

17 juin 2020



ESTIMATION DE LA QUEUE SUPÉRIEURE DE LA DISTRIBUTION DU PATRIMOINE FAMILIAL AU CANADA



BUREAU DU DIRECTEUR PARLEMENTAIRE DU BUDGET
OFFICE OF THE PARLIAMENTARY BUDGET OFFICER

Le directeur parlementaire du budget (DPB) appuie le Parlement en fournissant des analyses économiques et financières dans le but d'améliorer la qualité des débats parlementaires et de promouvoir une plus grande transparence et responsabilité en matière budgétaire.

Le DPB a conçu une approche de modélisation pour estimer la queue supérieure de la distribution du patrimoine familial au Canada. Cette approche donne lieu à une nouvelle microbase de données sur les familles à valeur nette élevée et permet d'entreprendre des analyses et des estimations de coûts. Le présent rapport décrit l'approche ayant servi à l'établissement de la base de données et fait la démonstration de ses capacités d'analyse.

Le DPB souhaite remercier le professeur Jim Davies de lui avoir fourni de précieuses précisions techniques sur l'estimation de la queue supérieure de la distribution du patrimoine familial, ainsi que les fonctionnaires membres de l'équipe responsable de l'Enquête sur la sécurité financière (ESF) de Statistique Canada d'avoir répondu à ses nombreuses questions au sujet de l'ESF.

Analyste principal :
Nigel Wodrich, analyste

Collaborateur :
Aidan Worswick, analyste

Ce rapport a été préparé sous la supervision de :
Xiaoyi Yan, directrice

Nancy Beauchamp, Carol Faucher, Jocelyne Scrim et Rémy Vanherweghem ont contribué à la préparation du rapport pour publication.

Pour obtenir de plus amples renseignements, veuillez communiquer avec nous à l'adresse suivante : dpb-pbo@parl.gc.ca.

Yves Giroux
Directeur parlementaire du budget

RP-2021-007-S_f

Table des matières

Résumé	1
1. Introduction	3
2. Calcul du patrimoine familial au Canada	4
3. Établissement de la base de données	8
4. Capacités de la base de données	9
Annexe A : Approche de modélisation et hypothèses	13
A.1 Première mise en concordance des données	13
A.2 Intégration des données de la liste des personnes riches	15
A.3 Interpolation de Pareto	17
A.4 Étalonnage par itération	21
Annexe B : Statistiques sommaires	22
Annexe C : Évolution de la base de données	24
Références	25
Notes	28

Résumé

Le directeur parlementaire du budget (DPB) a conçu une approche de modélisation pour estimer la queue supérieure de la distribution du patrimoine familial au Canada. Son principal objectif est d'examiner les données sous-déclarées et manquantes liées aux familles à valeur nette élevée dans le Fichier de microdonnées à grande diffusion de l'Enquête sur la sécurité financière (FMGD de l'ESF). S'appuyant sur les Comptes du bilan national (CBN), le modèle réétalonne les données du FMGD de l'ESF afin d'ajouter un ensemble de données synthétiques sur les familles dont le patrimoine est supérieur à trois millions de dollars.

Ce travail de modélisation a abouti à une nouvelle ressource analytique : la *Base de données sur les familles à valeur nette élevée* (BDFE). Cette dernière permet au DPB de produire des estimations des coûts et des analyses des mesures qui touchent les familles canadiennes dont le patrimoine s'élève à des millions et à des milliards de dollars.

Grâce à la BDFE, le DPB constate que les familles les plus riches du Canada possèdent un patrimoine nettement supérieur à celui déclaré dans le FMGD de l'ESF. La part du patrimoine du 1 % des familles les plus riches est de 12 points de pourcentage plus élevée dans la BDFE que dans le FMGD de l'ESF (tableau 1 du résumé). Cet écart s'explique par les erreurs d'échantillonnage et les erreurs non dues à l'échantillonnage parmi les familles à valeur nette élevée dans le FMGD.

Tableau 1 du résumé

Distribution du patrimoine familial, FMGD de l'ESF et BDFE, par quantile choisi, Canada, 2016

Quantile du patrimoine familial	FMGD de l'ESF	BDFE
	Part du patrimoine total (%)	Part du patrimoine total (%)
0,01 % supérieur	0,4	5,6
0,1 % supérieur	3,1	12,1
0,5 % supérieur	9,2	20,5
1 % supérieur	13,7	25,6
5 % supérieur	33,0	43,4
10 % supérieur	47,6	56,4
20 % supérieur	67,2	73,5
40 % moyen	30,5	25,3
40 % inférieur	2,3	1,2

Sources : Calculs réalisés par le DPB à l'aide du FMGD de l'ESF; Base de données du DPB sur les familles à valeur nette élevée.

Le présent rapport décrit l'approche de modélisation utilisée pour produire l'ensemble de données synthétiques sur les familles à valeur nette élevée et l'intégrer au FMGD de l'ESF, puis pour mettre en concordance les valeurs agrégées de l'ensemble de données combinées avec celles des CBN. Il servira de référence aux futurs travaux du DPB sur le sujet.

La BDFE a été établie grâce à des données publiques. D'autres documents sont disponibles sur demande.

1. Introduction

Durant la campagne électorale fédérale de 2019, le directeur parlementaire du budget (DPB) a estimé le coût des propositions faites par les partis politiques, à leur demande¹. L'une de ces estimations portait sur les recettes fiscales qui découleraient de la perception d'un impôt annuel sur le patrimoine net des familles fortunées, soit celles dont la valeur nette est supérieure à 20 millions de dollars².

Le DPB s'est heurté à un obstacle de taille pour répondre à la demande, soit l'absence d'une microbase de données publiques qui permet d'évaluer de manière fiable les familles à valeur nette élevée du Canada. Par exemple, le principal produit de microdonnées sur le patrimoine familial de Statistique Canada, le Fichier de microdonnées à grande diffusion de l'Enquête sur la sécurité financière (FMGD de l'ESF), retrace les familles dont le patrimoine s'élève à tout au plus 27 millions de dollars. À titre de comparaison, le patrimoine de la personne figurant à la dernière position du classement des 100 personnes les plus riches du magazine *Canadian Business* s'élève à 875 millions de dollars.

Pour remédier au manque de données, le DPB a mis au point une approche de modélisation qui permet d'estimer de manière fiable la queue supérieure de la distribution du patrimoine familial au Canada. Il a pour ce faire adapté la technique d'interpolation directe de Pareto expliquée dans Bach et coll. (2014) et Saez et Zucman (2019). Cette technique permet de créer un ensemble de données synthétiques qui établit un lien entre les microdonnées sur le patrimoine provenant du FMGD de l'ESF et de la liste des personnes les plus riches du magazine *Canadian Business* (CB). Cet ensemble de données synthétiques a permis au DPB de répondre à la demande et de publier une [estimation des coûts](#) de deux pages en septembre 2019.

Après l'élection fédérale, le DPB a décidé de tabler sur le travail qu'il avait fait pour concevoir un outil d'analyse fonctionnelle des familles à valeur nette élevée. Pour y arriver, il a apporté plusieurs améliorations à l'approche de modélisation utilisée pendant la campagne électorale. La plus importante d'entre elles a consisté à suivre la procédure de régression des moindres carrés ordinaires (MCO) modifiée et d'étalonnage par itération expliquée dans Vermeulen (2016) et (2018). L'approche améliorée établit une concordance entre les valeurs agrégées des actifs et des passifs et les valeurs nettes agrégées de la distribution réestimée du patrimoine familial avec celles des Comptes du bilan national (CBN). Du fait de ces améliorations, on ne peut pas établir de comparaison directe entre les estimations de coûts réalisées par le DPB pendant la campagne électorale et les résultats figurant dans le présent rapport.

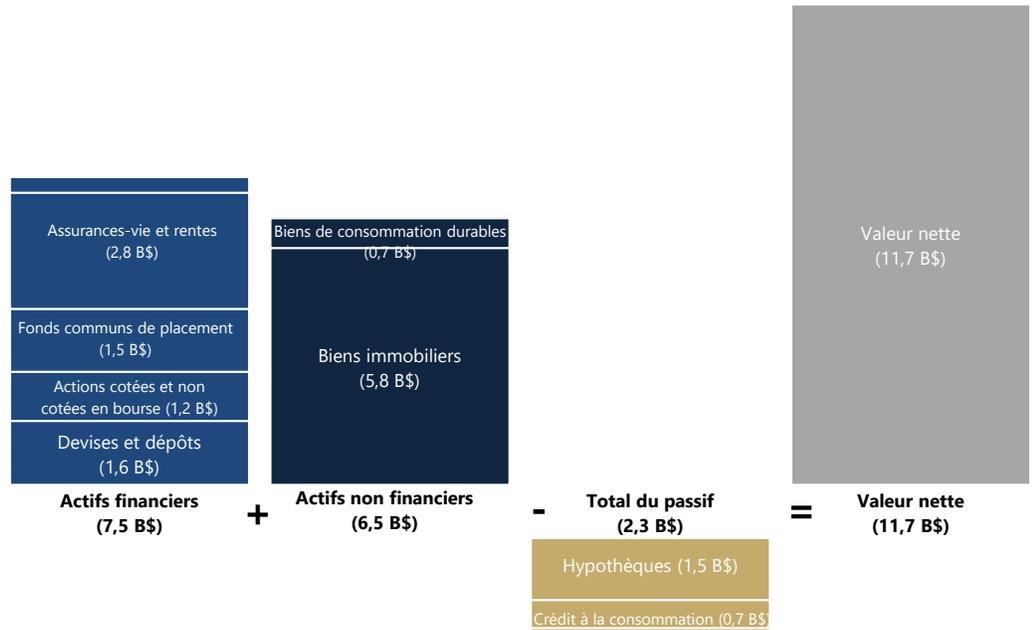
Tout ce travail de modélisation a abouti à un produit final : la *Base de données sur les familles à valeur nette élevée* (BDFE). La BDFE a été élaborée à partir de données publiées à la fin de l'année 2016, date la plus récente à laquelle toutes les sources ont communiqué des données. Elle servira à la réalisation d'analyses et d'estimations de coûts sur les familles à valeur nette élevée au fur et à mesure qu'elles seront nécessaires.

Pour présenter le type de travail analytique qui peut être réalisé à l'aide de la BDFE, des statistiques sommaires de la base de données sont présentées dans la section 4 et l'annexe B du rapport. Ces résultats sont présentés à titre d'exemple et peuvent différer de l'analyse d'une mesure particulière à l'aide de la BDFE.

2. Calcul du patrimoine familial au Canada

Pour les besoins du présent rapport, le DPB a estimé le patrimoine familial en fonction de la valeur nette négociable, soit le montant d'argent qui reste à une famille si elle liquide tous ses actifs financiers et non financiers et s'acquitte de toutes ses dettes^{3,4}. Les familles canadiennes détiennent collectivement un grand patrimoine. Selon les Comptes du bilan national (CBN) de Statistique Canada, qui comptabilisent les stocks d'actifs et de passifs et la valeur nette dans chaque secteur institutionnel, la valeur nette totale du secteur des ménages du Canada s'élevait à 11,7 billions de dollars à la fin de 2019. Ce chiffre est environ cinq fois plus élevé que le PIB annuel du pays⁵. Les biens immobiliers (5,8 billions de dollars) et les hypothèques sur ces derniers (1,5 billion de dollars) sont les catégories d'actifs et de passifs les plus importantes, respectivement (figure 2-1).

Figure 2-1 Actifs, passifs et valeur nette des ménages, Canada, T4, 2019

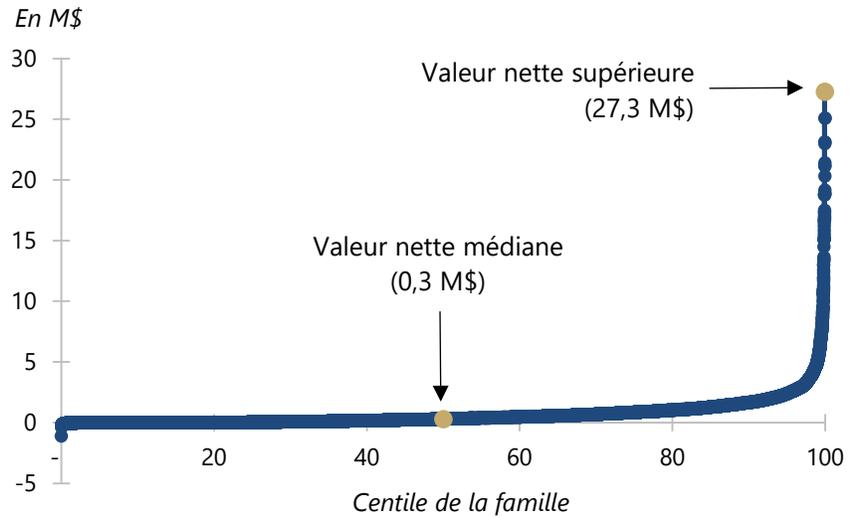


Source : Calculs réalisés par le DPB à l'aide du Tableau 36-10-0580-01 de Statistique Canada (Comptes du bilan national pour le secteur des ménages, T4, 2019).

La distribution du patrimoine entre les ménages est fortement biaisée en faveur des familles les plus riches⁶. Au Canada, une faible proportion des familles se trouvant au sommet de la distribution possède une valeur nette d'une ampleur sans égale à la valeur nette médiane du pays (figure 2-2). La forte concentration du patrimoine au sein d'un petit nombre de familles rend difficile une évaluation fiable du patrimoine au sommet de la distribution. Cette difficulté est palpable dans le Fichier de microdonnées à grande diffusion de l'Enquête sur la sécurité financière (FMGD de l'ESF), l'enquête nationale de Statistique Canada visant à calculer la valeur nette des Canadiens. La valeur nette de la famille la plus riche observée dans le FMGD de l'ESF de 2016 ne s'élevait qu'à 27 millions de dollars⁷; l'enquête ne fait état d'aucune famille plus riche pour plusieurs raisons possibles (encadré 2-1).

Figure 2-2

Distribution de la valeur nette familiale, Fichier de microdonnées à grande diffusion, Enquête sur la sécurité financière, 2016



Source : Calculs réalisés par le DPB à l'aide du FMGD de l'ESF de 2016.

Pour améliorer l'évaluation de la queue supérieure de la distribution du patrimoine familial, on peut adopter au moins quatre approches d'ordre général. La première consiste à compiler des dossiers sur chaque famille à valeur nette élevée, à la manière de la liste des milliardaires du monde établie par le magazine Forbes. La deuxième consiste à utiliser les déclarations d'impôt sur le revenu des particuliers pour calculer les revenus déclarés par les contribuables. La troisième consiste à utiliser les dossiers de l'impôt sur les biens transmis par décès pour connaître le patrimoine déclaré d'un défunt et émettre certaines hypothèses sur le lien entre ce patrimoine et le véritable patrimoine des vivants. La quatrième consiste à rajuster la distribution du patrimoine familial dans les enquêtes nationales comme le FMGD de l'ESF en utilisant des données provenant d'autres sources. Cette dernière approche est celle que le DPB a retenue et est expliquée plus en détail dans la section suivante.

Encadré 2-1 Limites des enquêtes nationales sur le patrimoine visant à dénombrer les familles à valeur nette élevée

Plusieurs raisons plausibles expliquent pourquoi les enquêtes nationales sur le patrimoine, comme l'ESF du Canada, parviennent mal à dénombrer et à analyser les familles à valeur nette élevée.

Des **erreurs d'échantillonnage** risquent de se glisser dans les enquêtes si l'échantillon n'est pas représentatif de la population, y compris celle au sommet de la distribution du patrimoine familial.

Les **erreurs de réponse**, c'est-à-dire les familles qui déclarent de manière inexacte, volontairement ou non, la valeur de leurs actifs et de leur passif, peuvent influencer les estimations relatives aux familles à valeur nette élevée.

Certaines grandes valeurs liées aux actifs et aux passifs dans le FMGD de l'ESG sont également susceptibles d'être **tronquées par le haut**, c'est-à-dire remplacées par une valeur maximale. Si cette procédure garantit la confidentialité des données publiées, elle réduit aussi les parts supérieures du patrimoine (voir l'annexe A.3).

Par ailleurs, la limite qui a le plus grand effet peut être la **non-réponse différentielle**, c'est-à-dire la tendance des familles à valeur nette élevée à moins participer aux enquêtes. Si ces familles sont sous-échantillonnées et que le poids de celles qui sont échantillonnées n'est pas suffisamment augmenté, les parts supérieures du patrimoine seront sous-estimées.

Même si Statistique Canada fait état d'un taux de réponse global de 70,3 % pour l'ESF de 2016, on sait peu de choses sur l'incidence de la non-réponse différentielle dans l'ESF. Aux États-Unis, des données montrent qu'il existe une corrélation positive entre le patrimoine et le taux de non-réponse dans la principale enquête nationale sur le patrimoine, l'Enquête sur les finances des consommateurs (*Survey of Consumer Finances*) (Kennickell et Woodburn, 1997).

Statistique Canada tente de remédier à la non-réponse différentielle des familles à valeur nette élevée en suréchantillonnant les zones géographiques dont on sait que les familles occupantes ont un revenu plus élevé et dont on pense qu'elles ont un patrimoine plus élevé (Statistique Canada, 2018a). Toutefois, des approches similaires de suréchantillonnage des familles à valeur nette élevée grâce à de l'information géographique ou à de l'information géographique stratifiée en fonction du revenu dans plusieurs pays européens se sont révélées d'une efficacité limitée pour calculer avec précision le patrimoine des familles à valeur nette élevée (Vermeulen, 2018).

3. Établissement de la base de données

Le DPB a établi la *Base de données sur les familles à valeur nette élevée* (BDFE) à l'aide de données provenant de trois sources :

1. **Fichier de microdonnées à grande diffusion de l'Enquête sur la sécurité financière (FMGD de l'ESF)**⁸ : L'ESF (FMGD) est l'enquête nationale sur la valeur nette des Canadiens. Statistique Canada mène une enquête sur les principaux actifs financiers et non financiers, ainsi que sur les dettes d'un échantillon représentatif de plus de 12 000 familles économiques résidentes⁹. La BDFE se fonde sur la dernière version publiée du FMGD de l'ESF en 2016.
2. **Comptes du bilan national (CBN)** : Les CBN regroupent les bilans individuels des ménages dans l'ensemble de l'économie et font état de leurs actifs financiers, de leurs actifs non financiers, de leur passif et, en fin de compte, de leur valeur nette¹⁰. La BDFE se fonde sur les données des CBN du quatrième trimestre de 2016, date qui concorde le mieux avec les dates auxquelles le FMGD de l'ESF et la liste des personnes les plus riches du magazine *Canadian Business* utilisés dans la base de données ont été établis¹¹.
3. **Liste des personnes les plus riches du magazine *Canadian Business*** : Le magazine *Canadian Business* réalise des recherches journalistiques et des études de marché en vue de dresser une liste des 100 Canadiens les plus riches¹². La BDFE utilise la liste des personnes les plus riches établie par le magazine en 2017, qui a été publiée en décembre 2016 et qui concorde le mieux avec le FMGD de l'ESF établi en 2016.

Le DPB a suivi l'approche élégante de Vermeulen (2016) pour remédier aux données manquantes et sous-déclarées des familles à valeur nette élevée dans le FMGD de l'ESF et créer la BDFE. Premièrement, les valeurs agrégées des actifs financiers, des actifs non financiers et du total du passif dans le FMGD de l'ESF ont été rajustées pour concorder avec les totaux correspondants par catégorie dans les CBN. Deuxièmement, les données de la liste des personnes les plus riches du magazine CB ont été ajoutées au FMGD de l'ESF. Troisièmement, l'ensemble de données communes qui en a résulté a été utilisé pour effectuer une régression MCO modifiée et ainsi déterminer la forme de la distribution du patrimoine des familles manquantes et sous-déclarées et faire le lien entre les familles figurant dans la partie supérieure du FMGD de l'ESF et celles figurant dans la partie inférieure de la liste des personnes les plus riches du magazine CB.

Quatrièmement, les résultats de la régression MCO modifiée ont servi à créer un nouvel ensemble de données synthétiques sur les familles à valeur nette élevée. Cinquièmement, l'ensemble de données synthétiques a été fusionné avec l'ensemble de données communes de la deuxième étape. L'ajout de l'ensemble de données synthétiques crée généralement plus d'actifs et de passifs que ce qui est prévu dans les CBN, ce qui mène à la sixième étape : réduire (ou augmenter) chacun des actifs financiers, des actifs non financiers et le total du passif dans le FMGD de l'ESF par un facteur d'ajustement et revenir à la deuxième étape pour répéter la procédure de façon itérative jusqu'à ce que la valeur des actifs financiers, des actifs non financiers et du total du passif dans l'ensemble final de données intégrées (combinant le FMGD de l'ESF rajusté en fonction des CBN, l'ensemble de données synthétiques et la liste des personnes les plus riches du magazine CB) soit égale à celle des CBN.

L'approche de modélisation utilisée pour établir la BDFE est décrite plus en détail à l'annexe A.

4. Capacités de la base de données

Le DPB a généré les statistiques sommaires de la BDFE pour mettre en évidence ses capacités d'analyse.

Grâce à la BDFE, il constate que les familles les plus riches du Canada possèdent un patrimoine nettement supérieur à celui déclaré dans le FMGD de l'ESF. La part du patrimoine du 1 % des familles les plus riches est de 12 points de pourcentage plus élevée dans la BDFE que dans le FMGD de l'ESF (tableau 4-1).

Tableau 4-1 Distribution du patrimoine familial, FMGD de l'ESF et BDFE, par quantile choisi, Canada, 2016

Quantile de patrimoine familial	FMGD de l'ESF		BDFE
	Part du patrimoine total	Part du patrimoine total	
	(%)		(%)
0,01 % supérieur	0,4		5,6
0,1 % supérieur	3,1		12,1
0,5 % supérieur	9,2		20,5
1 % supérieur	13,7		25,6
5 % supérieur	33,0		43,4
10 % supérieur	47,6		56,4
20 % supérieur	67,2		73,5
40 % moyen	30,5		25,3
40 % inférieur	2,3		1,2

Sources : Calculs réalisés par le DPB à l'aide du FMGD de l'ESF; Base de données du DPB sur les familles à valeur nette élevée.

L'annexe B présente d'autres statistiques sommaires pour la fin de l'année 2016, soit la période de « référence » de la BDFE, lorsque chacune de ses sources a communiqué ses données les plus récentes.

Pour analyser les familles à valeur nette élevée au cours des périodes suivantes, il faut formuler certaines hypothèses sur l'évolution des familles et de leur patrimoine depuis la fin de 2016. Afin d'illustrer le type d'hypothèses requises pour faire évoluer la BDFE, le DPB a également généré des statistiques sommaires sur les familles à valeur nette élevée pour la fin de l'année 2019. Il a supposé que, depuis 2016 :

- la composition des familles (nombre de personnes, âge, etc.) est demeurée constante dans l'ensemble de la distribution du patrimoine¹³;
- le nombre de familles a augmenté au même rythme que le nombre de personnes, et cette croissance a été uniforme dans l'ensemble de la distribution du patrimoine¹⁴;
- la valeur agrégée des actifs financiers, des actifs non financiers et du total du passif a augmenté au même rythme que celui prévu dans les CBN, et cette croissance a été proportionnelle dans l'ensemble de la distribution du patrimoine familial.

Suivant ces hypothèses, le DPB a fait deux rajustements à la BDFE. Tout d'abord, le poids de chaque observation a été augmenté en fonction du taux

de croissance démographique du Canada entre le quatrième trimestre de 2016 et le quatrième trimestre de 2019. Ensuite, les actifs financiers, les actifs non financiers et le total du passif de chaque observation ont été augmentés de manière proportionnelle jusqu'à ce que leurs totaux globaux correspondent à ceux des CBN au quatrième trimestre de 2019.

Les statistiques sommaires qui en découlent sont présentées dans les tableaux 4-2 et 4-3. Les deux tableaux mettent en évidence la forte concentration de patrimoine parmi les familles à valeur nette élevée du Canada.

Tableau 4-2 Distribution du patrimoine familial, par quantile choisi, Canada, 2019

Quantile de patrimoine familial	Seuil de patrimoine	Nombre de familles	Patrimoine total	Part du patrimoine total
	(M\$)	(milliers)	(G\$)	(%)
0,01 % supérieur	143,1	1,6	654	5,6
0,1 % supérieur	29,3	16,0	1 427	12,2
0,5 % supérieur	9,7	79,7	2 410	20,6
1 % supérieur	6,1	159,3	3 010	25,7
5 % supérieur	2,3	796,7	5 107	43,7
10 % supérieur	1,6	1 593,5	6 629	56,7
20 % supérieur	1,0	3 186,9	8 633	73,8
40 % moyen	0,1-1,0	6 373,8	2 932	25,1
40 % inférieur	Moins de 0,1	6 373,8	132	1,1

Sources : Base de données du DPB sur les familles à valeur nette élevée; calculs du DPB fondés sur les Estimations démographiques trimestrielles de Statistique Canada et les CBN.

Tableau 4-3 Distribution du patrimoine, par seuil de patrimoine choisi, Canada, 2019

Seuil de patrimoine familial	Nombre de familles	Patrimoine total	Part du patrimoine total
Familles dont le patrimoine est supérieur à :	(milliers)	(G\$)	(%)
1 G\$	0,1	221	1,9
500 M\$	0,2	333	2,8
250 M\$	0,7	488	4,2
100 M\$	2,7	785	6,7
50 M\$	7,2	1 097	9,4
25 M\$	19,4	1 525	13,0
10 M\$	76,3	2 377	20,3
5 M\$	206,6	3 271	28,0
2,5 M\$	699,1	4 871	41,6
1 M\$	3 123,7	8 570	73,3

Sources : Base de données du DPB sur les familles à valeur nette élevée; calculs du DPB fondés sur les Estimations démographiques trimestrielles de Statistique Canada et les CBN.

Annexe A : Approche de modélisation et hypothèses

A.1 Première mise en concordance des données

Le DPB a procédé à un premier rajustement des microdonnées du FMGD de l'ESF afin que les valeurs agrégées des actifs et des passifs et les valeurs nettes agrégées concordent avec les totaux correspondants par catégorie pour le secteur des ménages dans les CBN.

Même si le FMGD de l'ESF et les CBN estiment tous deux la valeur nette des ménages, il existe des distinctions procédurales et conceptuelles entre les deux sources qui conduisent à des estimations légèrement différentes¹⁵. De toute évidence, le FMGD découle d'une enquête dont les estimations sont assorties d'intervalles de confiance; les CBN estiment quant à eux les stocks et les flux dans les comptes de capital et les comptes financiers, mais puisque certaines catégories de ménages sont calculées comme résidu d'autres secteurs, les CBN ont une marge d'erreur qui leur est propre. L'ESF n'échantillonne pas les territoires et certains groupes de population, ce qui représente deux pour cent de la population. Certains actifs et passifs sont également calculés différemment. Par exemple, les CBN ne font pas état de la valeur des objets de collection, comme les œuvres d'art. Aussi, les deux sources calculent différemment les dettes contractées par carte de crédit. Il s'agit là de la principale raison qui explique, selon Statistique Canada (2019a), la sous-couverture du total du passif dans le FMGD de l'ESF (tableau A1-1).

Tableau A1-1 **Concordance entre le FMGD de l'ESF et le secteur des ménages des CBN, 2016**

	FMGD de l'ESF (G\$)	CBN (G\$)	Couverture (ESF/CBN)
Actifs financiers	5 845	6 468	0,904
Actifs non financiers	6 193	5 934	1,043
Total du passif	1 751	2 062	0,850
Valeur nette	10 287	10 339	0,995

Sources : Calculs réalisés par le DPB à l'aide du FMGD de l'ESF de 2016 et Tableau 36-10-0580-01 de Statistique Canada.

Notes : Les totaux des CBN correspondent aux résultats du quatrième trimestre de 2016.

Les biens mobiliers ont été comptés comme un actif financier.

Les chiffres ayant été arrondis, leur somme peut ne pas correspondre au total indiqué.

Néanmoins, il existe plusieurs raisons pour lesquelles il est souhaitable de faire concorder le FMGD de l'ESF avec les CBN. Cette mise en concordance peut compenser les sous-déclarations dans les enquêtes nationales sur le patrimoine (Vermeulen, 2016). Contrairement à l'ESF et à celle qui l'a précédée, l'Enquête sur les finances des consommateurs (EFC), les CBN ont été estimés de manière cohérente au fil du temps (Davies et Di Matteo, 2020). Cette cohérence permet de mieux comparer les estimations de la distribution du patrimoine familial dans le temps. Les CBN sont également estimés et publiés plus souvent (tous les trimestres) que les ESF (tous les trois ans). Cette fréquence plus élevée permet d'actualiser les estimations les années où aucune ESF n'est menée. Enfin, la concordance avec le secteur des ménages des CBN permet d'analyser la position globale des ménages par rapport aux autres secteurs économiques inclus dans les CBN (Statistique Canada, 2019a). C'est pour ces mêmes raisons que Statistique Canada procède également à la mise en concordance de l'ESF avec les CBN dans l'ensemble de données des Comptes économiques répartis pour le secteur des ménages (CERSM).

Afin de faire concorder le FMGD de l'ESF et les CBN, le DPB a d'abord classé chaque variable d'actif et de dette du FMGD de l'ESF en trois grandes catégories : les actifs financiers¹⁶, les actifs non financiers et le total du passif¹⁷. Pour chaque catégorie, il a calculé un facteur d'ajustement qui est l'inverse du calcul de la « couverture » présentée dans le tableau A1-1. Il a augmenté (ou diminué) la valeur des actifs financiers, des actifs non financiers et du total du passif pour chaque famille figurant dans le FMGD de l'ESF en appliquant le facteur d'ajustement pertinent. Puisque chaque famille possède un portefeuille unique d'actifs et de passifs, sa valeur nette varie de manière différente avec cette procédure d'ajustement¹⁸.

A.2 Intégration des données de la liste des personnes riches

L'étape suivante a consisté à ajouter les données sur le patrimoine provenant de la liste des personnes riches au FMGD de l'ESF rajusté en fonction des CBN.

Cet ajout se justifie par le souci d'améliorer la précision de l'analyse de régression subséquente, qui estime les paramètres de Pareto utilisés dans l'imputation des familles à valeur nette élevée manquantes et sous-déclarées¹⁹. Vermeulen (2018) montre que l'ajout ne serait-ce que d'un petit nombre d'entrées provenant de la liste des personnes riches améliore considérablement la précision des estimations interpolées de la queue supérieure de la distribution, suffisamment pour qu'il n'y ait pratiquement aucun biais d'estimation²⁰.

Au Canada, il existe deux grandes listes de personnes riches accessibles au public : la liste des milliardaires du monde du magazine Forbes, qui comprend des entrées canadiennes, et la liste des personnes les plus riches du magazine *Canadian Business* (CB). Le DPB a choisi cette dernière pour établir la BDFE, à l'instar de Davies et Di Matteo (2020). Ces derniers soulignent que la liste du magazine CB comprend des milliardaires qui ne figurent pas sur celle du magazine Forbes, ainsi que des entrées sous le seuil d'un milliard de dollars américains fixé par ce dernier²¹.

Avant que l'on puisse les ajouter au FMGD de l'ESF rajusté en fonction des CBN, les données de la liste des personnes les plus riches du magazine CB ont toutefois dû être rajustées²².

Contrairement au FMGD de l'ESF, le magazine CB inclut les Canadiens non-résidents dans son dénombrement des 100 Canadiens les plus riches. C'est pourquoi le DPB a retiré ces Canadiens de l'ensemble de données du magazine CB, comme l'a aussi fait MacDonald (2018).

En outre, plusieurs entrées du magazine CB font référence à des familles élargies qui comprennent plusieurs unités familiales. Il s'agit notamment des entrées intitulées « famille », « frères » et « succession » ou qui comprennent plusieurs personnes non mariées. À l'opposé, l'unité familiale du FMGD de l'ESF se compose de familles économiques et de personnes ne faisant pas partie d'une famille économique (personnes seules).

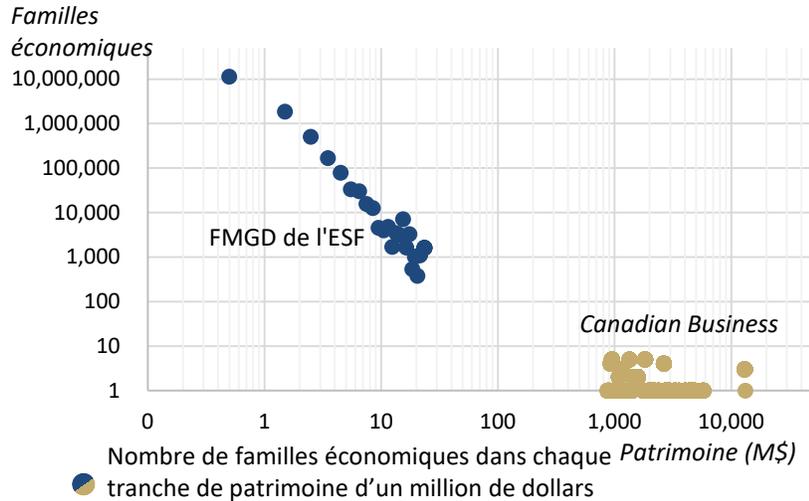
Le DPB a mis au point une méthode pour scinder les familles élargies figurant dans la liste du magazine CB en familles économiques constitutives. Il a utilisé des sources publiques pour déterminer la ou les générations qui gèrent le patrimoine familial. Chaque frère ou sœur (et cousin ou cousine, le cas échéant) de la ou des générations gestionnaires ainsi que leur(s) parent(s) vivant(s) (et oncles et tantes, le cas échéant) ont été traités comme une seule famille économique. Le DPB est parti du principe que le patrimoine déclaré de la famille élargie appartenait exclusivement et entièrement aux

familles économiques constitutives cernées. Il a aussi supposé que le patrimoine de la famille élargie est réparti de manière égale entre les familles économiques qui la composent. Enfin, il a retiré toutes les entrées issues de la scission qui étaient inférieures à l'entrée la plus basse (875 millions de dollars) de la liste du magazine CB. Cette dernière procédure lui a permis de s'assurer que la queue supérieure de la distribution du patrimoine (de plus de 875 millions de dollars) comprenait une population complète de familles au-dessus de ce seuil afin d'éviter tout biais dans l'analyse de régression subséquente visant à estimer les paramètres de Pareto.

À la suite de cette scission, l'ensemble de données épurées du magazine CB comprenait 80 familles économiques résidentes. Chacune d'entre elles détenait un patrimoine d'au moins 875 millions de dollars, et collectivement, elles détenaient un patrimoine de 197 milliards de dollars.

Le DPB a ajouté les données épurées du magazine CB au FMGD de l'ESF rajusté en fonction des CBN, créant ainsi un *ensemble de données communes* (voir la figure A2-1). Chaque observation du magazine CB s'est vu attribuer un poids de 1 (ce qui signifie que chacune d'entre elles représente une unité familiale).

Figure A2-1 Distribution du patrimoine familial dans l'ensemble de données communes²³, 2016



Sources : Calculs réalisés par le DPB à l'aide du FMGD de l'ESF de 2016 et liste des personnes les plus riches du magazine *Canadian Business*, 2017

A.3 Interpolation de Pareto

Le DPB a utilisé l'*ensemble de données communes* pour imputer les familles à valeur nette élevée manquantes et sous-déclarées. Pour ce faire, il a appliqué l'approche de régression MCO modifiée pour les plans d'enquête complexes de Vermeulen (2018). Il a ensuite interpolé les familles à valeur nette élevée manquantes et sous-déclarées à l'aide de l'estimation des paramètres de Pareto qui en a découlée.

Une hypothèse clé de cette technique d'imputation est que la queue supérieure de la distribution du patrimoine familial se conforme à une distribution de Pareto. Cette hypothèse a été largement appliquée dans la littérature sur les distributions de patrimoine, y compris au Canada. Davies et Shorrocks (1999) caractérisent la notion selon laquelle la queue supérieure de la distribution du patrimoine se conforme à une distribution de Pareto comme étant une « caractéristique permanente » de la distribution du patrimoine. Dans leur comparaison de différentes méthodes statistiques visant à imputer les observations tronquées par le haut dans le FMGD de l'ESF, Brzozowski et coll. (2010) supposent que le décile supérieur de l'ESF est distribué selon le principe de Pareto. Pour sa part, Ogwang (2011) constate, en recourant à des méthodes d'estimation MCO modifiée et EMV, que les listes des personnes les plus riches dressées par le magazine CB de 1999 à 2008 se comportent selon la loi de Pareto²⁴. Enfin, dans leur analyse de l'évolution des parts supérieures de patrimoine familial au Canada entre 1892 et 2016, Davies et Di Matteo (2020) partent du principe que la queue supérieure de la distribution du patrimoine se conforme à une distribution de Pareto.

Vermeulen (2018) émet une autre hypothèse fondamentale : les ensembles de données de l'enquête nationale sur le patrimoine et de la liste des personnes riches « se conforment à la même distribution de Pareto ». Le DPB formule aussi cette hypothèse, mais en faisant preuve de prudence pour deux raisons. La première est liée à la fiabilité et au contenu de la documentation disponible sur la méthodologie employée par le magazine CB. La plus récente méthodologie sur laquelle le DPB a pu mettre la main date de 2006²⁵. Cette méthodologie fournit de précieux renseignements sur l'approche suivie par le magazine CB, du moins pour l'année en question. On y souligne que certaines dettes (sociétés privées, biens immobiliers) sont à tout le moins estimées et déduites du total des actifs. Toutefois, on y indique également que des estimations intentionnellement prudentes sont utilisées pour évaluer les investissements privés et qu'on peut raisonnablement supposer que les 100 personnes les plus riches valent plus que la somme indiquée (Canadian Business, 2006). Davies et Di Matteo (2020) soulignent que les problèmes liés à la compilation des données des listes de personnes riches sont en partie résolus grâce à l'examen minutieux auquel les listes donnent lieu et aux améliorations qu'elles subissent

lorsqu'elles sont présentées chaque année (le magazine CB dresse la liste des personnes riches depuis 1999). S'il est raisonnable de supposer que le magazine CB a estimé le patrimoine des Canadiens dont la valeur nette est la plus élevée, rien ne permet d'affirmer avec certitude qu'il existe un biais dans ses données, le cas échéant.

La deuxième raison pour laquelle il faut faire preuve de prudence lorsqu'on suppose que l'*ensemble de données communes* repose sur le même modèle de distribution de Pareto a trait à la troncature par le haut réalisée dans le FMGD de l'ESF. En effet, ce dernier subit une troncature par le haut de telle sorte qu'un certain nombre de valeurs d'actifs et de dettes parmi les plus élevées sont remplacées par une valeur maximale afin d'y garantir la confidentialité de chaque observation divulguée. Cependant, cette façon de procéder réduit également le patrimoine des familles les plus riches dans ces fichiers par rapport aux données de l'ESF accessibles dans un centre de données de recherche (CDR), qui ne sont pas tronquées par le haut²⁶. Brzozowski et coll. (2010) font valoir que la part de patrimoine du 1 % des familles les plus riches était inférieure d'environ 1,5 point de pourcentage dans le FMGD de l'ESF de 1999 (13,2 %) par rapport aux données de l'ESF de 1999 accessibles dans un CDR (15,7 %). Le degré de troncature par le haut dans le FMGD de l'ESF de 2016 ne relève pas du domaine public, et le DPB n'a pas analysé l'étendue de la troncature par le haut ou de son possible biais sur les estimations de Pareto. En théorie, ce possible biais est atténué si l'on estime les paramètres de Pareto sur un segment suffisamment grand de la queue supérieure de la distribution de l'*ensemble de données communes*; un segment plus important devrait inclure, toutes proportions gardées, moins de familles tronquées par le haut, ce qui atténue ainsi le possible biais que ces familles peuvent introduire.

Pour appliquer l'approche de régression de Vermeulen (2018), le DPB a d'abord isolé un sous-ensemble d'observations i de l'*ensemble de données communes* du patrimoine $w_i \geq w_{min}$ visé par la régression. Le choix d'un w_{min} approprié, voire idéal, est incertain et se fait souvent au cas par cas²⁷. Dans la réestimation que fait Vermeulen (2018) des parts supérieures de patrimoine dans 10 pays européens et aux États-Unis, le choix de w_{min} dépend en partie de la méthode utilisée dans l'enquête nationale sur le patrimoine de chaque pays pour suréchantillonner les familles à valeur nette élevée, qui sont moins susceptibles de participer à de telles enquêtes. Les pays qui suréchantillonnent en utilisant des renseignements personnels, comme l'information sur l'impôt sur le revenu (États-Unis) ou sur le patrimoine imposable (Espagne, France), ont chacun été mis à l'essai avec des w_{min} allant de 500 000 € à 10 millions d'euros. À l'opposé, les pays qui suréchantillonnent en utilisant de l'information géographique stratifiée en fonction du revenu (Allemagne, Belgique) ou uniquement de l'information géographique (Autriche, Portugal), ou qui ne font aucun suréchantillonnage (Italie, Pays-Bas) ont chacun été mis à l'essai avec un w_{min} allant de 500 000 € à 2 millions d'euros seulement. Dans ces pays, il n'y avait pas assez

d'observations au-dessus du seuil de 2 millions d'euros pour estimer avec précision les paramètres de Pareto.

Dans l'ESF du Canada, on ne semble pas recourir à des renseignements personnels pour suréchantillonner les familles à valeur nette élevée. On y stratifie chaque province en zones rurales et urbaines. Dans les zones rurales, on utilise de l'information géographique pour sélectionner un échantillon à plusieurs degrés à partir de la base aréolaire de l'Enquête sur la population active. Dans les zones urbaines, on utilise de l'information provenant du Fichier des familles T1 (FFT1) du Fichier des indicateurs socioéconomiques (FSE), comme l'âge et le revenu, pour stratifier le registre des adresses en groupes de logements dont on prévoit que les familles occupantes auront un patrimoine semblable (Statistique Canada, 2018b). La strate urbaine où le patrimoine est le plus élevé représente les 5 % les plus riches de chaque province²⁸.

Le DPB a ainsi réduit la fourchette du w_{min} convenant au contexte canadien à une valeur comprise entre 750 000 \$ et 3 millions de dollars. Il s'agit d'une conversion approximative des valeurs de w_{min} en euros utilisées par Vermeulen (2018) pour les pays qui recourent également à de l'information géographique et à de l'information géographique stratifiée en fonction du revenu pour suréchantillonner les familles à valeur nette élevée. Vermeulen (2018) et Chakraborty et coll. (2019) attirent l'attention sur le compromis à faire dans le choix d'un w_{min} en particulier : un seuil plus bas augmente la taille de l'échantillon pour la régression, ce qui donne lieu à une estimation de Pareto plus fiable, quitte à inclure des observations qui ne suivent pas le comportement de la queue de distribution de Pareto.

En fin de compte, le DPB a choisi la limite supérieure de la fourchette du w_{min} , soit 3 millions de dollars. Par rapport aux enquêtes nationales sur le patrimoine mené dans les pays européens, qui procèdent à un suréchantillonnage en se servant d'informations géographiques ou d'informations géographiques stratifiées en fonction du revenu, le FMGD de l'ESF du Canada compte beaucoup plus d'observations au seuil de 2 millions d'euros, ou 3 millions de dollars²⁹. Le choix d'un w_{min} plus élevé permet de conserver un plus grand nombre d'observations du FMGD de l'ESF rajusté en fonction des CBN après l'interpolation, tout en maintenant une taille d'échantillon suffisante pour procéder à l'estimation par régression des paramètres de Pareto.

Après avoir choisi un w_{min} , le DPB a classé toutes les observations n liées au patrimoine $\{w_i, i = 1, \dots, n \mid w_i \geq w_{min}\}$ provenant de l'ensemble de données communes en ordre décroissant de patrimoine. Chaque observation liée au patrimoine w_i et au poids N_i a été définie en fonction de \bar{N} , soit le poids moyen de toutes les observations dont le patrimoine est égal ou supérieur à w_{min} , et aussi de \bar{N}_{f_l} , soit le poids moyen des observations i les plus riches

$$(\bar{N}_{f_l} = \frac{\sum_{j=1}^i N_j}{i}).$$

Vermeulen (2018) présente une dernière caractéristique de la régression. Gabaix et Ibragimov (2011) ont constaté que les régressions MCO du logarithme du rang et de la taille étaient systématiquement et fortement biaisées dans les petits échantillons. Vermeulen (2018) réduit donc de moitié ($1/2$) le classement de chaque observation de la régression. Cette modification réduit le biais à un ordre principal.

La régression MCO modifiée se décrit donc de la façon suivante :

$$\ln\left((i - 0.5) \frac{\overline{N}_{f_i}}{\overline{N}}\right) = \ln(n) + \alpha(\ln(w_{min}) - \ln(w_i))$$

Le coefficient estimé α provenant de la régression ci-dessus correspond au paramètre de Pareto, qui détermine la forme de la queue supérieure réestimée de la distribution du patrimoine familial. En général, plus α est élevé, plus la queue supérieure de la distribution est large et plus la concentration du patrimoine est élevée. Le coefficient estimé α a été appliqué à une fonction de distribution cumulative standard de Pareto sur un intervalle de patrimoine donné $\{[w_k, w_i] \mid w_i \geq w_k \geq w_{min}\}$:

$$f(w_i, w_k) = -w_{min}^\alpha * \left(\left(\frac{1}{w_i^\alpha} \right) - \left(\frac{1}{w_k^\alpha} \right) \right)$$

La fonction de distribution cumulative $f(w_i, w_k)$ ci-dessus donne des estimations entre 0 et 1 pour la probabilité qu'une famille située dans la queue supérieure de la distribution ait un patrimoine compris entre w_k et w_i . Suivant Chakraborty et coll. (2019), les estimations liées à la distribution cumulative ont été converties en nombre de familles synthétiques dans l'intervalle de patrimoine $[w_k, w_i]$ en multipliant la fonction de distribution $f(w_i, w_k)$ par $n\overline{N}$, soit la somme des poids de toutes les observations n liées au patrimoine $\{w_i, i = 1, \dots, n \mid w_i \geq w_{min}\}$. À l'instar de Davies et Di Matteo (2020), le DPB a conservé les entrées épurées du magazine CB sans interpolation. L'ensemble de données synthétiques qui en découle se compose de familles dont le patrimoine oscille entre w_{min} (3 millions de dollars) et w_{cbm} , soit le patrimoine le moins élevé de l'ensemble de données épurées du magazine CB (875 millions de dollars).

Après avoir remplacé les observations de l'ensemble de données communes liées au patrimoine $\{w_i, i = 1, \dots, n \mid w_{min} \leq w_i < w_{cbm}\}$ par l'ensemble de données synthétiques, le DPB a créé un ensemble de données intégrées. Ce dernier combine les familles du FMGD de l'ESF rajusté en fonction des CBN dont le patrimoine est inférieur à w_{min} , les familles de l'ensemble de données synthétiques dont le patrimoine oscille entre w_{min} et w_{cbm} et les familles dont les données ont été épurées par le magazine CB et dont le patrimoine est égal ou supérieur à w_{cbm} .

A.4 Étalonnage par itération

Le remplacement des familles à valeur nette élevée du FMGD de l'ESF rajusté en fonction des CBN par l'ensemble de données synthétiques interpolées et l'ensemble de données épurées du magazine CB crée un problème : le patrimoine global des familles du nouvel *ensemble de données intégrées* est plus élevé que celui du secteur des ménages des CBN.

Le DPB a imité Vermeulen (2016) en procédant à un étalonnage par itération qui met en concordance les valeurs agrégées des actifs et des passifs et les valeurs nettes agrégées de l'*ensemble de données intégrées* avec celles des CBN.

La première étape consiste à revenir au FMGD de l'ESF rajusté en fonction des CBN (le produit de l'annexe A.1). Pour les familles dont le patrimoine s'élève à $w_i \geq w_{min}$ dans le FMGD de l'ESF rajusté en fonction des CBN, le DPB a calculé trois ratios : actifs financiers agrégés par rapport à la valeur nette agrégée; actifs non financiers agrégés par rapport à la valeur nette agrégée; passif total agrégé par rapport à la valeur nette agrégée. Il a ensuite appliqué les ratios de manière uniforme aux familles synthétiques et à celles du magazine CB dans l'*ensemble de données intégrées* pour diviser leur patrimoine en valeurs d'actifs et de passifs constitutifs.

Dans l'étape suivante, le DPB a calculé les valeurs agrégées des actifs financiers, des actifs non financiers et du total du passif pour la totalité de l'*ensemble de données intégrées*. Les valeurs agrégées de cet ensemble de données ont été comparées à leurs valeurs correspondantes dans les CBN. Dans la mesure où les valeurs agrégées de l'*ensemble de données intégrées* étaient supérieures (ou inférieures) à celles des CBN, le DPB a révisé à la baisse (ou à la hausse) les facteurs d'ajustement appliqués aux données originales du FMGD de l'ESF présentées à l'annexe A.1. Puis, il a réestimé les paramètres de Pareto figurant à l'annexe A.3 et répété cette procédure d'ajustement et de réestimation de manière itérative jusqu'à ce que les valeurs agrégées des actifs financiers, des actifs non financiers et du total du passif concordent avec leurs valeurs correspondantes dans les CBN. Cette procédure a nécessité en général plusieurs répétitions pour produire des valeurs étalonnées en fonction des CBN pour tous les actifs et passifs. La valeur finale étalonnée du paramètre α , qui détermine la forme de la distribution du patrimoine familial, était de 1,45³⁰.

Les derniers facteurs d'ajustement appliqués aux données du FMGD de l'ESF pour mettre en concordance l'*ensemble de données intégrées* avec les CBN sont présentés dans le tableau A4-1. Globalement, les actifs financiers et non financiers ont été réduits de 5,8 et 13 % respectivement. Le total du passif a été ajusté à la hausse de 12,8 %, ce qui correspond aux passifs déclarés nettement moins élevés dans le FMGD de l'ESF par rapport aux CBN.

Tableau A4-1

Facteurs d'ajustement appliqués au FMGD de l'ESF pour la mise en concordance des valeurs agrégées des actifs et des passifs dans l'ensemble de données intégrées avec les CBN

	Première mise en concordance Annexe A.1	Étalonnage par itération Annexe A.4	Total (A.1*A.4)
Actifs financiers	1,106	0,852	0,942
Actifs non financiers	0,959	0,907	0,870
Total du passif	1,176	0,959	1,128

Source : Calculs du DPB

La procédure d'étalonnage par itération a été répétée jusqu'à ce que les valeurs agrégées des actifs et des dettes se situent dans une fourchette de 0,00001 % des valeurs correspondantes dans les CBN.

Le DPB a effectué un ajustement proportionnel très limité des actifs financiers, des actifs non financiers et du total du passif de toutes les familles dans l'ensemble de données intégrées afin de mettre entièrement en concordance leurs valeurs agrégées avec celles des CBN.

L'ensemble de données qui en a découlé est la *Base de données sur les familles à valeur nette élevée* (BDFE).

Annexe B : Statistiques sommaires

Les tableaux B-1 et B-2 présentent les statistiques sommaires de la BDFE pour son année de référence 2016. Les résultats de la BDFE se comparent à d'autres précédents dans la littérature : les parts de patrimoine figurant dans le tableau B-1 ressemblent à celles de Davies et Di Matteo (2020); le nombre de familles et leur patrimoine, présentés dans le tableau B-2, sont similaires à ceux de Wealth-X (2017); l'ensemble des résultats de la BDFE concernant les révisions significatives à la hausse des parts supérieures de patrimoine par rapport aux enquêtes nationales sur le patrimoine s'harmonise avec les résultats de Bach et coll. (2015), de Vermeulen (2016) et (2018), et de Davies et Di Matteo (2020).

À titre de référence pour l'interprétation des statistiques sommaires, la BDFE étalonnée compte environ 15 349 000 familles qui possèdent collectivement un patrimoine de 10,3 billions de dollars.

Tableau B-1 Distribution du patrimoine familial, par quantile choisi, Canada, 2016

Quantile du patrimoine familial	Seuil de patrimoine	Nombre de familles	Patrimoine total	Part du patrimoine total
	(M\$)	(milliers)	(G\$)	(%)
0,01 % supérieur	130,5	1,5	574	5,6
0,1 % supérieur	26,7	15,3	1 254	12,1
0,5 % supérieur	8,9	76,7	2 117	20,5
1 % supérieur	5,5	153,4	2 644	25,6
5 % supérieur	2,1	767,5	4 488	43,4
10 % supérieur	1,4	1 534,9	5 829	56,4
20 % supérieur	0,9	3 069,9	7 599	73,5
40 % moyen	0,1-0,9	6 139,7	2 613	25,3
40 % inférieur	Moins de 0,1	6 139,7	128	1,2

Source : Base de données du DPB sur les familles à valeur nette élevée.

Tableau B-2 Distribution du patrimoine familial, par seuil de patrimoine choisi, Canada, 2016

Seuil de patrimoine	Nombre de familles	Patrimoine total	Part du patrimoine total
Familles dont le patrimoine est supérieur à :	(milliers)	(G\$)	(%)
1 G\$	0,1	184	1,8
500 M\$	0,2	277	2,7
250 M\$	0,6	408	3,9
100 M\$	2,2	656	6,3
50 M\$	6,2	925	8,9
25 M\$	16,7	1 287	12,5
10 M\$	63,7	1 994	19,3
5 M\$	173,8	2 751	26,6
2,5 M\$	549,5	3 983	38,5
1 M\$	2 699,0	7 246	70,1

Source : Base de données du DPB sur les familles à valeur nette élevée.

Annexe C : Évolution de la base de données

Les futurs travaux sur la BDFE seront axés sur des sujets revêtant de l'importance pour les parlementaires, la disponibilité de nouvelles sources de données et l'évolution des recherches des universitaires sur le calcul des parts supérieures de patrimoine.

Le DPB souhaite vérifier si la troncature par le haut effectuée dans le FMGD de l'ESF introduit un biais dans l'estimation des paramètres de Pareto. Cette analyse pourrait être réalisée si on établissait une BDFE à l'aide des données de l'ESF provenant d'un centre de données de recherche (CDR) de Statistique Canada, où les observations ne sont pas tronquées par le haut, puis si on comparait le FMGD de l'ESF et les versions de la BDFE établie à l'aide des données de l'ESF provenant d'un CDR.

Statistique Canada a recueilli les données de la prochaine ESF, qui a eu lieu de septembre à décembre 2019 (Statistique Canada, 2019b). On ne sait pas exactement quand le nouveau fichier de microdonnées à grande diffusion sera rendu public. Le DPB envisage d'adapter la BDFE en fonction de la plus récente version publique de l'ESF, qui sera désormais réalisée tous les trois ans.

Aussi à l'avenir, la base de données pourrait voir son évolution freinée par la perte potentielle d'une source de données. La liste des personnes les plus riches utilisée pour établir la BDFE est celle dressée par le magazine CB en 2017, dont les données ont été recueillies en 2016. Bien que le magazine CB ait publié une liste en 2018 (dont les données ont été recueillies en 2017), le DPB n'a pas été en mesure de trouver la liste de 2019. Si le magazine CB cesse de publier tous les ans une liste des personnes les plus riches, le DPB devra se tourner vers d'autres listes, comme celle des milliardaires du monde du magazine Forbes, pour mettre à jour la BDFE.

Enfin, on pourrait améliorer l'approche de modélisation utilisée pour établir la BDFE en réalisant d'autres recherches. Celles-ci pourraient porter sur les sujets d'intérêt suivants : l'établissement d'un meilleur w_{min} , la fixation du seuil de patrimoine auquel l'interpolation de Pareto commence; le perfectionnement de l'approche servant à diviser le patrimoine des familles à valeur nette élevée synthétiques en catégories d'actifs et de passifs constitutifs; l'estimation possible de catégories plus précises liées aux actifs et au passif des familles à valeur nette élevée; l'examen de l'intégration des formes de patrimoine non négociable dans les estimations de la queue supérieure de la distribution du patrimoine familial. Le DPB compte bien suivre les travaux des universitaires pour découvrir de nouvelles théories et méthodologies qui lui permettraient d'améliorer la BDFE.

Références

- Auten, G. et Splinter, D., *Income Inequality in the United States: Using Tax Data to Measure Long-term Trends* (document de travail), 2019. Consulté à l'adresse suivante : http://davidsplinter.com/AutenSplinter-Tax_Data_and_Inequality.pdf. [en anglais seulement]
- Bach, S., Beznoska, M. et Steiner, V., « A Wealth Tax on the Rich to Bring Down Public Debt? Revenue and Distributional Effects of a Capital Levy in Germany », *Fiscal Studies*, vol. 35, n° 1, 2014, p. 67-89. [en anglais seulement]
- Bach, S., Thiemann, A. et Zucco, A., *The Top Tail of the Wealth Distribution in Germany, France, Spain, and Greece*, 2015. Consulté sur le site Web du Deutsches Institut für Wirtschaftsforschung à l'adresse suivante : https://www.diw.de/documents/publikationen/73/diw_01.c.513261.de/dp1502.pdf. [en anglais seulement]
- Brzozowski, M., Gervais, M., Klein, P. et Suzuki, M., « Consumption, income, and wealth inequality in Canada », *Review of Economic Dynamics*, vol. 13, n° 1, 2010, p. 52-75. [en anglais seulement]
- Canadian Business, *Rich 100 Methodology*, 2006. Consulté sur le site Web de Internet Archive à l'adresse suivante : https://web.archive.org/web/20060318021430/http://www.canadianbusiness.com/after_hours/lifestyle_activities/article.jsp?content=20051213_122048_4652. [en anglais seulement]
- Canadian Business, *Canada's Richest People: The Complete Top 100 Ranking (2017)*, 2016. Consulté à l'adresse suivante : <https://www.canadianbusiness.com/lists-and-rankings/richest-people/100-richest-canadians-complete-list-2017/>. [en anglais seulement]
- Catherine, S., Miller, M. et Sarin, N., *Social Security and Trends in Inequality*, 2020 (en cours d'écriture). Consulté sur le site Web de SSRN à l'adresse suivante : <https://ssrn.com/abstract=3546668>. [en anglais seulement]
- Chakraborty, R., Kavonius, I. K., Pérez-Duarte, S. et Vermeulen, P., « Is the Top Tail of the Wealth Distribution the Missing Link between the Household Finance and Consumption Survey and National Accounts? », *Journal of Official Statistics*, vol. 35, n° 1, 2019, p. 31-65. [en anglais seulement]
- Davies, J., « The Distribution of Wealth in Canada », *Research in Economic Inequality*, vol. 4, n° 1, 1993, p. 159-180. [en anglais seulement]
- Davies, J. et Shorrocks, A., « The Distribution of Wealth » dans A. B. Atkinson et F. Bourguignon (dir.), *Handbook of Income Distribution: Volume 1*, New York, NY: North-Holland, 1999, p. 605-675. [en anglais seulement]
- Davies, J., Shorrocks, A. et Lluberá, R., *Global Wealth Databook 2018*, 2018. Consulté sur le site Web du Credit Suisse à l'adresse suivante : <https://www.credit-suisse.com/media/assets/corporate/docs/about-us/research/publications/global-wealth-databook-2018.pdf>. [en anglais seulement]
- Davies, J. et Di Matteo, L., « Long Run Canadian Wealth Inequality in International Context », *Review of Income and Wealth*, 2020. Prépublication en ligne : <https://doi.org/10.1111/roiw.12453>. [en anglais seulement]

Gabaix, X. et Ibragimov, R., « Rank – 1/2: A Simple Way to Improve the OLS Estimation of Tail Exponents », *Journal of Business & Economic Statistics*, vol. 29, n° 1, 2011, p. 24-39. [en anglais seulement]

Gu, W. et Wong, A., *Les estimations du capital humain au Canada : approche fondée sur le revenu de la vie entière* (no 11F0027M au catalogue de Statistique Canada – N° 062), 2010. Consulté à l'adresse suivante : <https://www150.statcan.gc.ca/n1/fr/pub/11f0027m/11f0027m2010062-fra.pdf?st=5C89HjQP>.

Kennickel, A. et Woodburn, R. L., *Consistent Weight Design for the 1989, 1992 and 1995 SCFs, and the Distribution of Wealth* (document de travail), 1997. Consulté sur le site Web de la Réserve fédérale à l'adresse suivante : <https://www.federalreserve.gov/econresdata/scf/files/wgt95.pdf>. [en anglais seulement]

Kennickel, A. et Woodburn, R. L., *Consistent Weight Design for the 1989, 1992 and 1995 SCFs, and the Distribution of Wealth* (document de travail) [matériel supplémentaire], 1997. Consulté sur le site Web de la Réserve fédérale à l'adresse suivante : <https://www.federalreserve.gov/econresdata/scf/files/wgt95app.pdf>. [en anglais seulement]

MacDonald, D., *Born to Win: Wealth concentration in Canada since 1999*, 2018. Consulté sur le site Web du Centre canadien de politiques alternatives à l'adresse suivante : <https://www.policyalternatives.ca/sites/default/files/uploads/publications/National%20Office/2018/07/Born%20to%20Win.pdf>. [en anglais seulement]

Ogwang, T., « Power laws in top wealth distributions: evidence from Canada », *Empirical Economics*, vol. 41, n° 2, 2011, p. 473-486. [en anglais seulement]

Directeur parlementaire du budget, *Estimation des coûts d'une proposition faite en campagne électorale : Impôt sur le patrimoine net*, 2019a. Consulté à l'adresse suivante : https://www.pbo-dpb.gc.ca/web/default/files/Documents/ElectionProposalCosting/Results/326302_02_FR.pdf?timestamp=1568818986141.

Directeur parlementaire du budget, *Performance de l'estimation du coût des promesses électorales 2019*, 2019b. Consulté à l'adresse suivante : https://www.pbo-dpb.gc.ca/web/default/files/Documents/Reports/ADM001/ADM001_fr.pdf.

Saez, E. et Zucman, G., 18 janvier 2019. [Lettre à la sénatrice Elizabeth Warren]. Consulté à l'adresse suivante : <https://gabriel-zucman.eu/files/saez-zucman-wealthtax-warren.pdf>. [en anglais seulement]

Statistique Canada, *Estimations démographiques trimestrielles : octobre à décembre 2016* (n° 91-002-X au catalogue), 2017. Consulté à l'adresse suivante : <https://www150.statcan.gc.ca/n1/fr/pub/91-002-x/91-002-x2016004-fra.pdf?st=6pE9ffng>.

Statistique Canada, *Enquête sur la sécurité financière : Information détaillée pour 2016*, 2018a. Consulté à l'adresse suivante : https://www23.statcan.gc.ca/imdb/p2SV_f.pl?Function=getSurvey&SDDS=2620.

Statistique Canada, *Enquête sur la sécurité financière de 2016 : guide de l'utilisateur pour le fichier de microdonnées à grande diffusion*, 2018b. Tiré de l'Enquête sur la sécurité financière : fichier de microdonnées à grande diffusion de Statistique Canada.

Statistique Canada, *Guide de l'utilisateur : Système canadien des comptes macroéconomiques* (n° 13-606-G au catalogue), 2018c. Consulté à l'adresse suivante : <https://www150.statcan.gc.ca/n1/fr/catalogue/13-606-G>.

Statistique Canada, *Comptes économiques répartis pour le secteur des ménages, estimations des répartitions des actifs, des passifs et de la valeur nette, 2010 à 2018 : rapport technique sur la méthodologie et la qualité* (n° 13-604-M2019001 au catalogue), 2019a. Consulté à l'adresse suivante : <https://www150.statcan.gc.ca/n1/fr/pub/13-604-m/13-604-m2019001-fra.pdf?st=NR7VbCfH>.

Statistique Canada, *Enquête sur la sécurité financière*, 2019b. Consulté à l'adresse suivante : <https://www.statcan.gc.ca/fra/enquete/menages/2620>.

Statistique Canada, *Produit intérieur brut, revenus et dépenses, quatrième trimestre de 2019* [n° 11-001-X au catalogue], 28 février 2020a. Consulté à l'adresse suivante : <https://www150.statcan.gc.ca/n1/fr/daily-quotidien/200228/dq200228a-fra.pdf?st=SjHJHBRW>.

Statistique Canada, *Comptes du bilan national* [tableau : 36-10-0580-01], 2020b. Consulté à l'adresse suivante : https://www150.statcan.gc.ca/t1/tbl1/fr/cv.action?pid=3610058001&request_locale=fr.

Statistique Canada, *Estimations démographiques trimestrielles : octobre à décembre 2019* (n° 91-002-X au catalogue), 2020c. Consulté à l'adresse suivante : <https://www150.statcan.gc.ca/n1/fr/pub/91-002-x/91-002-x2019004-fra.pdf?st=CxSrE1dx>.

Vermeulen, P., « Estimating the Top Tail of the Wealth Distribution », *The American Economic Review*, vol. 106, n° 5, 2016, p. 646-650. [en anglais seulement]

Vermeulen, P., « How Fat is the Top Tail of the Wealth Distribution? », *Review of Income and Wealth*, vol. 64, n° 2, 2018, p. 357-387. [en anglais seulement]

Wealth-X, *World Ultra Wealth Report 2017*, 2017. Consulté à l'adresse suivante : <https://www.wealthx.com/report/exclusive-uhnwi-analysis-world-ultra-wealth-report-2017/>. [en anglais seulement]

Weil, D., « Capital and Wealth in the Twenty-First Century », *The American Economic Review*, vol. 105, n° 5, 2015, p. 34-37. [en anglais seulement]

Notes

¹ Directeur parlementaire du budget (2019b). Dans les quatre mois précédant l'élection fédérale de 2019, le DPB a estimé les coûts de plus de 200 propositions à la demande des partis politiques.

² Directeur parlementaire du budget (2019a). Plus précisément, le DPB a été appelé à estimer les recettes qui découleraient de la perception « d'un impôt [annuel] sur le patrimoine net pour les familles économiques résidant au Canada qui équivaudra à 1 % du patrimoine net au-delà de 20 millions de dollars » pour lequel tous « les actifs et les passifs seront inclus dans l'assiette fiscale de l'impôt sur le patrimoine net, à l'exception des gains de loterie ».

³ Cette définition est identique à celle figurant dans l'Enquête sur la sécurité financière (Statistique Canada, 2018b) et forme le fondement statistique du modèle du DPB dans le présent rapport. Pour les besoins de ce dernier, la définition s'applique également aux termes « valeur nette » et « patrimoine », qui sont utilisés de façon interchangeable.

⁴ De nouvelles recherches sont publiées sur la question de savoir s'il faut inclure les formes de patrimoine non négociable dans l'estimation du patrimoine et des parts de patrimoine des ménages et, dans l'affirmative, comment il faut procéder. Weil (2015) décrit le patrimoine issu du capital humain et des transferts publics comme les deux formes les plus importantes, sur le plan quantitatif, d'« objets de patrimoine » que les calculs du patrimoine négociable ne permettent pas de saisir. Catherine et coll. (2020) se concentrent sur le patrimoine issu des transferts publics; ils développent une approche visant à intégrer le patrimoine de la sécurité sociale dans le calcul du patrimoine des ménages aux États-Unis. Ils constatent que cet ajout atténue les augmentations de l'inégalité de richesse depuis 1989 et réduit les parts supérieures de patrimoine par rapport aux résultats d'autres recherches récentes.

Même si ces autres formes de patrimoine peuvent être moins concrètes et difficiles à calculer en raison de leur nature non négociable, elles représentent également des réserves de valeur importantes au Canada. Gu et Wong (2010) ont estimé le patrimoine de capital humain au Canada en utilisant une approche fondée sur les gains au cours de la vie entière; ils ont constaté qu'en 2007, le stock de patrimoine de capital humain s'élevait à 16,4 billions de dollars. À titre de comparaison, la valeur nette (négociable) du secteur des ménages cette même année, telle qu'elle figure dans les Comptes du bilan national (CBN), n'était que de 6 billions de dollars (Statistique Canada, 2020b).

La sécurité sociale représente également une réserve de valeur importante au Canada. Les CBN incluent dans leur sous-secteur des fonds de sécurité sociale le Régime de pensions du Canada (RPC) et le Régime de rentes du Québec (RRQ) (Statistique Canada, 2018c). À la fin de l'année 2019, la valeur nette de ce secteur était évaluée à 0,5 billion de dollars (Statistique Canada, 2020b). Les CBN n'accordent pas cette valeur nette au secteur des ménages, mais plutôt au secteur des administrations publiques. D'autres programmes de protection sociale de paiement à l'utilisation, comme la Sécurité de la vieillesse (SV) et le Supplément de revenu garanti (SRG), ne sont pas inclus dans le sous-secteur de

la sécurité sociale des CBN parce que ces programmes ne détiennent pas d'actifs accumulés; cependant, ces programmes de transfert constituent sans doute une forme de patrimoine pour les ménages (Catherine et coll., 2020).

⁵ Le PIB du Canada aux prix courants au quatrième trimestre de 2019 s'élevait à 2,3 billions de dollars (Statistique Canada, 2020a).

⁶ Dans le présent rapport, le concept de « famille » équivaut au concept d'« unité familiale » employé par Statistique Canada (2018b). Cela comprend les familles économiques et les personnes ne faisant pas partie d'une famille économique (personnes seules).

Statistique Canada (2018b) définit une famille économique comme étant « un groupe de deux personnes ou plus habitant dans le même logement et apparentées par le sang, par alliance, par union libre ou par adoption ». De même, une personne qui ne fait pas partie d'une famille économique est « une personne qui vit seule ou avec d'autres personnes avec lesquelles elle n'a aucun lien de parenté, par exemple un colocataire ou un pensionnaire ».

⁷ Cette observation sur les familles les plus riches dans le FMGD de l'ESF de 2016 porte sur 965 familles économiques dans la population générale.

⁸ La présente analyse repose sur le Fichier de microdonnées à grande diffusion de Statistique Canada, qui contient des données anonymes recueillies au cours de de l'Enquête sur la sécurité financière de 2016. Tous les calculs réalisés avec ces microdonnées l'ont été par le directeur parlementaire du budget (DPB). Ce dernier accepte toute responsabilité pour l'utilisation et l'interprétation de ces données.

⁹ Pour obtenir de plus amples renseignements, voir *l'Enquête sur la sécurité financière de 2016 : guide de l'utilisateur pour le fichier de microdonnées à grande diffusion*, 2016 (Statistique Canada, 2018b).

¹⁰ Pour obtenir de plus amples renseignements, voir le Guide de l'utilisateur du Système canadien des comptes macroéconomiques (Statistique Canada, 2018c).

¹¹ Le quatrième trimestre de 2016 est le trimestre qui correspond le mieux à la période pendant laquelle les données de l'ESF de 2016 ont été recueillies. En effet, selon Statistique Canada (2019b), les données de l'ESF de 2016 ont été recueillies entre le 9 septembre et le 6 décembre 2016. En outre, comme il est indiqué dans le texte principal, la liste des personnes les plus riches du magazine CB a également été publiée au quatrième trimestre de 2016 (décembre 2016).

¹² Pour obtenir de plus amples renseignements, voir la méthodologie employée par le magazine CB pour dresser la liste des 100 personnes les plus riches (Canadian Business, 2006).

¹³ Auten et Splinter (2019) expliquent qu'il est important de formuler des hypothèses sur l'évolution de la composition des familles lorsqu'on estime les parts supérieures de revenu dans le temps. Leurs données font état de changements différentiels à la composition des familles dans la distribution des revenus (p. ex. en dehors de la partie supérieure de la distribution, on observe une baisse du taux de nuptialité, une diminution de la taille des familles et une augmentation du nombre de ménages monoparentaux). Toutes choses étant égales par ailleurs, on peut s'attendre à ce que ces changements différentiels à la composition des familles au fil du temps entraînent des changements dans la distribution des revenus entre les familles. Il ne serait pas étonnant de constater que les changements différentiels à la composition des familles peuvent également influencer sur les parts supérieures de patrimoine.

¹⁴ L'augmentation du nombre de familles a été calculée approximativement en fonction du taux de croissance démographique figurant dans les Estimations démographiques trimestrielles réalisées par Statistique Canada entre le quatrième trimestre de 2016 et le quatrième trimestre de 2019. Ce calcul approximatif a été nécessaire, car on ne connaissait pas le nombre de familles économiques au quatrième trimestre de 2019 à la date de publication. Il existe un écart de 0,3 % entre les taux de croissance annuels de la population et du nombre de familles économiques depuis 2012.

¹⁵ Statistique Canada (2019a) fournit une excellente explication des différences conceptuelles entre l'ESF et les CBN.

¹⁶ On a calculé les actifs financiers en utilisant les régimes de retraite des employeurs évalués sur la base de la cessation, plutôt qu'en fonction d'une approche de continuité. Statistique Canada (2018b) fournit une description utile de la distinction entre les deux méthodes d'évaluation.

¹⁷ Les variables ont été placées selon la mise en correspondance présentée dans Statistique Canada (2019a), sauf que le DPB a conservé la valeur des objets de collection dans le FMGD de l'ESF.

¹⁸ Vermeulen (2016) avance que cette procédure d'ajustement tend à augmenter de manière disproportionnée le patrimoine des ménages les plus riches, puisque les enquêtes nationales sur le patrimoine ont tendance à sous-déclarer les actifs financiers et que ces derniers représentent une part plus importante du portefeuille des familles les plus riches par rapport à celui des familles les plus pauvres. Cette dynamique est également présente dans les données canadiennes.

¹⁹ Davies (1993), Bach et coll. (2014), Bach et coll. (2015), Vermeulen (2016), Davies et coll. (2017), Vermeulen (2018), Chakraborty et coll. (2019) et Davies et Di Matteo (2020) recourent à une approche qui permet de tirer parti des données provenant des enquêtes sur le patrimoine des ménages et de celles des listes de personnes riches pour estimer la queue supérieure de la distribution du patrimoine familial.

²⁰ Vermeulen (2018) réalise une étude Monte-Carlo pour démontrer qu'il est utile d'ajouter les listes de personnes riches dans les estimations de la queue supérieure d'une distribution. Selon les résultats de son étude, l'ajout d'une liste de personnes riches aux données d'une enquête dans la régression visant à estimer les paramètres de Pareto amène une estimation de la queue supérieure interpolée de la distribution avec un biais à la hausse ou à la baisse de seulement 0,01.

²¹ Davies et Di Matteo (2020) présentent une analyse utile des différences entre la liste du magazine Forbes et la liste des personnes les plus riches du magazine *Canadian Business*, ainsi qu'une comparaison des entrées de chaque liste.

²² Bach et coll. (2014), Bach et coll. (2015), et Davies et Di Matteo (2020) procèdent aussi à l'épuration de la liste des personnes riches avant d'en intégrer les données dans les microdonnées des enquêtes sur le patrimoine national.

²³ Les familles économiques ayant une valeur nette négative dans le FMGD de l'ESF ne sont pas présentées dans la figure A2-1. Selon le FMGD de l'ESF, on comptait 878 482 familles économiques avec une valeur nette négative en 2016.

²⁴ Dans des termes non fondés sur une formule, selon la loi de Pareto appliquée à la distribution du patrimoine familial, le patrimoine de la n ième famille la plus riche de la population est inversement proportionnel à son rang.

²⁵ Canadian Business (2006). La méthodologie a été trouvée sur le site Web d'archivage « Wayback Machine ».

²⁶ Bien que le fichier principal de l'ESF ne soit pas tronqué par le haut, la procédure de pondération de la méthodologie d'enquête peut réduire le poids de certaines familles à valeur nette élevée, même dans le fichier principal. Statistique Canada (2018b) indique, dans le cadre de la procédure de pondération, qu'« on détermine [...] les observations influentes et [qu']on réduit les poids pour un petit nombre d'observations extrêmes ».

²⁷ Vermeulen (2018) explique qu'il ne sait pas très bien où commence la queue supérieure de la distribution du patrimoine de Pareto. Pour faire face à cette incertitude, il utilise dans ses estimations six seuils w_{min} différents allant de 500 000 € à 10 millions d'euros.

²⁸ Information communiquée au DPB dans une correspondance avec des analystes de l'équipe responsable de l'ESF de Statistique Canada.

²⁹ Le FMGD de l'ESF de 2016 comprend 638 observations dont le patrimoine est supérieur à 3 millions de dollars. À l'opposé, dans aucun des pays de l'analyse de Vermeulen (2018) utilisant de l'information géographique ou de l'information géographique stratifiée en fonction du revenu pour suréchantillonner les familles à valeur nette élevée, on ne comptait plus de 100 observations dont le patrimoine était supérieur à 2 millions d'euros.

³⁰ Cette valeur α s'inscrit dans la fourchette des estimations de α dont se servent Davies et Di Matteo (2020) pour imputer les données de la queue supérieure à partir des données historiques des enquêtes sur le patrimoine au Canada.