des trois variables explicatives continues est utilisée afin de lisser les relations. Donc, nos modélisations retiendront l'ensemble des variables retrouvées dans la littérature précédemment décrite. Par contre, les dépenses publiques en sport par habitant, le produit intérieur brut par tête et la population seront introduits en proportion.

Soit M_{it} qui est défini comme le rapport entre le nombre de médailles remportées par un pays i et la somme des médailles gagnées lors d'une édition t des JO d'été. En supposant pour le moment que M_{it} est strictement positif, le modèle liant la performance olympique d'un pays à ses déterminants peut s'écrire comme suit :

$$M_{it} = \beta_0 + \beta_1 INV_{it} + \beta_2 PIB_{it} + \beta_3 POP_{it} + \beta_4 COM_i + \beta_5 PCOM_i + \beta_6 H\hat{O}TE_{it} + \alpha_i + \epsilon_{it}$$

où INV_{it} , PIB_{it} et POP_{it} sont des rapports entre la valeur propre aux pays et celle de l'agrégation de ces derniers pour une période donnée. Respectivement, INV_{it} représente le ratio d'investissement public en sport par tête réalisé précédemment aux Jeux, PIB_{it} rapporte l'information sur le logarithme produit intérieur brut par habitant et POP_{it} est le logarithme de la population. COM_i , $PCOM_i$ et $H\hat{O}TE_{it}$ sont des variables dichotomiques indiquant si le pays est considéré comme communiste, post communiste ou est l'hôte des Jeux. L'annexe A présente la construction détaillée de chacune des variables ainsi que certaines statistiques descriptives qui leur sont associées. Le paramètre α_i représente les effets invariables dans le temps qui influencent le succès olympique et qui sont propres à chaque pays. La section suivante propose deux techniques pour considérer ce paramètre dans nos spécifications du modèle.

Chapitre 3

Analyse économétrique

Cinq problèmes économétriques surviennent suite à l'ajout de l'investissement gouvernemental en sport par habitant à la modélisation du succès olympique. Ce chapitre discute de ces derniers et propose des modèles afin de résoudre chacun d'eux. L'annexe B présente un tableau liant ces problèmes aux modèles qui permettent leurs résolutions. Il est à noter qu'aucune modélisation ne permet de résoudre l'ensemble de ceux-ci. En revanche, nous présentons une analyse de sensibilité des résultats aux différentes approches retenues.

3.1 Données censurées

Dans les six olympiades traitées, en moyenne 17 pays de notre échantillon n'obtiennent aucune médaille lors d'une édition des JO. Cependant, lorsqu'elle est nulle, la variable dépendante n'a pas la même signification pour l'ensemble des pays étudiés. Pour une édition des JO donnée, certains devraient augmenter grandement la taille de leurs populations, leurs niveaux de richesse ainsi que leurs investissements en sport afin de remporter une première médaille. D'autres, par contre, se trouvent sur la frontière entre le succès olympique et l'absence de médaille et l'effort marginal à déployer pour obtenir leur premier succès est très faible. Les données sont censurées en raison du grand nombre de valeurs nulles observées et de l'impossibilité d'observer un compte négatif de médailles. Dans le cas présent, la censure biaise négativement les estimateurs obtenus à l'aide d'une régression linéaire. Il est nécessaire de considérer un modèle latent qui exprime la relation entre le ratio du nombre de médailles désirées M_{it}^* par un pays et le vecteur de variables explicatives \mathbf{x}_{it} :

$$M_{it}^* = \mathbf{x}_{it}'\boldsymbol{\beta} + \epsilon_i, \quad \epsilon_i \sim N(0, \sigma^2).$$

Cependant, la variable dépendante observée M_{it} , est égale à la valeur obtenue par le modèle linéaire latent seulement lorsque le nombre de médailles est strictement positif.

$$M_{it} = \begin{cases} M_{it}^* & \text{si } M_{it}^* > 0, \\ 0, & \text{si } M_{it}^* \leq 0. \end{cases}$$

Bernard et Busse (2004) proposent l'utilisation d'un modèle Tobit afin de considérer la censure du résultat olympique à $M_{it} = 0$. L'avantage du modèle Tobit est qu'il permet une modélisation simultanée de deux parties. La première identifie si la nation participante obtiendra un compte positif ou nul de médailles selon ses caractéristiques. La deuxième partie est une régression linéaire identifiant le nombre de médailles qui seront remportées si et seulement si le pays s'est qualifié, dans un premier temps, pour être médaillé. Ce travail utilise trois types de modélisation Tobit dans le but de rendre compte de la censure de nos données ; le modèle Tobit classique, instrumenté et panel. Toutefois, le modèle Tobit impose une modélisation identique de la probabilité d'obtenir une première médaille et du compte de médailles remportées.

En plus de ces spécifications Tobit, une procédure d'estimation en deux étapes d'Heckman est proposée afin de capter la non linéarité de la relation à $M_{it} = 0$. Cette dernière est spécifiée d'abord par une équation de sélection inobservée puis par une régression pour les résultats strictement positifs ce qui permet d'obtenir des estimations différenciées pour chacune de ces deux parties. En employant le même vecteur de variables explicatives pour les deux étapes du processus, le modèle remplit les conditions minimales d'identification¹. L'équation de sélection prend la forme :

$$\begin{aligned} z^* &= \mathbf{x}'_{it}\boldsymbol{\gamma} + u, \quad u \sim N(0,1) \\ z &= 1, \quad \text{si } z^* > 0 \\ z &= 0, \quad \text{si } z^* \leq 0. \end{aligned}$$

Tandis que la partie observée M_{it} s'écrit pour les résultats sélectionnés (si et seulement si $z = 1$) :

$$M_{it} = \mathbf{x}'_{it}\boldsymbol{\beta} + \lambda + \epsilon, \quad \epsilon \sim N(0, \sigma^2).$$

Cette spécification offre l'avantage de considérer les effets qui sont propres à chaque pays à travers un paramètre λ déterminé à l'aide des résidus de l'équation de sélection. Ce dernier correspond à l'inverse du ratio de Mills et est ajouté comme variable explicative dans l'équation du modèle observé afin de contrôler pour le biais de sélection des pays médaillés. Il est à

1. Il est généralement recommandé d'utiliser une variable d'exclusion qui est présente dans l'équation de sélection, mais qui ne se retrouve pas dans la seconde étape.

noter que le modèle de sélection d'Heckman est à considérer avec précaution puisqu'il repose sur l'hypothèse forte de normalité des résidus.

3.2 Hétérogénéité des nations participantes

Certains pays possèdent un avantage comparatif quant à l'obtention de médailles olympiques. En fait, pour un pays donné, des caractéristiques inobservables et invariantes dans le temps peuvent expliquer son succès aux Jeux olympiques. Dans le langage courant, ces attributs se voient souvent désignés par le nom de "culture sportive" d'un pays. Bien que le concept puisse sembler flou et vaste, la culture sportive représente les caractéristiques nationales qui facilitent la récolte de médaille olympique. L'histoire olympique et la spécialisation sportive d'un pays, l'intensité de la participation d'une nation, de même que la génétique avantageuse de certains peuples, sont de bons exemples de ces avantages. On peut modéliser ce phénomène grâce à un modèle à erreurs composées :

$$M_{it} = \mathbf{x}'_{it}\boldsymbol{\beta} + \underbrace{\alpha_i + \epsilon_{it}}_{\tilde{\epsilon}_{it}}$$

Toutefois, en supposant que $E(\epsilon_{it}|\mathbf{x}_{it}) = 0$, les estimateurs obtenus par les MCO seront biaisés dans la mesure où le terme d'erreur invariant α_i est corrélé avec le vecteur de variables explicatives. Il est possible de considérer l'avantage invariant d'une nation de deux façons distinctes : soit par les effets fixes ou par les effets aléatoires.

Pour les modélisations avec données groupées², il est proposé d'intégrer une variable indicatrice (d_i) pour chaque pays et d'estimer les paramètres κ_i qui captent l'hétérogénéité des nations face aux gains de médailles olympiques.

$$M_{it} = \mathbf{x}'_{it}\boldsymbol{\beta} + \sum_{i=1}^{53} \kappa_i d_i + \epsilon_{it}$$

Cependant, cette prise en considération du terme d'erreur invariant dans le temps par la méthode des effets fixes nous met face à face avec le problème des paramètres incidents. Ce dernier survient lorsque l'estimation se fait sur une période fixe et que le nombre de paramètres à estimer croît considérablement. Pour les modèles linéaires, ce problème se répercute en la non convergence des paramètres κ_i captant l'avantage qu'ont certains pays à remporter des

2. *Pooled data.*

médailles olympiques. Ces spécifications linéaires obtiennent tout de même des estimateurs convergents des paramètres associés aux autres variables explicatives du modèle. Par contre, pour les modèles non linéaires (comme pour le modèle Tobit par exemple) l'ensemble des paramètres estimés ne convergent pas. Par conséquent, les valeurs obtenues par cette méthode sont moins propices à l'analyse. Pour estimer les effets fixes par cette méthode, la variable indicatrice de la Tunisie, un pays aux caractéristiques moyennes, a été retirée. Notons que dans l'échantillon seul un pays est communiste, la Chine, et que par conséquent cette variable est colinéaire avec la variable dichotomique indiquant ce pays. Par convention, la variable indiquant le pays communiste est conservée dans les tableaux des résultats. Néanmoins, il faut noter que ce paramètre estimé représente l'effet joint d'être un pays communiste et des effets invariants dans le temps propre à la culture chinoise sans possibilité de distinguer chacun de ces deux paramètres.

Une seconde hypothèse peut être émise par rapport au terme d'erreur invariant. La prise en considération par effet aléatoire impose que l'effet spécifique à un pays ne soit pas corrélé avec les variables explicatives du modèle ($E(\alpha_i|\mathbf{x}) = 0$). Cette spécification du modèle par effets aléatoires est moins parcimonieuse puisqu'elle ne force pas l'estimation de paramètres pour chacun des pays participants aux JO. Le problème de paramètres incidents que nous posait l'estimation des modèles d'effets fixes est ainsi évité. Ainsi, si le terme d'erreur invariant α_i et des variables explicatives ne sont pas corrélés, les effets aléatoires obtiennent des estimateurs convergents et efficaces.

Nous avons effectué un test de Hausman qui permet de départager entre le choix de ces deux méthodes. L'hypothèse nulle de ce dernier est que les coefficients estimés par l'estimateur efficace des effets aléatoires sont les mêmes que ceux estimés par l'estimateur des effets fixes ($H_0 : \beta_{RE} = \beta_{FE}$). Ainsi, si l'hypothèse nulle est rejetée, les estimateurs obtenus à l'aide des effets fixes seront préférés. Cependant, l'impossibilité de rejeter que les coefficients soient similaires rend préférable l'utilisation des effets aléatoires puisque ceux-ci obtiennent des estimateurs convergents, mais davantage efficaces. Dans ce cas, l'orthogonalité du terme d'erreur invariant et des variables explicatives suppose que la réalisation d'un modèle avec données groupées plutôt que des données en panel permet d'obtenir des estimateurs convergents et efficaces.

3.3 Faible taille de l'échantillon

Comme il a été discuté à la section 2, l'information sur l'investissement a été recensée pour 53 pays sur une période couvrant 6 olympiades. Ce petit nombre de périodes couvertes peut

causer un problème de convergence pour les estimateurs de nos modèles. Par exemple, les estimateurs des variables dichotomiques ajoutées aux modélisations pour capter l'hétérogénéité des pays n'ont que six informations afin d'être adéquatement estimés. Conséquemment, les propriétés asymptotiques sur des grands échantillons ne sont pas applicables et un doute plane sur la convergence des paramètres estimés.

3.4 Jeu relatif

À la différence des auteurs précédents, nous avons construit les variables explicatives continues de nos spécifications en proportion de la même variable au niveau mondial afin de considérer l'effet des pays concurrents sur son résultat olympique. Au numérateur se trouve l'information sur le pays quant à chacune de ces variables continues. Au dénominateur se retrouve la somme des performances de l'ensemble des pays pour ces variables. Soit $continues_{it}$ décrivant l'information suivante : la population, le produit intérieur brut par tête ou l'investissement gouvernemental en sport par habitant d'un pays i à l'édition t des JO, alors le logarithme du ratio suivant est utilisé comme variable explicative continue :

$$\left(\frac{continues_{it}}{\sum_i continues_{it}} \right).$$

Cette manière de faire permet aux modèles d'expliquer qu'une augmentation pour un pays donné peut être accentuée ou compensée par une fluctuation de cette information pour ses adversaires. Un pays dont le bassin d'athlètes se voit grandit obtient donc davantage de médailles olympiques en autant que la taille de la population de ces compétiteurs n'augmentent pas d'autant. Cette construction des variables permet notamment d'expliquer pourquoi certains pays comme la Chine, dont la richesse relative a explosé vers la fin du dernier siècle, ont amélioré considérablement leur positionnement sur la scène sportive internationale par rapport à d'autres dont la croissance économique a été plus lente.

Cette construction est tout particulièrement intéressante lorsque l'on s'attarde à la variable d'investissement public par habitant réalisé par un pays. En effet, si cette dernière représente l'effort déployé par un pays afin de rayonner aux JO, alors il devient important de regarder si ses investissements se démarquent de ceux engagés par ses opposants. Bien qu'un pays détermine un montant à investir dans les performances sportives, il peut tout de même se faire déclasser aux Olympiques si le montant n'est pas suffisamment élevé par rapport à ses adversaires.

3.5 Endogénéité de l'investissement

L'un des objectifs de ce mémoire est d'inclure le logarithme des dépenses dans le sport par habitant et de corriger l'estimation de sa source probable d'endogénéité. En fait, un biais dû à l'ajout de cette variable explicative rend les estimateurs non convergents. Le problème en cause provient de la simultanéité entre la variable identifiant les dépenses publiques et le nombre de médailles remportées. L'investissement dans le sport amène fort probablement un succès plus grand aux Olympiques. En contrepartie, un mauvais résultat aux JO peut inciter un pays à investir dans l'intention d'améliorer son résultat. La relation causale est donc inversible et il est nécessaire de corriger ce biais de simultanéité dans le but d'obtenir un estimateur convergent des dépenses gouvernementales en sport.

La modélisation à l'aide d'une variable instrumentale est proposée afin de résoudre le biais de simultanéité et d'isoler le sens de la relation. Un bon instrument se doit de répondre à deux caractéristiques bien précises. Dans un premier temps, pour que l'instrument (Z) soit valide il se doit d'être exogène au modèle structurel. Ainsi, cette variable ne doit pas être corrélée avec le terme d'erreur (u) du modèle.

$$\text{corr}(Z,u) = 0$$

En outre, il est nécessaire que l'instrument soit informatif, c'est-à-dire fortement corrélé avec la variable explicative endogène. Dans le cas de notre modélisation, l'instrument proposé ne doit pas expliquer directement le succès olympique actuel d'une nation hormis par son influence sur la variable endogène, l'investissement public dans les activités sportives.

Dans ce mémoire, le financement public du sport retardé d'une édition des JO est suggéré comme variable instrumentale. La justification de cet instrument est a priori très intuitive. Le niveau de financement public passé explique dans une certaine mesure l'effort de financement actuel des activités sportives par le gouvernement. Bien sûr, l'instrument retardé proposé pourrait être corrélé avec le terme d'erreur s'il explique le succès actuel aux Jeux, ce qui peut poser un problème. Cependant, la variable explicative du financement a été construite en moyenne d'investissement pour les 4 années antérieures à une édition des JO. De ce fait, l'instrument recueille l'information sur les investissements qui ont été réalisés 4 années avant une édition des JO. Ainsi, la corrélation entre le succès olympique à un moment donné et l'instrument retardé devrait être faible.

Un problème classique découlant de l'approche par variable instrumentale est la possibilité d'avoir un instrument faible; soit un instrument faiblement corrélé avec la variable explicative endogène. Dans ce cas, les estimateurs des variables instrumentales obtenus par doubles

moindres carrés (DMC) ainsi que les principaux tests d'hypothèses sont biaisés par la taille de l'échantillon. D'autre part, Stock et Yogo (2005) ont proposé des seuils critiques du test décrit par Cragg et Donald (1993) en deçà duquel une VI est considérée invalide. Cependant, ce dernier test requiert la suridentification du modèle et par conséquent, l'utilisation d'au moins deux instruments valides. L'investissement retardé de deux périodes ainsi que le cycle électoral des pays ont tous les deux été testés et rejetés comme instrument valide dû à leur faible corrélation avec la variable endogène. Bien que notre VI proposée semble intuitivement valide, nous n'avons pu trouver d'instrument supplémentaire afin de confirmer la force de celle-ci.

Chapitre 4

Résultats et analyse

Ce mémoire met l'accent sur deux catégories de spécifications du modèle explicatif du succès olympique dont les résultats seront présentés dans cette section. Les modèles linéaires offrent l'avantage d'être aisément interprétables et programmables. Par contre, ces derniers ne permettent pas de prendre en considération le problème de censure dans les données. Par conséquent, l'usage de procédures d'estimation non linéaires doit être employé afin d'obtenir des estimateurs qui ne soient pas biaisés. Les méthodes d'estimations seront présentées sans et avec prise en considération des effets invariants propres à chacun des pays. De plus, les résultats obtenus suivant l'instrumentalisation de l'investissement public en sport à l'aide de cette dernière retardée d'une période sont présentés pour la spécification linéaire et pour le Tobit. L'annexe B présente les tableaux de résultats et les statistiques pertinentes à l'analyse de ceux-ci.

4.1 Spécifications linéaires

Le tableau B.1 présente les paramètres estimés par différentes spécifications linéaires testées. Dans la première colonne se trouvent les résultats de l'estimation par la méthode des moindres carrés ordinaires (MCO). Cependant, cette spécification du modèle ne considère pas la culture sportive des pays contrairement aux spécifications qui seront présentées par la suite. Dans la seconde colonne se retrouve l'estimation de spécification panel avec effets aléatoires (E.A.). Dans la troisième colonne se retrouvent les coefficients obtenus par la spécification ajoutant les variables dichotomiques pour chacun des pays afin de considérer les effets fixes (E.F.). Comme les coefficients obtenus pour les variables indicatrices des pays ne sont pas convergents dû au problème de paramètres incidents, ils ne se retrouveront pas dans le tableau des résultats. La réalisation d'un test de Hausman ne permet pas de rejeter l'hypothèse nulle du test¹. Ainsi, on ne peut rejeter que les estimateurs des effets aléatoires sont convergents et

1. Prob $> \chi^2 = 0,1098$.

efficaces. C'est pourquoi la dernière colonne présente les estimateurs à effets aléatoires avec variable instrumentale (E.A.+VI). La variable d'investissement sportif est alors instrumentée avec les dépenses gouvernementales réalisées pour les JO précédents.

Bien que des différences se remarquent entre les valeurs obtenues, le sens de l'impact et l'ordre de grandeur des déterminants du succès olympique sont similaires, peu importe la spécification du modèle retenue. Sans surprise, les coefficients estimés pour le pays hôte, pour le logarithme de la population et pour le logarithme de l'investissement sportif par habitant sont positifs et significatifs. Ainsi, selon les spécifications linéaires réalisées, *ceteris paribus*, une augmentation de 1% des dépenses publiques en sport de son pays augmentera en moyenne sa part de médailles remportées de 1 à 2 unités de pourcentage. En contrepartie, si un pays n'augmente pas son financement public en sport et que les nations adverses augmentent toutes leurs dépenses de 1%, toutes choses étant égales, alors ce dernier peut s'attendre à une baisse de 1 à 2 unités de pourcentage de son résultat olympique. L'impact est davantage important en réponse à un accroissement de son nombre de citoyens par rapport à la taille de la population mondiale (6,5 à 7 unités de pourcentage) ou alors, lorsque l'on reçoit les JO à domicile (20 à 35 unités de pourcentage). Pour finir, la valeur du paramètre ρ se trouvant sous les deux spécifications par effets aléatoires indique la part de la variance du terme d'erreur constant α_i dans la variance totale.

Il est à noter que lorsque l'on contrôle pour la culture sportive, le produit intérieur brut par tête ne semble plus influencer de façon significative le positionnement des nations concurrentes aux JO. Ce résultat diffère grandement de la littérature qui ne considérait pas le financement gouvernemental du sport et qui trouvait un effet important pour cette variable. Nous avons réalisé des estimations excluant le financement gouvernemental du sport et ces dernières obtenaient des coefficients élevés et significatifs du PIB par habitant. Ce constat soutient l'hypothèse que le financement public du sport reflète la volonté d'un pays à s'illustrer sur la scène internationale. Par conséquent, puisqu'il décrit un effort matérialisé d'un pays afin d'augmenter son positionnement aux JO d'été, il s'agit d'un déterminant plus adapté du succès olympique que la richesse relative d'un pays.

4.2 Spécification non linéaires

Les spécifications précédentes ne considéraient pas le problème de censure au niveau des données. Respectivement dans le tableau B.2 se trouvent le Tobit, le Tobit panel avec effets aléatoires (E.A.), le Tobit qui inclut des variables dichotomiques pour chacun des pays (Tobit E.F.) ainsi que ce dernier Tobit instrumenté (Tobit E.F.+VI) permettent, tous les quatre, de

palier, du moins en partie, ce problème de non linéarité. À ces spécifications du modèle est ajoutée la procédure d'estimation en deux étapes d'Heckman. Cependant, comme il l'a été relevé dans le chapitre précédent, les spécifications Tobit à effets fixes font face au problème de paramètres incidents. Par conséquent, les coefficients estimés à l'aide des effets fixes sont non convergents. Ces derniers se retrouvent dans la troisième (E.F.) de même que dans la quatrième colonne (E.F.+VI). Le tableau B.2 présente les effets marginaux sur l'espérance de médailles des estimations non linéaires mesurées en calculant la moyenne des effets marginaux pour chaque observation ainsi que les statistiques pertinentes à l'analyse de ceux-ci.

4.2.1 Tobit

Contrairement aux spécifications linéaires précédemment discutées, les estimations non linéaires permettent de tenir compte de la censure des données. Les coefficients obtenus sont de signes et de magnitudes qui ne se comparent que très faiblement entre eux. Lorsque l'on considère le cas le plus classique de la spécification Tobit, l'investissement public en sport influence le succès olympique de façon très similaire au modèle estimé par les MCO. Dans ce même cas, la population, que la littérature a consacré comme un déterminant principal du résultat olympique, ne semble alors que très peu influencer le nombre de médailles remportées par un pays aux JO. Lorsque l'on prend en considération la culture sportive des pays à l'aide d'un Tobit en panel ou en ajoutant des variables indicatrices pour chacun des pays, le financement du sport semble influencer davantage son succès olympique. Selon les spécifications Tobit estimées, un gouvernement qui augmente de 1% ses dépenses sportives par habitant obtiendra, toute chose étant égale par ailleurs, une hausse d'approximativement 1,5 à 3 unités de pourcentage sa part des médailles. Cependant, il n'est pas possible de se fier aux coefficients estimés suite à l'ajout de variables dichotomiques. Par ailleurs, l'effet d'un accroissement de son bassin d'athlètes potentiels par rapport à la population mondiale semble influencer de façon bien différente le succès olympique. Cela démontre bien la non convergence des spécifications due au problème de paramètres incidents. Il est à noter que la colonne E.F.+VI présentant les coefficients obtenus, suite à l'inclusion de notre instrument et des effets fixes, diffère grandement des paramètres estimés pour la spécification incluant seulement les effets fixes. L'inclusion de l'investissement retardé comme instrument du financement actuel du sport semble donc influencer sur les résultats d'estimation. Cependant, on ne peut encore ici être assuré de la validité de cette opération.

4.2.2 Procédure en deux étapes de Heckman

La procédure d'estimation en deux étapes d'Heckman permet de corriger le biais de sélection. Cependant, cette méthode permet de contrôler les variables explicatives des effets qui sont propres à chaque pays et ignorés du processus de sélection. C'est pourquoi le processus de sélection des pays médaillés a été identifié à l'aide des mêmes variables que celles utilisées

dans la seconde étape de la méthode sans contrôler pour les caractéristiques inobservables et invariantes dans le temps d'un pays. De cette façon, la seconde étape se trouve adéquatement identifiée. Les résultats obtenus par cette estimation sont présentés dans les deux dernières colonnes du tableau B.2. Dans la première colonne se trouve les effets marginaux sur la probabilité d'obtenir une première médaille. La seconde colonne présente les effets marginaux évalués à la moyenne des effets marginaux pour chaque observation pour les pays médaillés.

Les paramètres estimés par ce processus sont différents de ceux obtenus par les autres méthodes d'estimation. On remarque avec étonnement que selon cette spécification le pays hôte des Jeux ne semble pas posséder un avantage sur ses compétiteurs. Ce phénomène que possède le pays organisateur des JO semble pourtant bien documenté dans la littérature ce qui pourrait nous porter à exclure ces résultats. Cependant, les coefficients estimés pour l'investissement et la population sont toujours significatifs et leurs valeurs estimées sont d'un ordre de grandeur conséquent avec les résultats des spécifications non linéaires précédentes. Ainsi, selon cette estimation et toutes choses étant égales par ailleurs, un pays moyen qui augmente ses dépenses en sport de 1% haussera en moyenne de 3 unité de % sa part des médailles à financement gouvernemental mondial en sport constant. Il est intéressant d'ajouter que les coefficients estimés par l'équation de sélection des pays médaillés sont tous significatifs, ce qui n'est pourtant pas le cas pour les paramètres estimés de l'équation observée. Ainsi, l'information qui ne semble pas expliquer le nombre de médailles obtenues aux JO permet tout de même de différencier les nations médaillées des pays moins privilégiés. Ainsi, le produit intérieur brut par habitant qui ne semble plus influencer le résultat olympique dans notre modèle semble tout de même être une information importante puisqu'elle permet de différencier les pays lors de l'étape de sélection des nations médaillées.

4.3 Estimation du coût d'une médaille

Les résultats des spécifications précédentes nous permettent d'estimer le coût marginal d'une médaille olympique. Selon la spécification utilisée, nous estimons que, *ceteris paribus*, un pays aux caractéristiques moyennes doit investir entre 35,6 et 106,7 millions de dollars afin d'accroître d'une médaille son résultat olympique. Pour les pays qui ne sont pas médaillés, c'est en moyenne entre 8,4 et 25,2 millions de dollars supplémentaires que le gouvernement doit allonger afin de connaître un premier succès. Pour sa part, nous estimons qu'une somme supplémentaire de 38,5 à 115,5 millions de dollars aurait dû être investie par le Canada afin de remporter une médaille additionnelle aux Jeux olympiques de Londres en 2012. Ce résultat est nettement supérieur au coût moyen d'investissement par médaille engagé par le pays afin de remporter les 18 médailles qu'il a obtenues à cette édition des JO, le coût moyen par médaille s'élevant alors à près de 12,5 millions de dollars. L'ensemble de ces résultats suppose que l'ef-

fort d'investissement devant être réalisé pour obtenir des médailles olympiques est croissant. Néanmoins, cette observation découle de la forme logarithmique qui est imposée au modèle utilisé.

Conclusion

La relation entre le financement public accordé aux sports et les perspectives de gain aux JO est très peu documentée. Des sommes substantielles sont pourtant investies par l'ensemble des nations dans l'attente des exploits de leurs athlètes sur la scène sportive internationale. Depuis ses balbutiements, ce mémoire souhaite déterminer l'effet des politiques d'investissements en sport dont les gouvernements se dotent sur la récolte de médailles aux Jeux olympiques d'été. À notre connaissance, seule la modélisation de Flores *et al.* (2008) entreprenait d'inclure une telle variable comme facteur explicatif du succès olympique. Cependant, ces chercheurs ne traitent pas le biais dû à l'endogénéité découlant de l'inclusion de cette dernière. Sans pouvoir le soumettre au test de Stock et Yogo, nous suggérons le financement public du sport par habitant retardé d'une période dans le but de résoudre ce biais de simultanéité et d'arriver à des estimateurs qui soient convergents. De plus, ce travail se démarque par la caractérisation des déterminants du succès olympique continus sous forme de proportion. Celle-ci offre l'avantage de considérer l'incidence qu'ont les pays adverses sur son propre résultat olympique. En dernier lieu, plusieurs spécifications du modèle expliquant les résultats obtenus par les pays participants aux JO ont été estimées afin de résoudre les diverses problématiques qui font suite à l'ajout de la variable d'investissement.

Selon les spécifications utilisées, nous estimons qu'une augmentation de 1% des dépenses gouvernementales en sport pour un pays donné lui permettra, *ceteris paribus*, d'augmenter en moyenne de 1 à 3 unité de % sa part des médailles remportées au JO d'été à condition que les investissements sportifs de ses adversaires soient constants. Ainsi, selon nos résultats, la volonté des nations à s'illustrer sur la scène sportive internationale motive donc, dans une certaine mesure, la mise en place de politique de financement gouvernementale du sport. Par ailleurs, une augmentation d'un pour cent des dépenses élèvera à la marge de 0,15 pour cent la probabilité de connaître un premier succès pour les pays qui ne remportent pas de médailles. À l'aide des diverses spécifications utilisées, nous estimons qu'un pays aux caractéristiques moyennes devra accroître ses dépenses en sport de 35,6 à 106,7 millions de dollars afin de remporter une médaille supplémentaire aux JO d'été. À notre connaissance, ce mémoire est le premier travail identifiant clairement la politique gouvernementale de financement du sport comme déterminant du succès aux Jeux olympiques d'été.

Un point très important de ce mémoire est de gérer l'hétérogénéité des pays quant au processus de production des médailles olympiques d'un point de vue d'ensemble. Faisant suite aux résultats de ce mémoire, on peut penser qu'une part de l'avantage des pays face au succès sportif découle du choix auquel ils font face à travers l'allocation des ressources financières aux différentes institutions sportives nationales. Ainsi, l'étude et la compréhension du processus interne d'optimisation des investissements gouvernementaux en sports devrait faire suite à ce travail.

Qu'en est-il de la valeur prédictive des différents modèles proposés ? Trivedi et Zimmer (2012) argumentent la nécessité d'une modélisation plus complexe que la modélisation Tobit lorsque l'on souhaite prévoir les résultats des pays aux JO. Notamment, en plus de la nécessité d'estimer des paramètres convergents et non biaisés, il est d'intérêt de minimiser la variance et le bruit blanc. Ce mémoire examinait le rôle du financement gouvernemental en sport comme facteur explicatif du succès olympique. L'inclusion de cette variable explicative, le traitement de la simultanéité à l'aide de VI et l'attention portée à la culture sportive des pays en était donc l'objectif principal. L'étude des caractéristiques des pays médaillés et non médaillés permettrait assurément d'obtenir une spécification plus complète de la censure et diminuerait certainement la variance des estimateurs obtenus par le processus d'estimation en deux étapes d'Heckman. De plus, l'échantillon de données de l'investissement est relativement faible et ne couvre que 6 Jeux olympiques d'été. Pour l'ensemble de ces raisons, il est jugé hasardeux d'utiliser nos modélisations dans le but de prévoir le succès olympique futur. L'amélioration de la valeur prédictive du modèle serait, par ailleurs, une avenue intéressante que pourrait prendre un travail futur.

Finalement, comme les valeurs prises par le nombre de médailles olympiques remportées par les pays sont discrètes, il serait intéressant de modéliser ce compte à l'aide de la loi de Poisson ou de la loi binomiale négative. À notre connaissance, peu de recherches se sont attardées sur ce type de spécification. L'utilisation de lois discrètes dans un contexte de données de panel serait donc une route que devraient prendre de futures recherches sur le sujet.

Bibliographie

- Andreff, M. et W. Andreff. 2012, «Is Hosting the Games Enough to Win ? A Predictive Economic Model of Medal Wins at 2014 Winter Olympics», *Papeles de Europa*, vol. 25, p. 51–75. URL <http://search.proquest.com/docview/1322235616?accountid=12008>.
- Andreff, M., W. Andreff et S. Poupaux. 2008, «Les déterminants économiques de la performance olympiques : Prevision des medailles qui seront gagnes aux Jeux de Pekin.», *Revue d'Economie Politique*, vol. 118, n° 2, p. 135–169. URL <http://search.proquest.com/docview/56785910?accountid=12008>.
- Ball, D. W. 1972, «Olympic Games Competition : Structural Correlates of National Success», *International Journal of Comparative Sociology*, vol. 13, n° 3-4, p. 186–200. URL <http://search.proquest.com/docview/60893889?accountid=12008>.
- Bernard, A. B. et M. R. Busse. 2004, «Who Wins the Olympic Games : Economic Resources and Medal Totals», *Review of Economics and Statistics*, vol. 86, n° 1, p. 413–417. URL <http://search.proquest.com/docview/56238405?accountid=12008>.
- Buts, C., C. Du Bois, B. Heyndels et M. Jegers. 2011, «Socioeconomic Determinants of Success at the Summer Paralympics», *Journal of Sports Economics*, vol. 14, n° 2, doi :10.1177/1527002511416511, p. 133–147, ISSN 1527-0025. URL <http://jse.sagepub.com/cgi/doi/10.1177/1527002511416511>.
- Clarke, S. R. 2000, «Home advantage in the Olympic Games», dans *Proceedings of the Fifth Australian conference on Mathematics and Computers in Sport*, édité par G. Cohen et T. Langtry, University of Technology Sydney, Sydney, p. 76–85. URL <http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Home+advantage+in+the+Olympic+Games#1>.
- Condon, E. M., B. L. Golden et E. A. Wasil. 1999, «Predicting the success of nations at the Summer Olympics using neural networks», *Computers & Operations Research*, vol. 26, n° 13, doi :[http://dx.doi.org/10.1016/S0305-0548\(99\)00003-9](http://dx.doi.org/10.1016/S0305-0548(99)00003-9), p. 1243–1265. URL <http://www.sciencedirect.com/science/article/pii/S0305054899000039>.

- Cragg, J. G. et S. G. Donald. 1993, «Testing Identifiability and Specification in Instrumental Variable Models», doi :10.1017/S0266466600007519.
- Flores, R. J., D. Forrest, I. Sanz et J. de Dios Tena. 2008, «Can Governments Buy Olympic Medals?», *Workshop : The Economic of the Olympic Games*, p. 8.
- Johnson, D. K. N. et A. Ali. 2000, «Coming to play or coming to win : Participation and success at the Olympic Games», *Working Paper*, n° Wellesley College Dept. of Economics, p. 21. URL http://papers.ssrn.com/sol3/papers.cfm?abstract_id=242818.
- Johnson, D. K. N. et A. Ali. 2004, «A Tale of Two Seasons : Participation and Medal Counts at the Summer and Winter Olympic Games», *Social Science Quarterly*, vol. 85, n° 4, p. 974–993. URL <http://search.proquest.com/docview/56460777?accountid=12008>.
- Krishna, A. et E. Haglund. 2008, «Why do some countries win more Olympic medals? Lessons for social mobility and poverty reduction», *Economic and Political Weekly*, p. 143–151. URL <http://www.jstor.org/stable/40277720>.
- Levine, N. 1974, «Why do Countries Win Olympic Medals? Some Structural Correlates of Olympic Games Success : 1972», *Sociology and Social Research*, vol. 58, n° 4, p. 353–361. URL <http://search.proquest.com/docview/60858225?accountid=12008>.
- Li, Y., L. Liang, Y. Chen et H. Morita. 2008, «Models for measuring and benchmarking olympics achievements», *Omega*, vol. 36, n° 6, doi :10.1016/j.omega.2007.05.003, p. 933–940, ISSN 03050483. URL <http://linkinghub.elsevier.com/retrieve/pii/S0305048307000886>.
- Lozano, S., G. Villa, F. Guerrero et P. Cortés. 2002, «Measuring the performance of nations at the Summer Olympics using data envelopment analysis», *Journal of the Operational Research Society*, vol. 53, n° 5, doi :10.1057/palgrave/jors/2601327, p. 501–511, ISSN 0160-5682. URL <http://www.palgrave-journals.com/doi/10.1057/palgrave/jors/2601327>.
- Patrimoine Canadien. 2012, «Rapport sur les Plans et les Priorités», cahier de recherche, Gouvernement Canadien. URL http://www.olympic.org/Documents/Reports/FR/fr_report_670.pdf.
- Pfau, W. 2006, «Predicting the Medal Wins by Country at the 2006 Winter Olympic Games : An Econometrics Approach», *Korean Economic Review*, vol. 22, n° 18829, p. 233–247. URL http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1267529.
- Poupaux, S. 2006, «Soviet and post-Soviet sports», dans *Handbook on the Economic of Sport*, édité par W. Andreff et S. Szymanski, Edward Elgar Publishing, p. 316–324.

- Stock, J. H. et M. Yogo. 2005, «Testing for Weak Instruments in Linear IV Regression», dans *Identification and Inference for Econometric Models*, édité par D. W. Andrews, chap. 5, Cambridge University Press, p. 573.
- Tcha, M. 2004, «The Color of Medals : An Economic Analysis of the Eastern and Western Blocs' Performance in the Olympics», *Journal of Sports Economics*, vol. 5, n° 4, doi : 10.1177/1527002503257212, p. 311–328, ISSN 1527-0025. URL <http://jse.sagepub.com/cgi/doi/10.1177/1527002503257212>.
- The Nielsen Company. 2008, «The Final Tally - 4.7 Billion Tunes in to Beijing 2008 : More Than Two in Three People Worldwide», cahier de recherche, The Nielsen Company. URL http://www.nielsen.com/content/dam/corporate/us/en/newswire/uploads/2008/09/press_release3.pdf.
- Trivedi, P. et D. Zimmer. 2012, «Success at the Summer Olympics : Do Economic Factors Have High Predictive Value?», *Working Paper*, p. 39. URL http://people.wku.edu/david.zimmer/index_files/draft_olympic.pdf.
- Wu, J., Z. Zhou et L. Liang. 2010, «Measuring the Performance of Nations at Beijing Summer Olympics Using Integer-Valued DEA Model», *Journal of Sports Economics*, p. 549–566. URL <http://search.proquest.com/docview/760190605?accountid=12008>.

Annexe A

Définition des variables

Tableau A.1 – Statistiques descriptives des variables
basées sur 318 observations : 53 pays (i) et 6 olympiades (t)

| Symbole | Variable | Définition | Moyenne | Écart-type |
|-------------|----------------------------------------|----------------------------------------------------------------------------------------------|---------|------------|
| M_{it} | Nombre de médailles | $\frac{med_{it}}{\sum_i med_{it}}$ | 0,019 | 0,034 |
| INV_{it} | Investissement par habitant | $\ln \left(\frac{Inv_{it}/Population_{it}}{\sum_i Inv_{it}/\sum_i Population_{it}} \right)$ | -5,342 | 2,229 |
| PIB_{it} | Produit intérieur brut par habitant | $\ln \left(\frac{PIB_{it}/Population_{it}}{\sum_i PIB_{it}/\sum_i Population_{it}} \right)$ | -4,679 | 1,429 |
| POP_{it} | Population | $\ln \left(\frac{Population_{it}}{\sum_i Population_{it}} \right)$ | -5,631 | 1,734 |
| COM_i | Communiste | = 1 si communiste ; 0 sinon. | 0,019 | 0,136 |
| $PCOM_i$ | Postcommuniste | = 1 si postcommuniste ; 0 sinon. | 0,057 | 0,231 |
| $HÔTE_{it}$ | Hôte | = 1 si pays hôte ; 0 sinon. | 0,016 | 0,146 |

Annexe B

Résultats des spécifications

Tableau B.1 – Régressions Linéaires

| | MCO | E.A. | E.F. | E.A.+VI | |
|----------------------------------------|----------------------|----------------------|-----------------------|-----------------------|-----------------------|
| | | | | 1 ^{re} étape | 2 ^e étape |
| log(PIB per-capita/PIB-tot per-capita) | 3,572*** (0,836) | 0,812 (0,529) | 0,257 (0,553) | 0,124* (0,064) | 0,630 (0,505) |
| communiste | 31,245*** (5,982) | 28,426** (13,221) | 31,537*** (14,091) | 0,351 (1,435) | 30,065*** (11,664) |
| post-communiste | 2,772 (3,390) | 1,117 (7,353) | 9,672** (4,335) | 0,485 (0,799) | 0,361 (6,477) |
| pays hôte | 34,650*** (6,111) | 20,130*** (2,320) | 19,691*** (2,259) | -0,053 (0,307) | 23,037*** (2,492) |
| log(pop-pays/pop-tot) | 6,447*** (0,492) | 6,436*** (0,943) | 7,230** (3,047) | -0,098 (0,115) | 6,540*** (0,829) |
| log(invest-pays/invest-tot) | 1,161** (0,545) | 1,194*** (0,331) | 0,974* (0,345) | 0,650*** (0,044) | 2,009*** (0,551) |
| constante | 68,936*** (4,103) | 56,519*** (1,664) | 48,668*** (18,737) | -1,875** (0,798) | 60,545*** (1,464) |
| ρ | × | 0,864 | × | × | 0,845 |

Écart type entre parenthèses ; * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$;

ρ : fraction de la variance provenant du terme d'erreur α_i .

Tableau B.2 – Régressions Non Linéaires

| | Tobit | | | Heckman | | | |
|----------------------------------------|----------------------|----------------------|----------------------|-----------------------|----------------------|---------------------|----------------------|
| | E.A. | E.F. | | 1 ^{re} étape | 2 ^e étape | Sélection | Heckman |
| | | E.F. | E.F.+VI | | | | |
| log(PIB per-capita/PIB-tot per-capita) | -1,728** (0,802) | 0,836 (0,619) | -0,151 (0,409) | 0,081 (0,072) | 0,188 (0,512) | 0,071** (0,021) | 2,853 (1,681) |
| communiste | 27,207*** (5,521) | 14,372 (8,832) | 49,659*** (3,595) | 10,660*** (2,008) | -10,553 (17,360) | 0,759*** (0,065) | 9,358 (8,385) |
| post-communiste | -0,118 (3,124) | 5,422 (4,971) | 1,875 (2,204) | -0,973 (0,640) | 12,059*** (4,092) | 0,328*** (0,086) | 14,156*** (4,918) |
| pays hôte | 29,487*** (5,560) | 11,699*** (1,765) | 11,683*** (1,525) | -0,102 (0,318) | 13,688*** (1,656) | 0,642*** (0,055) | 10,251 (8,429) |
| log(pop-pays/pop-tot) | 0,523* (0,314) | 5,944*** (0,776) | -1,157** (0,575) | -2,037*** (0,441) | 11,754*** (3,680) | 0,143*** (0,009) | 7,739*** (0,884) |
| log(invest-pays/invest-tot) | 0,968* (0,567) | 2,133*** (0,453) | 1,516*** (0,439) | 0,635*** (0,051) | 3,265*** (0,769) | 0,060*** (0,013) | 3,164** (1,216) |
| Inverse du ratio de Mills : λ | × | × | × | × | × | × | 43,578** (14,236) |
| l-vraisemblance | 319 | 597 | 696 | × | 393 | × | × |

Écart type entre parenthèses ; * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$;

Effets marginaux évalués à la moyenne.