

Direction des Études et Synthèses Économiques

G 2011 / 09

**L'évaluation de l'activité dissimulée des entreprises
sur la base des contrôles fiscaux
et son insertion dans les comptes nationaux**

Claudie LOUVOT-RUNAVOT

Document de travail



Institut National de la Statistique et des Études Économiques

INSTITUT NATIONAL DE LA STATISTIQUE ET DES ÉTUDES ÉCONOMIQUES

*Série des documents de travail
de la Direction des Études et Synthèses Économiques*

G 2011 / 09

L'évaluation de l'activité dissimulée des entreprises sur la base des contrôles fiscaux et son insertion dans les comptes nationaux

Claudie LOUVOT-RUNAVOT*

MAI 2011

La mise à disposition de l'Insee des résultats des contrôles fiscaux de la Direction générale des finances publiques (Dgfp) a rendu possible la réalisation de cette étude. L'auteur remercie Arnaud GILBERT, en poste au bureau du contrôle fiscal au moment de ses travaux, de l'avoir éclairée sur le contrôle fiscal et sur l'interprétation des données qui en ressortent. Elle remercie également ses relecteurs, Fabrice LENGART, chef du département des comptes nationaux, et Didier BLANCHET, chef du département des études économiques d'ensemble, à l'Insee.

* Département des Comptes Nationaux - Timbre G401 - 15, bd Gabriel Péri - BP 100 - 92244 MALAKOFF CEDEX

L'évaluation de l'activité dissimulée des entreprises sur la base des contrôles fiscaux et son insertion dans les comptes nationaux

Résumé

La mesure de l'économie non observée est nécessaire à la comptabilité nationale pour rendre le contenu du produit intérieur brut aussi exhaustif que possible et comparable d'un pays à l'autre. L'activité que les entreprises dissimulent à l'administration fiscale se dérobe par définition à l'observation statistique. En France, il est possible de l'estimer en s'appuyant sur les contrôles fiscaux, au cours desquels les revenus non déclarés sont en général repérés. En assimilant l'échantillon des vérifications à une enquête, cette fraude observée a pu être extrapolée à l'ensemble du champ. Pour chaque type d'imposition, les probabilités de tirage ont été modélisées par des régressions logistiques, à partir de variables caractérisant les entreprises, telles que la taille ou le secteur d'activité. Les plus explicatives d'entre elles ont ensuite permis de définir une post-stratification sur laquelle se fonde l'évaluation macroéconomique de l'activité fraudée. Cette méthode d'extrapolation tient compte de la surreprésentation dans l'échantillon des entreprises susceptibles d'avoir dissimulé les masses d'impôts les plus importantes. Des variables comptables ont notamment été introduites dans la stratification pour améliorer la modélisation de ce « ciblage », avec pour conséquence d'abaisser sensiblement l'estimation de l'activité dissimulée.

L'objectif était de redresser la production et la valeur ajoutée des comptes nationaux. Parmi les « rectifications » opérées lors des contrôles fiscaux, seules ont été retenues celles qui concernaient les flux comptables réels de l'entreprise. En outre, lorsque les rectifications disponibles portaient sur l'impôt proprement dit, il a été nécessaire de remonter aux bases d'imposition. Cette procédure constitue le principal élément de fragilité du traitement des données de contrôle. Appliqués aux chiffres d'affaires de l'année 2006, les taux de fraude finalement obtenus conduisent à relever de 40 milliards d'euros la valeur ajoutée, soit un peu plus de 2 % du PIB.

Mots-clés : Activité dissimulée (ou fraudée) ; contrôle fiscal ; régression logistique ; post-stratification ; redressement ; comptabilité nationale

Evaluation of firms' dissimulated activity based on fiscal audits and integration in national accounts

Abstract

The measurement of the hidden economy is necessary to ensure exhaustiveness of Gross Domestic Product and thus be able to conduct comparisons between countries. The part of firms activity that is not declared to tax authorities inevitably evades statistical observation. In France, it is possible to evaluate this hidden activity by means of fiscal audits, during which concealed profits may be unveiled. By incorporating a fiscal audit sample into a survey, this type of fraud can then be extrapolated to the whole field. For every category of taxpayer, inclusion probabilities have been estimated with logistic models, making use of categorical predictor variables, like size or activity sector. The more significant of those variables have been used next to define a post-stratification upon which is based the macroeconomic evaluation of the dissimulated aggregates. This extrapolation method takes into account the overrepresentation of the firms that are supposed to conceal the biggest amounts of taxes. In particular, financial indicators have been introduced in the stratification to improve the modeling of this overrepresentation and they lead to a marked decrease in the estimation of the hidden production.

The aim was to rectify National Accounts output and value-added. Among the fiscal rectifications, only those that dealt with the real flows of the firms' book-keeping were taken into account. Moreover, when the available rectifications were concerned with taxes, it was necessary to go back to the corresponding taxable amount. This procedure is the main weakness of the fiscal information processing. Applied to 2006 turnover, the fraud rates finally obtained led to a 40 billion euro rise in the level of national value-added, that is to say a 2.2 % adjustment of GDP

Keywords: Dissimulated activity ; fiscal audit ; logistic regression ; post-stratification ; correction (rectification) ; National Accounts

Classification JEL : E26 ; C25 ; C81 ; E01

Sommaire

Synthèse	5
<i>De la fraude sur l'impôt à l'activité fraudée</i>	5
<i>De la fraude observée sur l'échantillon aux évaluations macroéconomiques</i>	6
<i>Une avancée vers une modélisation plus fine du « ciblage » des contrôles fiscaux</i>	7
<i>Le redressement des agrégats de la comptabilité nationale</i>	7
Introduction	9
<i>L'évaluation de l'activité fraudée des entreprises</i>	9
I - La fraude fiscale et les contrôles fiscaux	13
II - Les objectifs de l'étude, les grandes lignes de la méthode et les données mobilisées	15
<i>II.1 Quels agrégats comptables corriger et comment ?</i>	15
<i>II.2 Les grandes lignes de la méthode d'estimation</i>	16
<i>II.3 Les données mobilisées et leur organisation</i>	17
II.3.1 L'échantillon des contrôles fiscaux de la Dgfp	17
II.3.2 Le fichier Suse-Dgfp	18
III - Définition et contenu des taux de fraude	19
<i>III.1 Évaluation du chiffre d'affaires dissimulé</i>	19
<i>III.2 Évaluation de la fraude sur le résultat</i>	20
III.2.1 Le passage des rectifications en « droits rappelés » aux rectifications « en base »	21
III.2.2 La question des transferts de bénéfices	22
IV - Les aspects statistiques de la méthode d'estimation	24
<i>IV.1 Le processus de sélection de la Dgfp est-il modélisable ?</i>	24
<i>IV.2 La stratification : une représentation du plan de sondage</i>	25
IV.2.1 Une première stratification, croisant le secteur d'activité, la taille et la catégorie juridique...	26
IV.2.2 ...mais des variables comptables seraient pertinentes	26
IV.2.3 Deux méthodes d'investigation : les régressions logistiques et la statistique descriptive	27
IV.2.4 Quelques éléments concrets d'analyse statistique	28
<i>IV.3 Les estimations par post-stratification</i>	32
IV.3.1 Mise en œuvre de la stratification	32
IV.3.2 Interprétation de la post-stratification	34
IV.3.3 Le traitement des rectifications non renseignées	34
IV.3.4 Le changement de nomenclature : passage de la NES à la NAF-REV2	35
V - Les résultats des estimations	36

VI - Simulation du bouclage de la comptabilité nationale	44
<i>VI.1 Le bouclage théorique</i>	44
<i>VI.2 Les simulations et les retraitements spécifiques à la comptabilité nationale</i>	45
<i>VI.3 Le travail au noir</i>	46
Conclusion.....	48
Bibliographie	49
Glossaire.....	50
ANNEXE 1 Annexe descriptive : échantillon, contrôles, taux d'imposition	51
<i>Le fichier Suse-Dgfip</i>	51
<i>L'échantillon par type d'imposition, secteur d'activité et catégorie juridique</i>	51
ANNEXE 2 Description des stratifications	54
ANNEXE 3 Estimation économétrique des probabilités d'inclusion	59

Synthèse

Selon le Système de comptabilité nationale (SEC 1995 et SCN 1993), toute forme d'activité fraudée contribuant à enrichir des agents économiques doit être intégrée dans le calcul du produit intérieur brut, qu'il s'agisse de l'activité non déclarée des entreprises juridiquement reconnues ou d'un « travail au noir » plus clandestin. En France, le compte des entreprises étant principalement établi à partir des liasses fiscales, il n'englobe pas l'activité dissimulée au fisc. Il doit donc être redressé à ce titre, et l'on mobilise dans ce but les résultats des contrôles fiscaux effectués par la Direction générale des finances publiques. Cette méthode de correction de l'activité fraudée est dite « directe », puisqu'elle se fonde sur l'exploitation de données microéconomiques.

Les contrôles fiscaux s'appuient sur une expertise fouillée de la comptabilité des entreprises vérifiées, au cours de laquelle les revenus dissimulés sont repérés. Moyennant certaines précautions statistiques, il est possible d'extrapoler, à l'ensemble du champ concerné de la comptabilité nationale, l'activité fraudée mesurée sur un échantillon de vérifications.

Les contrôles fiscaux apportent toutefois sur l'activité dissimulée un éclairage particulier et ne suffisent pas à combler un vide d'information, inhérent à la collecte statistique sur l'activité économique. Le travail au noir des agents clandestins, en particulier, n'est pas dans le champ d'action de la Dgfi et reste mal connu, faute de sources. De plus, l'exploitation des résultats des contrôles fiscaux se heurte à de nombreuses difficultés statistiques, liées au « ciblage » de l'échantillon des entreprises vérifiées. Ce dernier n'est pas un échantillon d'enquête et ne peut donc pas être utilisé comme tel, sans précautions particulières. Les services fiscaux recherchent les entreprises susceptibles d'avoir dissimulé les masses d'impôts les plus importantes, leur but étant de recouvrer un maximum des impôts non versés. Cette démarche est ainsi orientée vers le repérage de cas d'espèce, particulièrement fraudeurs, alors qu'un échantillon d'enquête serait au contraire représentatif de l'ensemble des comportements de fraude des entreprises françaises.

L'évaluation macroéconomique de la fraude sur la base des contrôles fiscaux est une figure imposée de la comptabilité nationale, périodiquement reprise et actualisée à l'occasion des changements de bases successifs. La réévaluation présentée dans ce document a été l'occasion d'une refonte complète de la méthode d'estimation et de la manière d'intégrer dans les comptes les redressements ainsi obtenus. Ces changements s'imposaient du fait d'une importante réorganisation du système d'enregistrement de la Dgfi. Par ailleurs, pour la première fois, les résultats des contrôles ont pu être mis en regard des données comptables des entreprises vérifiées, ce qui a ouvert de nouvelles pistes d'investigation statistique.

De la fraude sur l'impôt à l'activité fraudée

Lors des contrôles fiscaux, le vérificateur procède à des corrections, dénommées « rectifications », portant, soit sur les bases d'imposition, c'est-à-dire sur certains flux comptables de l'entreprise, soit sur ses impôts. Ces rectifications sont répertoriées par référence au « thésaurus des contrôles fiscaux », selon des « chefs de rectification » définis sur la base du Code général des impôts. Les malversations comptables les plus courantes prennent la forme d'omissions de recettes, de majorations de frais généraux ou de charges, ou de falsifications d'amortissements et de provisions. La fraude sur l'impôt ne découle toutefois que partiellement de ces manipulations ; elle résulte fréquemment d'une mauvaise application de la législation fiscale, sans que les flux comptables sous-jacents soient affectés.

La finalité de l'exploitation statistique des données de contrôle est d'aboutir à une évaluation macroéconomique des parts du chiffre d'affaires et du résultat qui sont masquées par les entreprises dans le but de réduire le montant de leurs impôts. Il s'agit donc d'extrapoler des rectifications portant sur la base de calcul de l'impôt elle-même, et non pas sur les « droits rappelés » de l'impôt proprement dit. En outre, seule une partie des rectifications opérées par la Dgfi est à prendre en considération. En premier lieu, les rectifications uniquement

imputables à une application erronée de la législation fiscale doivent être éliminées, les comptes de l'entreprise n'étant pas, dans ce cas, falsifiés. Les rectifications portant sur les amortissements et sur les provisions, au demeurant fréquentes, sont également à exclure : les premières parce que les amortissements n'interviennent pas dans le calcul des soldes bruts de la comptabilité nationale, les secondes parce que les provisions n'ont pas d'équivalent en comptabilité nationale. Enfin, les rectifications portant sur les transferts de bénéfices des sociétés multinationales ont été écartées pour des raisons de cohérence avec l'information statistique disponible sur les échanges extérieurs. Les autres « chefs de rectification » ont été retenus, principalement les omissions de recettes, c'est-à-dire le chiffre d'affaires dissimulé, et la rectification globale du résultat, après déduction des rectifications mentionnées ci-dessus.

Une des difficultés des calculs tient à l'impossibilité technique de ne mobiliser que des bases d'imposition, ce qui serait l'idéal compte tenu de l'objectif visé. La rectification du résultat, en particulier, n'est disponible qu'en « droits rappelés ». Pour la traduire en base, il a été nécessaire de procéder à une estimation du taux d'imposition apparent moyen de l'entreprise rectifiée ; mais cette estimation reste fragile. L'utilisation conjointe de rectifications en base et en droits rappelés constitue le principal élément de fragilité du traitement des données de contrôle, dont témoigne la sensibilité des résultats aux hypothèses retenues pour passer des secondes aux premières.

L'échantillon des contrôles fiscaux a été constitué par regroupement de trois campagnes de contrôles annuelles, ce qui améliore sa représentativité. Les vérifications s'étalent de la fin des années 90 à 2006. L'échantillon comporte un peu plus de 117 000 enregistrements. Il a été extrapolé à la population de Suse 2006¹, laquelle compte plus de 2 600 000 entreprises sur le champ considéré.

De la fraude observée sur l'échantillon aux évaluations macroéconomiques

Projeter à l'ensemble des entreprises l'activité fraudée mesurée sur l'échantillon des contrôles fiscaux est un exercice qui s'apparente à l'extrapolation d'un fichier d'enquête. Le principe est d'étendre à toute une population les comportements de fraude observés lors des vérifications, en appliquant des coefficients multiplicatifs à la fraude constatée. Les services fiscaux visant plus particulièrement les entreprises supposées être les plus fraudeuses, ces dernières sont, en proportion, plus nombreuses que d'autres dans l'échantillon et doivent par conséquent peser d'autant moins dans l'extrapolation. D'une manière générale, plus la probabilité de tirage d'une entreprise est forte, plus faible doit être le poids de la fraude observée de cette entreprise dans ce calcul.

Pour extrapoler la fraude, il faut donc commencer par affecter à chaque entreprise une probabilité de tirage. Comme on ne la connaît pas, il faut l'estimer. Cette estimation doit être aussi conforme que possible à la probabilité de tirage réelle pour éviter des biais lors de l'extrapolation. C'est là que réside la principale difficulté de l'exercice.

La méthode la plus couramment mise en œuvre, et utilisée ici, se fonde sur une stratification a posteriori. Les entreprises supposées avoir des comportements de fraude similaires, tant celles de l'échantillon que celles de la population, sont regroupées en strates. Pour définir les strates, on relie ces similitudes de comportement à certaines caractéristiques des entreprises. On sait, par exemple, que les masses de fraude sont plus importantes pour les entreprises fraudeuses réalisant de gros chiffres d'affaires, que les petites entreprises individuelles ont une plus forte propension à frauder, que certains secteurs d'activité sont réputés plus fraudeurs que d'autres, etc. Des méthodes d'analyse statistique permettent de déterminer, de manière relativement objective, les liens entre la fraude et certaines caractéristiques des entreprises ; elles font ainsi émerger les variables de stratification les plus significatives.

¹ Suse : Système Unifié de Statistique d'Entreprises, de l'INSEE.

Ensuite, au sein de chaque strate ainsi définie, la probabilité de tirage des entreprises est supposée uniforme et égale au rapport entre les effectifs de cette strate, dans l'échantillon et dans la population, respectivement. On admet également que les propensions à frauder sur le chiffre d'affaires ou sur le résultat sont uniformes sur chaque strate ; elles se définissent par des taux de fraude, calculés sur l'échantillon, en rapportant les montants de fraude aux chiffres d'affaires. L'extrapolation consiste à appliquer ces taux observés aux chiffres d'affaires des strates correspondantes de la population. On obtient les évaluations macroéconomiques recherchées par agrégation des montants de fraude des différentes strates.

Une avancée vers une modélisation plus fine du « ciblage » des contrôles fiscaux

Depuis que la fraude est intégrée dans les comptes nationaux, elle a toujours été extrapolée par post-stratification. Les variables explicatives traditionnellement retenues sont le secteur d'activité, la catégorie juridique et la taille, mesurée par tranches de chiffre d'affaires. Cette stratification est à peu près incontournable et conduit à une représentation convenable du tirage. Elle a été reprise, dans un premier temps, pour effectuer les estimations de la base 2005. Toutefois, le fait d'avoir pu rapprocher les contrôles fiscaux des comptes des entreprises a rendu possible une modélisation plus fidèle du « ciblage » des contrôles. On sait qu'au cours de leur processus de sélection les services fiscaux utilisent des ratios comptables pour pointer des incohérences laissant présager un comportement de fraude. D'où l'idée d'introduire dans la stratification, en plus des variables traditionnelles, deux ratios comptables courants : le taux de valeur ajoutée et le taux de marge. Ces ratios se sont révélés pertinents au vu des régressions et des analyses statistiques, et leur impact sur les résultats macroéconomiques en témoigne également : si ces ratios n'avaient pas été introduits dans la post-stratification, la fraude sur le chiffre d'affaires serait surestimée de plus de 40 % et la fraude sur le résultat de 20 %. Pour toutes ces raisons, les évaluations finalement retenues pour la base 2005 s'appuient sur cette stratification enrichie.

Cette modélisation du ciblage des services fiscaux n'en demeure pas moins relativement fruste ; elle ignore, en particulier, des critères de sélection relevant typiquement de l'appréciation humaine. Mais elle est moins réductrice qu'il n'y paraît. La démarche statistique intervient en effet *a posteriori* : les entreprises contrôlées sont connues, ce qui permet de les décrire et de les comparer aux autres. Par une approche inductive, fondée sur l'observation, il est alors possible de repérer, parmi les éléments statistiques dont on dispose, ceux qui expliquent le mieux les probabilités de tirage. On peut ainsi parvenir à simuler le processus de sélection de manière fidèle, tout en le simplifiant énormément.

Le redressement des agrégats de la comptabilité nationale

Pour redresser le compte des entreprises de la comptabilité nationale, on dispose de deux rectifications macroéconomiques : une rectification portant sur le chiffre d'affaires, de l'ordre de 20 milliards d'euros, et une rectification relative à l'excédent brut d'exploitation (EBE), de l'ordre de 40 milliards d'euros. Cette dernière s'applique en réalité à la valeur ajoutée puisqu'il n'y a pas d'interventions sur les salaires lors des contrôles fiscaux. Ainsi, à la production des comptes nationaux, on ajoute le chiffre d'affaires dissimulé, c'est-à-dire la rectification du chiffre d'affaires pour omissions de recettes : $\Delta (P) = \text{rect} (CA) = 20$ milliards d'euros. À la valeur ajoutée des comptes, on ajoute la rectification de l'EBE : $\Delta (VA) = \text{rect} (EBE) = 40$ milliards d'euros. La nature des rectifications retenues² conduit à corriger les consommations intermédiaires de la comptabilité nationale de la différence entre la correction de la production et celle de la valeur ajoutée : cette correction est donc négative et égale à - 20 milliards d'euros. Les rectifications opérées sur les charges lors des contrôles fiscaux n'ont pas été mobilisées pour l'évaluer.

² Sont exclues, les rectifications de nature purement fiscale, celles qui touchent aux amortissements et aux provisions et les transferts de bénéfices.

Conceptuellement, en comptabilité nationale, la correction apportée à la valeur ajoutée ne doit pas être entièrement reportée sur l'excédent brut d'exploitation (EBE). En effet, la fraude sur l'activité va souvent de pair avec une dissimulation partielle des heures travaillées. On suppose, en première approche, que la production dissimulée est effectivement le fruit d'un travail dissimulé. En admettant que la productivité de ce travail dissimulé est identique à la productivité moyenne de l'entreprise, il devient possible de calculer la masse salariale dissimulée - c'est-à-dire le « travail au noir » effectué au sein des entreprises juridiquement déclarées - en appliquant des taux de salaire³ sectoriels à la valeur ajoutée dégagée par la production dissimulée de chaque secteur. Pour évaluer cette valeur ajoutée, on admet que le chiffre d'affaires dissimulé n'induit aucune correction sur les achats de l'entreprise ; les consommations intermédiaires afférentes étant inchangées, la valeur ajoutée issue de la production dissimulée est alors égale à cette dernière.

Au niveau de l'entreprise, il est peu probable que la fraude par omission de recettes coexiste avec la fraude par majoration induite des charges, même si une telle situation ne peut pas être totalement éliminée. D'un point de vue macroéconomique, on peut schématiser la réalité en considérant qu'il y a deux grandes catégories d'entreprises fraudeuses, au sens des comptes nationaux : celles qui fraudent sur le chiffre d'affaires et celles qui ne le font pas, mais fraudent en majorant indûment leurs charges. Sous cette hypothèse, la masse salariale dissimulée n'est le fait que des premières, pour lesquelles les consommations intermédiaires ne sont pas corrigées (*cf. supra*). La correction macroéconomique des consommations intermédiaires ne concerne donc que les entreprises fraudant par majoration de charges (- 20 milliards d'euros) (tableau 16).

In fine, l'intégration dans les comptes nationaux de l'activité dissimulée par les entreprises juridiquement reconnues comporte une manipulation comptable supplémentaire : elle vise à ne pas modifier la masse salariale distribuée par les entreprises, telle qu'elle ressort des liasses fiscales (SUSE). Pour ce faire, on considère que la masse salariale dissimulée est assimilable à un service, produit par des Entreprises individuelles (EI) virtuelles et consommé intermédiairement par les Sociétés non financières (SNF) et par les Entreprises individuelles juridiquement reconnues qui y ont recours. Dans ces conditions, cette masse salariale dissimulée doit être ajoutée, respectivement, aux corrections (initialement négatives) des consommations intermédiaires des SNF et des EI et ajoutée, par ailleurs, à la production des EI (il s'agit de la production d'EI « virtuelles »). Ce traitement comptable de la masse salariale dissimulée par des entreprises juridiquement reconnues n'a pas d'impact sur l'évaluation de la valeur ajoutée dissimulée, mais il en modifie le partage entre SNF et EI, à l'avantage de ces dernières et il gonfle la production dissimulée ; enfin, l'EBE n'est pas non plus corrigé globalement puisque le travail dissimulé a été comptabilisé en consommations intermédiaires.

Outre le travail non déclaré par les entreprises juridiquement reconnues, mentionné ci-dessus, il existe une autre forme de travail au noir : le travail totalement clandestin (ateliers clandestins, petits travaux domestiques pour l'entretien du logement, etc.) ; on l'attribue, conventionnellement, à des entreprises individuelles n'ayant pas d'existence juridique. Il doit aussi être intégré dans le calcul du PIB, mais il ne rentre pas dans le champ de cette étude. Pour donner un ordre de grandeur, en l'estimant à l'aide des taux de travail clandestin fixés par la base 2000, on fait passer le redressement total de la valeur ajoutée pour 2006 de 40,7 milliards d'euros à 55,1 milliards d'euros.

³ Le taux de salaire se définit comme la part des salaires dans la valeur ajoutée.

Introduction

L'évaluation de l'activité fraudée des entreprises

Depuis le début des années soixante, l'INSEE s'appuie sur les résultats des contrôles fiscaux effectués par la Direction générale des impôts (Dgi) - devenue en 2007 la Direction générale des finances publiques (Dgfp) - pour procéder à une évaluation macroéconomique de l'activité fraudée. Ce calcul est une figure imposée de la comptabilité nationale, périodiquement reprise et actualisée à l'occasion des changements de base successifs. L'objectif est de contribuer à rendre conforme aux normes internationales le compte des entreprises, en y introduisant l'activité non déclarée par ces dernières à l'administration fiscale.

Qu'il s'agisse du travail dissimulé d'une entreprise juridiquement reconnue ou d'un travail au noir, licite, mais clandestin, toute forme d'activité contribuant à enrichir des agents économiques doit en principe être intégrée dans le calcul du produit intérieur brut. C'est du moins ce que préconise le Système de comptabilité nationale (SEC 1995 et SCN 1993).

C'est en premier lieu pour satisfaire à ces recommandations conceptuelles que les comptables nationaux tentent d'estimer la fraude mais ils ont aussi des raisons techniques pour le faire, liées à l'algorithme de construction du produit intérieur brut. L'activité fraudée est de fait l'une des causes de divergence entre les trois approches spontanées du PIB (par la demande, la production ou les revenus) : chacune de ces mesures se fondant sur des sources spécifiques, les conséquences de la fraude sur les agrégats n'y sont pas prises en compte de la même manière. Par exemple, une activité industrielle dissimulée sera ignorée si le PIB est calculé par la valeur ajoutée puisque les flux comptables correspondants ne figurent pas dans la source mobilisée à cet effet, en l'occurrence la liasse fiscale. En revanche, si les revenus afférents à cette activité sont consommés ou investis dans l'économie nationale, ils peuvent se trouver inclus dans les données sur lesquelles s'appuie l'estimation du PIB par la demande.

Certains pays vont jusqu'à considérer que l'écart entre l'approche « production » du PIB et l'approche « demande » est uniquement dû à l'activité fraudée et c'est ainsi qu'ils estiment cette dernière (*encadré 1*). Cette mesure de la fraude, dite « indirecte », est extrêmement fruste ; la méthode de construction des comptes français ne permettrait d'ailleurs pas de la mettre en œuvre.

En France, pour évaluer la fraude, on utilise une méthode dite « directe », dont le principe est de l'estimer à partir de données microéconomiques. On sait que l'activité dissimulée ne figure pas, par définition, dans la liasse fiscale sur laquelle se fonde, en majeure partie, le calcul du PIB par la valeur ajoutée. Ce n'est donc qu'après avoir redressé le compte de secteur des entreprises de l'activité dissimulée, estimée de manière exogène, que sont équilibrées les différentes approches du PIB.

La fraude défiant par nature la collecte statistique, la mesurer est un exercice délicat qui ne peut reposer que sur une information lacunaire, parfois inadaptée. C'est pourquoi les évaluations macroéconomiques de ses différentes composantes restent fragiles, lorsqu'elles existent. La fraude sur les prélèvements fiscaux est l'une des mieux appréhendée. Des campagnes de contrôles sont en effet régulièrement menées par la Dgfp pour connaître le montant des impôts non perçus, en vue de les recouvrer. Ces contrôles s'appuient sur une expertise fouillée de la comptabilité des entreprises vérifiées. En particulier, les revenus que les entreprises n'ont pas déclarés au fisc sont décelés lors des vérifications, la fraude sur l'impôt passant en partie par une dissimulation d'activité. À partir du résultat des contrôles fiscaux, il est alors possible d'extrapoler à l'ensemble des entreprises l'activité non déclarée au fisc. Ainsi, les comptables nationaux disposent d'un gisement de données particulièrement bien adapté à leurs besoins puisqu'il leur permet de redresser la liasse fiscale, et par conséquent le compte des entreprises qui en découle.

On remarquera toutefois que les contrôles fiscaux apportent sur l'activité dissimulée un éclairage particulier ; ils ne suffisent pas, loin s'en faut, à combler un vide d'information qui demeure très important en l'état actuel du système statistique. Le travail au noir clandestin notamment, c'est-à-dire l'activité des agents économiques qui ne sont pas identifiés comme des producteurs, n'est pas dans le champ d'action de la Dgfp et reste mal connu. En outre, l'exploitation des données de contrôle se heurte à de nombreuses difficultés statistiques. En premier lieu, le contenu de l'information répertoriée se révèle parfois inadapté, celle-ci étant destinée à mesurer l'impôt et non pas les flux comptables sous-jacents. Cette information n'est pas non plus collectée dans le but d'être étendue à toutes les entreprises : la Dgfp constate mais son objectif n'est en aucun cas de projeter. Or, le principe de base d'une évaluation macroéconomique de l'activité dissimulée consiste à extrapoler à l'ensemble du champ la fraude des entreprises contrôlées. Se posent donc également des problèmes d'échantillonnage, d'estimation et d'extrapolation. Ils sont présentés dans ce document dont l'objet est principalement de décrire les méthodes d'estimation, les difficultés rencontrées pour les mettre en œuvre et la manière dont elles ont été résolues.

Encadré 1

L'activité non observée dans le monde : définitions et mesures (« non observed- economy » ou NOE)

La mesure de l'économie non observée s'impose désormais comme une nécessité

L'activité non déclarée à l'administration fiscale, celle d'un atelier clandestin, la rénovation d'un appartement « au noir », la contrebande ou le baby-sitting sont autant d'activités qui ont en commun le fait de se dérober à l'observation statistique, tout en participant à la création de richesse d'une économie et en donnant lieu à des paiements, puisque ce sont des activités marchandes. La variété des vocables utilisés pour désigner cette économie obscure témoigne du manque de précision de ses contours : on parle tantôt d'économie souterraine, d'économie informelle, d'économie parallèle, tantôt d'économie de l'ombre ou d'économie non observée, etc. Pour autant, en ignorer l'existence dans le calcul du PIB conduirait à exclure de la description et de la mesure des performances économiques des pays des pans entiers de leur activité, d'importance inégale, ce qui fausserait les comparaisons internationales. Selon une enquête de l'Organisation des Nations Unies (ONU), au début des années 2000, la part de l'économie non observée dans le PIB aurait atteint les 30 % au Kazakhstan, en Russie elle aurait avoisiné les 25 %, en Italie, en Pologne, en Lettonie, elle serait de l'ordre de 15 %, au Brésil de 13 %, en République Tchèque de 9 % alors qu'on l'estime à 3 % en France et à moins encore dans les pays d'Europe du nord (*graphique 1*). Évalué sans tenir compte de l'économie de l'ombre, le niveau du PIB serait assurément sous-estimé. Or, celui-ci intervient au dénominateur de différents ratios très utilisés par les pouvoirs publics pour la conduite des politiques économiques : le déficit et la dette publics, les émissions de dioxyde de carbone, les dépenses de recherche et développement, sont rapportés au PIB, etc. La contribution des États membres au financement de l'Union européenne est aussi calculée en pourcentage du RNB (un agrégat directement lié au PIB). Le PIB en niveau par habitant constitue également un indicateur de prospérité très usité. Par ailleurs, si l'économie souterraine venait à régresser ou au contraire à se développer, on fausserait la mesure de la croissance en l'omettant du PIB. L'émergence d'activités nouvelles, notamment dans les secteurs des services, les contraintes de l'harmonisation européenne et, plus généralement, l'internationalisation de l'économie rendent de plus en plus prégnante la nécessité de mesurer les économies informelles. Ces dernières se développent également chez les géants mondiaux, tels que la Chine, l'Inde ou le Brésil, parallèlement à d'importantes mutations structurelles. Selon certaines études, notamment celles de Schneider et Enste (*FMI - 2000*), la part de l'économie souterraine se serait accrue dans les pays de l'OCDE au cours des trois dernières décennies. Cette évolution serait liée à l'accroissement de la pression fiscale et à la mise en place de systèmes de Sécurité sociale : l'écart entre le coût du travail et la rémunération nette des salariés - le coin fiscal-social - se creusant, la situation devenait propice à la progression du travail clandestin.

Un concept multiforme

L'exhaustivité du PIB n'est pas une nouveauté. C'est au contraire un concept de base de la Comptabilité nationale ; mais au sens du SCN de 1993, les contours de l'activité économique sont très larges, ce qui rend plus exigeante la recherche d'exhaustivité. En plus de l'activité traditionnelle, aisément quantifiable, ils englobent, entre autres, les activités illégales dès lors qu'elles sont volontairement exercées par les acheteurs et les vendeurs, l'autoproduction des ménages, les activités non déclarées à l'administration fiscale ou aux offices statistiques. Pour aider les statisticiens à mesurer aussi correctement que possible l'économie non observée, l'Organisation de coopération et de développement économique (OCDE) a publié en 2002 un manuel visant à définir la

nature de la dissimulation d'activité et présentant des méthodes d'estimation. Plus récemment, en 2005 - 2006, pour faire l'inventaire des méthodes de mesure actuellement en vigueur, l'Organisation des Nations Unies (ONU) a conduit une enquête dans 43 pays, parvenus à des stades différents de leur développement (cf. *bibliographie*).

L'OCDE distingue trois types « d'économie non observée » (ENO) :

- L'économie « souterraine », dite aussi économie dissimulée ou cachée ou « au noir » : il s'agit d'activités exercées dans la légalité mais que les producteurs dissimulent délibérément à l'administration fiscale pour essayer de se soustraire à l'impôt ou à d'autres formes de prélèvements, tels que les charges sociales. Sous cette rubrique se trouvent aussi les activités qui ne sont pas enregistrées dans les sources statistiques, faute de réponse aux questionnaires ou de remplissage des formulaires. Au sein de « l'économie au noir », les comptables nationaux français distinguent les unités de production juridiquement reconnues des unités de production clandestines.
- L'économie informelle : cette rubrique recouvre des activités parfaitement légales mais exercées dans un cadre peu organisé, opérant à petite échelle, sans véritable identification de la part de la production imputable au travail ou au capital (les dépenses de production sont souvent inséparables de celles du ménage exerçant l'activité et les locaux de « l'entreprise » sont situés dans l'habitation de ce dernier, etc.). Le secteur informel est fondé sur un système de relations d'affaires non conventionnel, de sorte qu'il échappe par nature à la collecte statistique.
- Les activités illégales : trafic de drogue, vol, contrebande, etc.

Des méthodes de mesure insuffisamment harmonisées et des résultats fragiles

Selon les situations qui sont à l'origine de la dissimulation d'activité, elle sera plus ou moins facilement récupérable. Par définition, les activités illégales sont les plus difficiles à repérer, faute de sources. Rares sont les pays qui en tiennent compte dans le calcul de leur PIB, mais une quinzaine procèdent néanmoins à des tentatives d'estimation à l'aide de statistiques détenues par la police, sur la base d'estimations d'expert ou, pour le trafic de drogue, à partir d'informations sur les quantités saisies par les autorités et sur la base des statistiques sur le nombre d'usagers établies par les centres de soin.

Ces cas mis à part, plusieurs méthodes sont employées pour essayer d'estimer l'économie souterraine. On distingue deux types d'approches : les approches, dites directes ou microéconomiques, qui se fondent sur des enquêtes ou sur des données fiscales, et les approches indirectes ou macroéconomiques, qui s'appuient sur la confrontation de différents agrégats.

L'exploitation des contrôles fiscaux est une méthode directe ; elle compte parmi les plus robustes, sans être pour autant la plus répandue. Néanmoins, le Canada, les États-Unis, l'Italie, et bien sûr la France, l'utilisent ou l'ont testée. La « méthode des flux de marchandises » en est une autre ; elle consiste à estimer la production brute et la valeur ajoutée dissimulées en appliquant des taux de marge aux « flux de marchandises » qui alimentent le processus de production. Certains pays (le Canada, les États-Unis) mobilisent également des résultats d'enquête ; ainsi, à l'occasion d'enquêtes sur les dépenses des ménages, les personnes interrogées sont invitées à s'exprimer sur d'éventuelles activités cachées.

Les méthodes indirectes, dites de « macro-modélisation », sont les plus employées. L'une d'elles, « l'approche monétaire », part de l'hypothèse que les transactions de l'économie de l'ombre se règlent exclusivement en espèce. Ainsi, la part de l'augmentation du ratio entre avoirs liquides et dépôts qui ne s'explique pas par les taux d'intérêt, la hausse des revenus ou l'évolution des habitudes, entre autres, est due à la croissance de l'économie de l'ombre. La méthode « par la demande d'électricité » consiste à supposer que l'économie totale, incluant l'économie souterraine, augmente parallèlement à la consommation d'électricité. Dans les deux cas, on estime l'évolution de l'économie non observée, mais il faut au départ la calibrer. Une autre approche se fonde sur la comparaison des agrégats de la comptabilité nationale, partant de l'hypothèse que l'économie souterraine est spontanément incluse dans la masse des revenus ou des dépenses alors qu'elle ne l'est pas dans la production, ni dans la mesure de la valeur ajoutée qui en découle. En France une telle méthode ne serait pas applicable (cf. *supra*). Les trois approches macroéconomiques ont en revanche été testées par plusieurs grands pays (Canada, États-Unis, Allemagne, Grande-Bretagne, Italie). La « méthode de la main d'œuvre utilisée » compare les effectifs enregistrés dans les fichiers administratifs à ceux des sources déclaratives. La disparité des résultats issus de ces différentes méthodes témoigne de leur fragilité. Selon Schneider et Enste (2000), en moyenne sur la période 1986-1990, le Canada évalue à 2,9 % la part de l'économie souterraine dans le PIB par l'audit fiscal, à 12,0 % par la demande de monnaie, à 1,2 % par la demande d'électricité. Aux États-Unis, les chiffres correspondants sont moins dispersés : respectivement 10,0 %, 6,2 %, et 9,9 % ; par la comptabilité nationale, le ratio américain serait de 10,2 %. En Italie, les quatre évaluations correspondantes sont, dans l'ordre, de 10,0 %, 21,3 %, 19,3 % et 9,3 %.

Graphique 1 : La part de l'économie non observée dans le PIB de différents pays



Source : ONU (2008) - *Non-observed economy in National Accounts - Survey of country practices*

En dépit des efforts menés par les organismes internationaux et par les pays eux-mêmes pour évaluer l'économie non observée (ENO), la diversité des approches méthodologiques et conceptuelles rend les comparaisons internationales peu fiables. *Selon l'OCDE, dans l'acception la plus large du terme, l'ENO regroupe toutes les activités qui échappent aux sources sur lesquelles se fonde la construction des comptes.*

Dans les pays en transition, la couverture et la qualité des sources s'améliorent à mesure que se construit leur système statistique. Une part de l'activité non observée s'intègre alors progressivement dans les procédures de mesure classiques, de sorte que l'évolution de la part de l'ENO peut aussi bien correspondre à une réalité qu'à une amélioration du système statistique.

Selon la manière dont est mesurée l'ENO, on distingue en général trois groupes de pays : ceux qui se conforment strictement aux recommandations d'Eurostat (Nouveaux pays de l'Union européenne et pays candidats, certains pays de la Communauté des États indépendants) ; ceux qui font des efforts approfondis et systématiques pour assurer l'exhaustivité du PIB sans pour autant procéder régulièrement à une mesure de l'ENO (USA, Canada, Australie, France, Allemagne, Royaume-Uni, Irlande, Pays-Bas, Norvège, Finlande) ; ceux qui ne procèdent qu'à des mesures partielles ou ont recours à de méthodes très spécifiques, tels que l'Italie.

Ces remarques incitent à interpréter avec la plus grande prudence les ratios du graphique 1. Ils ne sont mentionnés que pour indiquer des ordres de grandeurs. Les méthodes d'estimation sont très hétérogènes, les définitions de l'économie non observée ne sont pas totalement harmonisées, ni les modes de calcul des ratios ; beaucoup de pays fournissent d'ailleurs plusieurs évaluations. Les dates d'observation ne sont pas non plus les mêmes, bien qu'elles se situent en général au début des années 2000 ou à la fin des années 90. Un certain nombre de « grands pays » n'ont pas communiqué à l'ONU un ratio synthétique représentatif de la part de l'ENO dans le PIB, mais seulement des évaluations partielles et des indications méthodologiques ; c'est la raison pour laquelle ces pays ne figurent pas sur le diagramme. Les ratios du graphique 1 ne sont pas non plus comparables à ceux de Schneider et Enste mentionnés ci-dessus, principalement parce qu'ils se réfèrent une période d'observation postérieure de plus de dix ans ; en outre, comme ils ont aussi été établis plus récemment, les méthodes et les sources ont eu le temps d'évoluer.

I - La fraude fiscale et les contrôles fiscaux

Les contrôles fiscaux sont effectués sur tout le territoire par des inspecteurs de la Dgfi. Les données transmises à l'INSEE sont issues de contrôles effectués sur le terrain ; au cours de ces contrôles, le vérificateur passe au crible la comptabilité de l'entreprise dans le but de repérer des malversations qui se seraient traduites par une minoration de l'impôt versé à l'État. Ce travail d'expertise l'amène parfois à reconstruire une partie des comptes de l'entreprise à partir d'éléments externes. Toutes les rectifications doivent être étayées par des preuves. Le domaine d'intervention du vérificateur se limite aux flux comptables susceptibles d'avoir un impact sur l'impôt. La masse salariale, en particulier, ne fait l'objet d'aucune intervention, même si des pratiques apparentées au « travail au noir » sont repérées. Il peut arriver que des frais généraux ou des avantages en nature, qui constituent pour les cadres ou les dirigeants de l'entreprise une forme de rémunération déguisée, soient contestés lors de la vérification ; toutefois, en pareil cas, la rectification ne porte pas sur la masse salariale, mais sur les frais de personnel. Sur la base de la comptabilité ainsi « rectifiée », le vérificateur calcule le montant des impôts que l'entreprise était censée payer et il en déduit le montant des impôts non perçus. Les sociétés prises en défaut doivent acquitter les sommes dues et, dans certains cas, payer des pénalités. Un contrôle fiscal porte sur six mois à trois ans d'exercice, voire plus, et la durée de la vérification se compte en mois (*Annexe 1*).

L'impôt non perçu est la résultante d'une ou plusieurs rectifications, selon les causes qui sont à l'origine de la fraude, et elles sont multiples. Une entreprise peut, par exemple, dissimuler une partie de son chiffre d'affaires ; ainsi, elle réduit son bénéfice imposable et, si elle a néanmoins perçu la TVA, elle déroge également à son rôle de percepteur de cet impôt en détournant une partie de la TVA brute qu'elle aurait dû reverser à l'État⁴. Dans une autre configuration, une entreprise peut chercher à réduire son impôt sur les bénéfices en augmentant artificiellement certaines de ses charges, notamment ses frais généraux ; dans ce cas, elle peut simultanément accroître à tort sa TVA déductible. Il arrive également que la TVA soit calculée sur la base de taux erronés. Pour se soustraire à l'impôt, les entreprises ont recours à différents moyens, pouvant aller du maquillage de la réalité économique, par omission de recettes ou falsification de charges, à la manipulation d'écritures comptables telles que les amortissements ou les provisions, en passant par le transfert d'une partie de leurs bénéfices vers l'étranger. Outre le contournement des règles strictement comptables, les entreprises jouent aussi sur celles qui régissent le passage du résultat comptable au résultat fiscal, et plus généralement sur l'application de la législation fiscale, au point d'aboutir à des malversations. Il arrive enfin à des entreprises de se méprendre dans l'interprétation de la législation fiscale, sans avoir l'intention de frauder. Ces cas sont évidemment repérés par la Dgfi et traités comme de la fraude dans cette étude.

Au fil du temps, les procédés mis en œuvre pour frauder sur l'impôt ont évolué, face à une législation fiscale de plus en plus complexe et en liaison avec d'importantes transformations dans le fonctionnement de l'entreprise. En premier lieu, les modes de gestion se sont rationalisés, notamment depuis la mise en place de la TVA⁵, dont le mécanisme suppose que soient enregistrés, non seulement les ventes mais aussi les achats. Ils se sont aussi modernisés, avec le développement de l'informatique. La plupart des entreprises tiennent de nos jours une comptabilité rigoureuse et font appel, le plus souvent, aux services d'un comptable, ce qui n'était pas toujours le cas autrefois. Ces raisons contribuent probablement à expliquer le recul du travail sans facture et celui de la fraude par omission de recettes, son corollaire. D'autres formes d'évasion fiscale ont en revanche pris le relais. La fraude par

⁴ La TVA brute est la TVA que l'entreprise devrait reverser à l'État au titre du chiffre d'affaires écoulé sur le marché intérieur (les exportations sont exonérées de TVA). La TVA déductible est la TVA payée par l'entreprise à l'occasion de ses achats ; elle la déduit de sa TVA brute ce qui réduit d'autant le montant final de TVA qu'elle reverse effectivement au fisc. Au total, ce montant ne porte que sur la valeur ajoutée, comme l'indique le nom de l'impôt. En réalité la TVA est *in fine* payée par le consommateur, mais dans le processus de collecte de cet impôt l'entreprise s'interpose entre le consommateur et l'État, jouant le rôle du percepteur. Une entreprise qui fraude sur la TVA brute ou déductible s'approprie une partie de l'impôt qu'elle est censée « collecter ».

⁵ La TVA est un impôt relativement récent puisqu'elle a été mise en place en 1954. Elle ne concernait au départ que les grandes entreprises ; en 1966, elle a été étendue au commerce de détail.

contournement de la législation, sans aucune manipulation des flux comptables de l'entreprise, est désormais fréquemment pratiquée, notamment par les grandes entreprises. Parallèlement, l'économie a développé des ramifications bien au-delà des frontières françaises. Avec la mondialisation et le développement des sociétés multinationales qui en découle, de nouvelles possibilités d'échapper à l'impôt ont émergé. Les groupes s'implantent certes dans des zones où les coûts de production sont les plus avantageux mais ils recherchent aussi les pays dont la fiscalité est favorable aux entreprises, afin de pouvoir transférer une partie de leurs bénéfices vers les filiales les moins imposées ; cette pratique est aujourd'hui courante.

De manière schématique, à l'issue d'un contrôle fiscal, le vérificateur aura évalué l'impôt fraudé par l'entreprise en l'expliquant à l'aide d'un certain nombre de « chefs de rectification ». L'impôt non perçu donne lieu à des « droits rappelés » - dans le cas de la TVA, on parle de « rappels » : c'est l'objet de la vérification. Pour évaluer les droits rappelés, il est nécessaire de connaître les « bases » de calcul de ces droits ; on parle aussi « d'assiettes ». C'est pourquoi les rectifications sont déclinées en « droits rappelés » (il s'agit alors de l'impôt) et en « bases » (il s'agit alors de flux comptables). Les différents « chefs de rectification » sont décrits par une nomenclature très détaillée, comptant environ sept cents postes : le thésaurus des contrôles fiscaux. Les rectifications y sont répertoriées par référence à la législation fiscale, définie par le Code général des impôts (CGI).

Les « droits rappelés » et les « rectifications en base » sont actuellement enregistrés dans une application gérée par la Dgfp (dénommée ALPAGE). Le système d'enregistrement des données de contrôle de la Dgfp a été rénové à la fin des années quatre-vingt-dix pour devenir complètement opérationnel à partir de 2003. Le thésaurus des contrôles fiscaux a remplacé l'ancienne nomenclature de repérage des rectifications, beaucoup plus simple et plus facile à interpréter pour des profanes.

Par ailleurs, une partie des informations qui étaient jusqu'alors mobilisées par l'INSEE, principalement certaines bases d'imposition, n'ont pas pu lui être communiquées (*cf.* § III-2). Pour procéder aux évaluations présentées dans ce document, il a donc été nécessaire de remettre complètement à plat la méthode mise en œuvre à l'occasion de la base 2000 des comptes nationaux.

II - Les objectifs de l'étude, les grandes lignes de la méthode et les données mobilisées

II.1 Quels agrégats comptables corriger et comment ?

Le but de l'étude était de parvenir à mobiliser les résultats des contrôles fiscaux pour redresser les principaux agrégats et soldes des comptes nationaux, au titre de l'un des volets de l'activité fraudée, et non le moindre : l'activité non déclarée à l'administration fiscale par des entreprises juridiquement identifiées.

Dans la comptabilité des entreprises, la dissimulation d'activité se répercute à différents niveaux. Elle se traduit principalement par une sous-estimation du chiffre d'affaires et par une sous-estimation du résultat imposable. Les différentes formes de fraude s'imbriquent au sein des comptes, de sorte que ces deux sous-estimations ne sont pas en général de même ampleur. Ainsi, une entreprise qui fraude sur le chiffre d'affaires pourra parfois minorer quelque peu ses consommations intermédiaires, sous réserve de pouvoir acheter sans facture, ainsi que ses charges salariales, en ayant recours au travail au noir ; ses comptes seront alors plus vraisemblables et l'impact sur le résultat de la dissimulation d'activité sera moindre. À l'inverse, et cette situation est fréquente, beaucoup d'entreprises fraudent uniquement sur leurs frais généraux, en les majorant cette fois, toujours dans le but de réduire leur résultat imposable.

Dans les comptes nationaux, le redressement de la fraude impacte également plusieurs agrégats et soldes : principalement la production, les consommations intermédiaires, la valeur ajoutée. En toute logique, l'excédent brut d'exploitation (EBE) devrait aussi être affecté, la fraude sur l'activité allant souvent de pair avec une dissimulation partielle des heures travaillées. Hormis les consommations intermédiaires, ces agrégats ou soldes ont été corrigés par extrapolation directe de rectifications fiscales. Le bouclage comptable complet fait en outre intervenir des traitements spécifiques à la comptabilité nationale. Le principe en est présenté en fin de document (*cf. § VI*), mais quelques explications préliminaires sont indispensables à ce stade pour comprendre la manière dont ont été interprétées les données de contrôle (*tableau 1*).

En premier lieu, l'objectif étant de corriger des flux comptables, et non pas des impôts, seules des rectifications en base ont été utilisées. Par ailleurs, une part non négligeable des rectifications n'est pas liée à une manipulation de leurs comptes par les entreprises, mais à une application erronée de la législation fiscale. Concrètement, ceci se produit, par exemple, lorsqu'une entreprise s'octroie une exonération à laquelle elle n'a pas droit : elle réduit alors son impôt bien que son bénéfice imposable ne soit pas falsifié. Les chefs de rectification de ce type sont nombreux et ils sont, par nature, excessivement techniques et particulièrement hermétiques pour des profanes puisqu'ils ne s'appuient que sur l'interprétation de la législation. Ils ne seront pas davantage explicités dans ce document. On retiendra seulement que ces rectifications de nature purement fiscale n'ont pas à intervenir dans le calcul de l'activité fraudée, puisqu'elles ne découlent d'aucune malversation de nature comptable.

Les principales rectifications en base mobilisées sont très globales : elles concernent le chiffre d'affaires et le résultat. En outre, les rectifications touchant aux amortissements et aux provisions ont dû être défalquées de la rectification du résultat de l'exercice, au même titre que les rectifications de nature purement fiscale mentionnées ci-dessus, mais pour des raisons différentes, liées au fait que l'articulation des comptes de secteur de la comptabilité nationale n'est pas totalement calquée sur la comptabilité d'entreprise (*tableau 1 et § III-2*). On aboutit ainsi à une rectification de l'excédent brut d'exploitation (EBE), cohérente conceptuellement avec la correction de l'agrégat correspondant de la comptabilité nationale, **si ce n'est que les salaires ne faisant l'objet d'aucune rectification lors des contrôles fiscaux, la rectification du résultat de l'entreprise, ainsi aménagée, s'applique en réalité à la valeur ajoutée.**

**Tableau 1 : Organisation schématique de la comptabilité d'entreprise
et des comptes de secteur de la comptabilité nationale**

Comptes nationaux		Comptabilité d'entreprise	
Comptes de production et d'exploitation		Compte de résultat	
	Production + Δ (P)		Chiffre d'affaires+Rect.(CA)
Conso. intermédiaires + Δ (CI)		Conso. intermédiaires	
	Valeur ajoutée + Δ (VA)		Valeur ajoutée + Rec. (EBE)
Salaires + Δ (S)		Frais de personnel	
	Excédent brut d'exploitation + Δ (EBE)		Excédent brut d'exploitation + Rec. (EBE)
Comptes de répartition du revenu (impôts, prestations, intérêts et dividendes, etc.)		Amortissements + Rect. (amortissements)	
		Provisions + Rect. (provisions)	
	Revenu disponible brut (RDB) + Δ (EBE)	Charges et produits financiers	
			Résultat courant avant impôt (RCAI)
Consommation de capital fixe (CCF)		Charges et produits exceptionnels	
		Impôts sur les bénéfices	
	Revenu disponible net		Résultat de l'exercice + Rect. (résultat)

Lecture : Rect. : lire « rectification en base ».

Le tableau ci-dessus met en regard, de manière extrêmement schématique, l'organisation des comptes de secteur de la comptabilité nationale et celle de la comptabilité d'entreprise. Le but est de mettre en évidence les divergences qui auront un impact sur le traitement des données des contrôles fiscaux, et uniquement sur ces dernières. Y figurent également, parmi les rectifications opérées lors des contrôles, celles qui interviennent dans l'évaluation des redressements de la comptabilité nationale présentés dans ce document : elles sont mentionnées dans les cases hachurées.

II.2 Les grandes lignes de la méthode d'estimation

Le principe du calcul des corrections pour fraude de la comptabilité nationale est d'extrapoler, à l'ensemble des entreprises susceptibles d'être concernées, la fraude décelée par la Dgfi sur l'échantillon des entreprises contrôlées. Cette extrapolation passe par l'intermédiaire de deux jeux de taux de fraude rapportant, respectivement, la fraude sur le chiffre d'affaires à ce dernier, et la fraude sur le résultat au chiffre d'affaires également. Ces taux sont calculés sur l'échantillon des entreprises contrôlées, selon une stratification pertinente qui vise à regrouper, au sein d'une classe, des entreprises ayant la même propension à frauder. Pour extrapoler, il suffit d'appliquer les taux de fraude observés sur chaque strate de l'échantillon aux chiffres d'affaires de toutes les entreprises appartenant à la strate correspondante de la population ; cela signifie que ces dernières sont supposées avoir le même comportement de fraude que les entreprises similaires de l'échantillon.

Pour obtenir les évaluations macroéconomiques présentées dans ce document, la fraude a été extrapolée aux entreprises répertoriées dans le fichier de SUSE 2006 (Système Unifié de Statistiques d'entreprises). On en a déduit de nouveaux taux de fraude par secteur d'activité. Ces taux extrapolés seront ensuite appliqués aux chiffres d'affaires des comptes nationaux retenus pour chaque année de la période de base (2006 à 2011), chiffres d'affaires dérivés de ceux de SUSE après différents traitements, spécifiques à la comptabilité nationale. On suppose donc implicitement que la propension à frauder des entreprises ne change pas en cours de base. C'est une hypothèse légitime compte tenu de la relative inertie du phénomène observé. Elle se justifie en outre par la nature et la fragilité des données disponibles qui rendent préférable l'utilisation de taux de fraude moyens plutôt que celle de taux annuels.

Cette méthode s'applique à tous les secteurs d'activité, hormis aux secteurs agricoles, aux banques et aux sociétés d'assurance. En effet, la fraude, telle qu'elle ressort de cette étude, vise spécifiquement à redresser la liasse fiscale ; or, le compte de ces secteurs, en comptabilité nationale, n'est pas établi à partir de cette dernière. Dans le cas de l'agriculture, il n'y a pas lieu de corriger la production pour fraude puisque celle-ci est déterminée à partir des quantités produites, lesquelles englobent la production fraudée, par définition. S'agissant du compte de secteur des institutions financières et des sociétés d'assurance, il

est établi par la Banque de France, à partir de ses propres sources ; on peut donc admettre qu'il n'a pas non plus à être corrigé au titre de la fraude.

II.3 Les données mobilisées et leur organisation

Pour effectuer les évaluations des bases précédentes, les taux de fraude sur échantillon avaient été calculés en prenant, pour dénominateur, l'assiette de la TVA brute, laquelle constituait une approximation correcte du chiffre d'affaires de l'entreprise sur le marché intérieur. Ainsi, ces taux avaient pu être obtenus à partir des seules données de la DGI. Les chiffres d'affaires de l'INSEE n'intervenaient qu'au stade de l'extrapolation et du calage final. Le fait que les bases de la TVA n'aient pas pu être transmises à l'INSEE pour effectuer la présente étude est l'une des conséquences les plus gênantes de la nouvelle organisation des données de contrôle de la Dgfi. Il devenait impossible de calculer les taux de fraude dans l'univers Dgfi, ce qui interdisait de reproduire à l'identique la méthode de la base 2000. Les autres difficultés proviennent de changements dans les définitions des rectifications.

Pour parvenir à calculer les taux de fraude de la présente étude, la seule solution était de récupérer dans SUSE les chiffres d'affaires des entreprises contrôlées, ce qui nécessitait de mettre en regard les variables de l'échantillon de la Dgfi avec celles de SUSE. En rapprochant ainsi la population avec un échantillon tiré dans cette dernière, on allait en outre pouvoir procéder à des vérifications, calculer aisément des taux de sondage, mettre en œuvre des méthodes économétriques et des analyses statistiques faisant intervenir des variables de SUSE, notamment des ratios comptables. Le secteur d'activité a pu également être codifié en fonction des APE⁶ de SUSE, *a priori* plus fiables que ceux de la Dgfi et, d'une manière générale, plus cohérents avec les données sur les entreprises de l'INSEE. Le type d'imposition et la catégorie juridique sont également ceux de SUSE, alors que les évaluations antérieures utilisaient les variables correspondantes de la DGI.

II.3.1 L'échantillon des contrôles fiscaux de la Dgfi

Les contrôles fiscaux sont organisés en « campagnes » annuelles au cours desquelles 45 000 entreprises environ sont vérifiées. Une entreprise contrôlée au cours d'une année donnée ne doit pas l'être à nouveau au cours des trois années suivantes, sauf si la vérification porte sur un impôt qui n'a pas encore été contrôlé. On peut ainsi considérer que les échantillons de trois années successives sont à peu près disjoints.

Ces échantillons annuels ne sont pas, il est vrai, aléatoires (*cf.* § IV-1). Ils résultent d'une procédure de sélection complexe et confidentielle, fondée sur l'examen d'un dossier très complet. Toutefois, la Dgfi a le souci de balayer l'ensemble des secteurs d'activité et, en principe, toutes les grandes entreprises doivent être vérifiées en cinq ans⁷. En première approximation, on a donc considéré *qu'en regroupant les données de trois campagnes de contrôles successives, en l'occurrence 2004, 2005 et 2006*, on devait obtenir un échantillon d'entreprises ne comportant pas de redondances, de taille suffisante pour que la précision des estimations soit convenable et assurément plus proche d'un échantillon d'enquête que ne le sont les échantillons annuels. Les taux de fraude estimés à partir de la concaténation de trois échantillons annuels sont des taux moyens sur la période, ce qui n'est pas un inconvénient majeur puisqu'ils seront conservés pendant toute la durée de la base.

L'échantillon n'en demeure pas moins « ciblé » au regard de la mesure de la fraude. Le but de la Dgfi étant de recouvrer un maximum des impôts non perçus, elle recherche manifestement les entreprises qui risquent de frauder le plus, soit en raison de leur taille, soit

⁶ L'INSEE attribue à toutes les entreprises, lors de leur inscription au répertoire SIRENE, un code APE, caractérisant l'activité principale de l'entreprise par référence à la nomenclature d'activités française. L'APE est à la base du classement des entreprises par secteur d'activité.

⁷ Plus l'entreprise est importante, plus fréquents sont les contrôles.

parce qu'elles ont une forte tendance à le faire. Le point clé de la technique d'estimation des taux de fraude, et l'une des principales difficultés du travail, est de parvenir à modéliser ce « ciblage » aussi fidèlement que possible, sans quoi des biais d'extrapolation pourraient affecter la qualité des résultats (cf. § IV-2-a et § V).

II.3.2 Le fichier Suse-Dgfip

Le calcul de la fraude présenté dans ce document s'appuie sur deux gisements de données : celui des contrôles fiscaux fourni par la Dgfip, les fichiers de SUSE, restreints aux entreprises vérifiées, pour les années 2000 à 2006, et le fichier de SUSE 2006, pour l'ensemble des entreprises appartenant au champ. Ce champ couvre l'ensemble des entreprises françaises, hormis celles qui ne font pas l'objet d'un redressement pour fraude en comptabilité nationale (secteurs agricoles, banques et les sociétés d'assurance). Le fichier-échantillon est constitué d'un panel de 117 300 entreprises, non cylindré, car des entreprises disparaissent et se créent chaque année. Ce fichier compte, pour chaque entreprise, plus de 200 variables utiles. Le fichier de calage est celui de SUSE 2006. Après élimination des entreprises hors champ, il comporte plus de 2 600 000 enregistrements (*Annexe 1*).

III - Définition et contenu des taux de fraude

La manière d'interpréter les contrôles fiscaux pour aboutir aux corrections recherchées s'appuie très largement sur les indications communiquées par la Dgfp. L'objectif était d'adapter et de synthétiser les données issues des contrôles fiscaux, telles qu'elles sont définies et organisées dans le nouveau système d'enregistrement (*cf. § I*), pour les rendre aussi cohérentes que possible avec les informations transmises à l'INSEE en 2002, pour effectuer les estimations de la base 2000.

Les données de contrôle qui ont été mobilisées sont les suivantes :

- les omissions de recettes, en base ;
- les droits rappelés sur l'impôt sur les bénéfices, pour les entreprises assujetties à l'impôt sur les sociétés (IS) d'une part, à l'impôt sur le revenu (IR) d'autre part ;
- les rectifications en base du résultat des entreprises déficitaires assujetties à l'IS d'une part, à l'IR d'autre part ;
- les rectifications en base sur amortissements et provisions ;
- les transferts de bénéfices en base ;
- le « résultat établi par le service » ;
- les rectifications de nature purement fiscale en base et en « droits rappelés ».

Le but du travail étant d'estimer des flux comptables dissimulés (chiffres d'affaires et revenus), les estimations dériveront directement des bases d'imposition et non pas des droits rappelés (montants d'impôts à recouvrer). Le principal écueil auquel se heurte l'utilisation des informations énumérées ci-dessus est lié au fait qu'elles sont, de ce point de vue, hétérogènes : toutes les rectifications fournies ne sont pas en base, ce qui aurait été préférable.

Toutes ces rectifications sont relatives à l'ensemble de la période de contrôle.

Les taux de fraude observés ont été calculés par référence à la durée complète du contrôle, en rapportant les rectifications relatives à la période de contrôle, au chiffre d'affaires de l'entreprise correspondant à cette même période (source SUSE). Ce dernier a été obtenu en sommant les chiffres d'affaires, ou les fractions⁸ de chiffres d'affaires, des différents exercices vérifiés. D'autres variables utilisées dans les calculs ont été, de manière similaire, ramenées à la période de contrôle.

Les taux de fraude observés sont donc des taux moyens sur la période de contrôle, évalués sur une plage de temps assez large puisque la majeure partie de ces contrôles s'étale sur au moins trois exercices comptables, compris entre 2000 et 2006 (*Annexe 1*). Pour effectuer les extrapolations, on a admis que ces taux moyens pouvaient s'appliquer aux chiffres d'affaires de l'année 2006.

III.1 Évaluation du chiffre d'affaires dissimulé

La fraude sur le chiffre d'affaires est assez simple à définir ; elle repose principalement sur les rectifications en base pour omissions de recettes, utilisées telles qu'elles ont été

⁸ Calculées au prorata de la durée infra-annuelle du contrôle, si celui-ci ne couvrait qu'une partie de l'année (en début et en fin de contrôle).

communiquées⁹ et, pour un certain nombre d'entreprises, sur une partie du chiffre d'affaires dérivé du « résultat établi par le service ».

Le « résultat établi par le service » résulte de la reconstruction des comptes de l'entreprise par le vérificateur. Celui-ci se trouve confronté à cette situation extrême lorsque la comptabilité n'est pas crédible ou si l'entreprise ne dépose pas de comptes. Ces comptes « reconstruits » sont établis à partir de tous les éléments disponibles susceptibles d'éclairer le vérificateur, tels que des données sur les stocks de marchandises ou sur des quantités physiques de produits intermédiaires, par exemple. En pareil cas, les omissions de recettes, au demeurant fréquentes dans de telles situations, ne sont pas incluses dans le poste « omissions de recettes » mentionné ci-dessus. Il a donc été nécessaire d'en évaluer approximativement le montant à partir du « résultat établi par le service », ce dernier constituant la seule information mobilisable à cet effet. Les omissions de recettes correspondant spécifiquement à ces cas ont été calculées en considérant que le « résultat établi par le service », ou plus exactement une fraction de ce résultat¹⁰, est égal à 60 % du chiffre d'affaires, ce qui revient à admettre que, pour les entreprises concernées, les charges représentent 40 % du chiffre d'affaires. Le chiffre d'affaires dissimulé ainsi évalué a été ajouté, le cas échéant, aux omissions de recettes initiales.

La formule de calcul de la fraude sur le chiffre d'affaires est alors la suivante :

Fraude sur chiffre d'affaires = omissions de recettes + (résultat établi par le service des cas L55 et L73) / 0,6
--

III.2 Évaluation de la fraude sur le résultat

L'évaluation de la fraude sur le résultat à introduire dans le compte des entreprises de la comptabilité nationale est plus complexe puisqu'il ne faut retenir à cet effet qu'une partie seulement des rectifications contribuant à la rectification en base globale du résultat des entreprises (cf. § II-1). *Seules doivent intervenir les rectifications portant sur les flux comptables réels, intervenant en amont de l'EBE.*

Le tableau 1 met en évidence une différence d'organisation essentielle entre la comptabilité d'entreprise et le compte du secteur des entreprises de la comptabilité nationale : le positionnement des amortissements et des provisions. Ces flux interviennent systématiquement dans le calcul du résultat imposable des entreprises alors qu'ils ne sont pas pris en compte dans le calcul du PIB. *Par conséquent, les rectifications en base portant sur les amortissements et les provisions, qui sont parmi les plus fréquentes, n'ont pas lieu d'être retenues pour redresser l'EBE ou le revenu disponible brut (RDB) des comptes nationaux.* Plus précisément, en comptabilité d'entreprise, les amortissements sont des écritures comptables : ils représentent, pour l'entreprise, l'étalement du coût d'un investissement, tel que l'achat d'une machine, sur la durée d'utilisation du bien. Des provisions sont également inscrites au compte de résultat pour permettre à l'entreprise de prévoir le financement d'une dépense probable, supposée lui échoir dans l'avenir. Ces écritures jouent un rôle essentiel dans la gestion comptable de l'entreprise puisqu'elles ont un impact, entre autres, sur le calcul de l'impôt sur les bénéfices. Derrière la notion d'amortissement comptable, on retrouve bien l'idée d'amortissement du capital, si ce n'est que l'amortissement fiscal est calculé en fonction de règles qui lui sont propres. Dans les comptes nationaux, l'amortissement économique du capital (la consommation de capital fixe

⁹ Le poste « omissions de recettes » recouvre quatre rubriques du thésaurus : les recettes sur chiffre d'affaires non comptabilisées et non déclarées (poste 00383110), les recettes sur chiffre d'affaires comptabilisées et non déclarées (poste 0038020), le poste « recettes sur chiffre d'affaires - exercice rattachement erroné » (poste 0038030), le poste « BNC - redressements pour recettes » (poste 0091030).

¹⁰ On dispose d'une ventilation de ces « résultats établis par le service » selon quatre situations, dont deux seulement ont été retenues. Elles correspondent aux articles L55 et L73 du Code des impôts : dans le premier cas, la comptabilité est rejetée ou partiellement rejetée et il s'agit d'un supplément de résultat ; dans le second cas, les comptes ne sont pas déposés et il s'agit du résultat global.

ou CCF) intervient certes, mais pas dans le calcul du revenu disponible brut (RDB). Par ailleurs, les provisions qui sont, pour les entreprises, des outils d'anticipation comptable et financière, n'ont pas d'équivalent en comptabilité nationale. Doivent également être éliminées, toutes les rectifications qui résultent d'une mauvaise application de la législation fiscale, sans qu'un maquillage des flux réels en soit à l'origine. La Dgfi a isolé les chefs de rectification du thésaurus correspondant à ce type de situation.

III.2.1 Le passage des rectifications en « droits rappelés » aux rectifications « en base »

Parmi les rectifications qui sont mobilisées pour évaluer la fraude sur le résultat, certaines sont « en base » et d'autres en « droits rappelés », notamment la rectification globale de l'impôt sur les bénéfices, sur laquelle s'appuie le calcul de la rectification en base du résultat. Cet inconvénient est inévitable car il est lié à la nouvelle organisation du système d'information de la Dgfi dans lequel les bases antérieures à la vérification ne sont pas toutes enregistrées ; c'est le cas notamment des résultats d'exercices. Pour parvenir *in fine* aux rectifications en base recherchées, la seule méthode envisageable consistait à transformer « en bases » les rectifications « en droits », et cela ne pouvait se faire qu'en les divisant par des taux d'imposition *ad hoc*. Cette opération a nécessité des hypothèses sur les taux d'imposition, auxquelles les estimations se révèlent très sensibles, ce qui les fragilise quelque peu.

Le cas des entreprises assujetties à l'IS

Dans le cas des entreprises assujetties à l'IS, deux jeux d'hypothèses ont été testés. Dans un premier temps, on a admis que l'on avait une approximation correcte de la base de calcul de l'impôt sur les bénéfices en divisant le montant des droits rappelés sur l'impôt sur les bénéfices par le taux d'imposition de 33,33 %, en général appliqué, pour cette catégorie de contribuables. Cette méthode s'est par la suite révélée décevante, en raison de problèmes de cohérence entre la rectification en base globale du résultat ainsi obtenu et les rectifications partielles, en base, qui en sont défalquées. Ces incohérences laissaient craindre une sous-estimation de la rectification en base globale, le taux d'imposition de 33,33 % devant plutôt être perçu comme un majorant d'un taux d'imposition apparent plus faible.

Dans un deuxième temps, on a procédé à une estimation de ces taux d'imposition sur IS apparents. Pour chaque entreprise du panel, et pour chaque exercice (2000 à 2006), on dispose en effet, dans SUSE, du montant de l'impôt sur les bénéfices et du résultat après impôt ; l'information précisant le type d'imposition (IS ou IR) n'est pas dans SUSE, mais figure parmi les données de la Dgfi. Les taux d'imposition ont été évalués en moyenne, sur deux ou trois exercices. Ils se révèlent, de fait, sensiblement inférieurs au taux officiel de 33,33 % (*Annexe 1*). Ce dernier est effectivement appliqué dans la majeure partie des cas, mais les entreprises bénéficient de différentes exonérations, voire, dans quelques cas isolés, d'un taux plus réduit, de sorte que les taux finals apparents sont plus faibles que le taux officiel. Ce calcul présente toutefois des éléments de fragilité, liés aux données d'abord : les grandeurs mobilisées comportent de nombreuses valeurs manquantes, ce qui oblige à les redresser, et il peut exister des erreurs de mesure. Par ailleurs, si ces taux apparents s'écartent du taux officiel, c'est en raison de différentes mesures fiscales (déficits reportables, exonérations diverses, etc.), applicables et appliquées avant le passage du vérificateur, mais parfois à tort. Or, il n'est pas exclu que certains de ces abattements aient été remis en question lors de la vérification, le taux d'imposition apparent se trouvant alors rehaussé.

Cette seconde méthode de passage en base n'a rien d'arbitraire : elle se fonde sur des informations objectives et elle a le mérite de réduire le nombre et l'ampleur des incohérences entre rectifications globales et partielles. On peut donc penser, qu'en dépit de ses faiblesses, elle est préférable à la méthode initiale, consistant à retenir un taux d'imposition unique de 33,33 %.

Le cas des entreprises soumises à l'IR

Dans le cas des entreprises soumises à l'IR, pour passer des droits rappelés à la base, il fallait cette fois déterminer des taux marginaux d'imposition et non pas des taux moyens. La méthode précédente n'a pas pu être reprise pour deux raisons. D'une part, le nombre des informations manquantes sur l'impôt est élevé pour ces contribuables, ce qui rend le calcul très imprécis ; d'autre part, cette méthode aurait débouché sur des taux moyens d'imposition, et non pas sur des taux marginaux.

On a donc tenté d'estimer ces taux marginaux d'imposition à partir du profit brut courant avant impôt (PBCAI¹¹) de chaque entreprise, disponible dans SUSE pour tous les exercices contrôlés. En le ramenant à la période de contrôle, puis à l'année, et en le divisant par 2,5 pour tenir compte d'un nombre de parts, on obtient une approximation du revenu imposable annuel moyen de l'entreprise au cours de la période de contrôle, grossière certes mais néanmoins acceptable. En comparant ce revenu aux limites des tranches d'imposition de l'impôt sur le revenu (en application en 2005), on en déduit des taux marginaux d'imposition individuels.

Cette méthode comporte aussi des éléments de fragilité : le PBCAI n'est pas exactement le revenu imposable, le choix du nombre de parts est arbitraire, les données sont entachées d'erreurs de mesure. Toutefois, le taux d'imposition étant défini *in fine* en comparant un revenu imposable, certes approximatif, à des valeurs limites (celles des tranches d'imposition), les conséquences des imprécisions qui affectent ce dernier se trouvent atténuées.

III.2.2 La question des transferts de bénéfices

Ces opérations concernent les filiales des groupes. Par le biais de leurs comptes consolidés, les filiales procèdent à des « transferts de bénéfices » entre elles, le but visé étant de réduire globalement l'impôt payé par le groupe. La mise en œuvre des transferts de bénéfices passe par les « prix de transfert » des échanges intra-groupe. Les prix de ces échanges n'ont rien de commun avec des prix de marché. Ils sont définis par le groupe, lequel s'en sert pour optimiser ses prélèvements fiscaux. Il aura, par exemple, intérêt à rehausser les prix des ventes (respectivement, à minorer les prix des achats) d'une filiale irlandaise à destination d'une filiale française (respectivement, de la filiale irlandaise à la filiale française) de manière à gonfler le résultat de la filiale irlandaise au détriment de celui de la filiale française. Globalement, la ponction fiscale du groupe se trouve ainsi réduite puisque la filiale irlandaise est imposée moins lourdement que la filiale française.

La majeure partie des rectifications regroupées sous cette rubrique correspondrait à des situations de ce type, même si quelques manipulations fiscales s'y ajoutent. Ces pratiques revenant incontestablement à transférer vers l'étranger le produit d'une activité réalisée sur le territoire français, les rectifications pour transfert de bénéfices devraient en théorie être retenues. Dans les comptes nationaux, pour des raisons d'équilibrage, il faudrait alors corriger « en miroir » les échanges intra-groupe ; or, les échanges font partie des données qui ne sont pas, ou très marginalement, modifiées à l'initiative des comptes nationaux, principalement pour préserver au mieux la cohérence du commerce extérieur des comptes avec les résultats des sources servant à l'établir, en l'occurrence celles de la Direction générale des Douanes et de la Balance des paiements. Au final, on a donc renoncé à corriger le résultat des entreprises de ces transferts de bénéfices, lesquels n'ont, au demeurant, qu'un impact relativement limité sur la correction globale de l'EBE, de l'ordre de 7 % (soit environ 3 milliards d'euros).

Un dernier choix concerne le traitement des rectifications visant les dirigeants et les associés. Il est en effet fréquent qu'une entreprise rectifiée sur IS le soit aussi sur IR. La

¹¹ Le PBCAI de SUSE correspond à l'EBE dont on retranche les produits et charges financiers. Il renseigne sur le résultat de l'entreprise hors événements exceptionnels et incidence fiscale.

rectification sur IR concerne alors les dirigeants et les associés ; elle n'a pas été prise en compte.

Les formules de calcul de la fraude sur le résultat sont finalement les suivantes :

Rectification sur le résultat des entreprises assujetties à l'IS =
 (droits rappelés - rectifications fiscales en droits) / tximp_IS + déficit - rectifications sur amortissements
 et provisions en base - rectifications fiscales en base - transferts de bénéfices en base.

Rectification sur le résultat des entreprises assujetties à l'IR =
 (droits rappelés - rectifications fiscales en droits) / tximp_IR + déficit - rectifications sur amortissements
 et provisions en base - rectifications fiscales en base - transferts de bénéfices en base.

Les « droits rappelés » sont les droits rappelés totaux appliqués à l'impôt sur les bénéfices. Le terme « déficit » correspond à la rectification en base appliquée au résultat des entreprises déficitaires. Ces entreprises ne paient pas l'impôt sur les bénéfices ; néanmoins, le vérificateur peut fort bien remettre en cause l'ampleur de leur déficit ; il en résulte une rectification « en base » de ce déficit. Pour tximp_IS, lire « taux d'imposition des contribuables soumis à l'IS » ; pour tximp_IR, lire « taux d'imposition des contribuables soumis à l'IR ».

La principale incertitude de ce calcul est clairement liée à l'utilisation conjointe de rectifications en base et en droits rappelés. La transformation en base des droits rappelés totaux est particulièrement fragile, car le calcul de la part fraudée de l'impôt sur les bénéfices ne se résume pas à la simple application d'un taux d'imposition unique, aussi pertinent soit son mode de calcul, à une rectification en base unique. C'est pourtant ce que traduit la méthode du passage des droits rappelés aux bases correspondantes, formalisée ci-dessus.

IV - Les aspects statistiques de la méthode d'estimation

Projeter à l'ensemble des entreprises l'activité fraudée mesurée sur l'échantillon des contrôles fiscaux est un travail qui s'apparente à l'extrapolation d'un fichier d'enquête. C'est un exercice assez classique pour un statisticien, si ce n'est que l'échantillon lui est, dans le cas présent, imposé. Il faut alors imaginer un ou plusieurs plans de sondage virtuels qui permettraient de tirer l'échantillon de la Dgfp, tel qu'on l'observe ; c'est là que réside l'originalité du travail. Lorsqu'un statisticien élabore un plan de sondage, il cherche à affecter à chaque entreprise de la population une certaine probabilité d'appartenir à l'échantillon, dite « probabilité d'inclusion », qu'il définit *a priori* en fonction des objectifs de son enquête. Dans la démarche *a posteriori*, qui est celle de la présente étude, il doit essayer d'estimer des probabilités de tirage qu'il ne connaît pas, et faire au mieux pour éviter les biais d'extrapolation (cf. § IV-2).

IV.1 Le processus de sélection de la Dgfp est-il modélisable ?

Les entreprises vérifiées sont sélectionnées par la Dgfp, à plusieurs niveaux, au vu de différents éléments sur le fonctionnement de l'entreprise, son environnement fiscal et financier et sur ses dirigeants. Sont notamment examinées des informations comptables sur plusieurs années et des données patrimoniales. Certaines règles doivent être respectées, relatives à la fréquence des contrôles, par exemple. La sélection se faisant localement, il est clair que la connaissance du terrain et des appréciations portées à dire d'expert jouent un rôle important dans les choix opérés. La préoccupation dominante de l'administration fiscale est de recouvrer un maximum des sommes dues. Cette optique oriente à l'évidence le processus de sélection, lequel vise en priorité les entreprises susceptibles de devoir des masses d'impôts importantes ; parmi celles qui pourraient avoir la même propension à frauder, les grandes entreprises sont donc plus recherchées que les plus petites. Cette pratique est par ailleurs conforme à la manière de sonder une population dont les unités sont de tailles très différentes : il faut éviter de « rater » les grandes. La Dgfp semble également chercher à vérifier presque tous les secteurs d'activité sans trop tenir compte des propensions sectorielles à frauder ; ce n'est pas là que se situe le « ciblage », semble-t-il.

On peut douter de parvenir à modéliser un processus de sélection aussi complexe et, qui plus est, relativement opaque. En réalité, la position du statisticien qui s'attèle à cet exercice n'est pas du tout celle des services fiscaux. Il se place *a posteriori* : ainsi, il connaît les entreprises contrôlées et, comme elles sont identifiées, il peut les comparer à celles qui ne le sont pas. Le fait de disposer de la réalisation du tirage est en soi une information, à la fois très simple et très riche, car elle résume l'ensemble du processus de sélection. Par une approche inductive, fondée sur l'observation des entreprises sélectionnées, il est possible de repérer, parmi les éléments statistiques dont on dispose, ceux qui expliquent le mieux les probabilités de tirage, en se limitant aux plus significatifs d'entre eux. De prime abord, une telle démarche peut être perçue comme réductrice, notamment parce qu'elle ne prend pas en compte les critères de sélection relevant typiquement de l'appréciation humaine. Elle l'est cependant moins qu'il n'y paraît. En effet, le statisticien bénéficie implicitement du travail de sélection fait *a priori* par les services fiscaux. Ses raisonnements sont donc différents et d'une certaine manière simplifiés ; il peut se dispenser de procéder à des recoupements entre des critères de sélection qui ne font que se conforter, recoupements qui sont en revanche indispensables dans une démarche *a priori*. Dit autrement, cela signifie qu'il existe des redondances dans le matériel mobilisé par les services fiscaux que la démarche statistique élimine d'office¹². À partir des variables statistiques dont on dispose, en nombre obligatoirement limité, on peut ainsi parvenir à simuler le processus de sélection de manière relativement fidèle, tout en le simplifiant énormément.

¹² Cette remarque a une traduction technique en économétrie. Dans une régression, lorsque deux variables explicatives délivrent le même message, on dit qu'elles sont colinéaires et il faut n'en conserver qu'une ; certaines procédures de régression, notamment celle qui a été utilisée dans le cadre de la présente étude, éliminent d'office les variables explicatives redondantes (Annexe 3).

IV.2 La stratification : une représentation du plan de sondage

La manière la plus simple de modéliser le plan de sondage de la Dgfp consiste à stratifier la population en fonction d'un certain nombre de variables statistiques dont dépend le comportement de fraude des entreprises. En premier lieu, on admet que la propension à frauder d'une entreprise, représentée par son (ou ses) taux de fraude, est la même sur l'ensemble de la strate de l'échantillon à laquelle appartient cette entreprise et l'on suppose en outre qu'elle est identique pour celles de la strate correspondante de la population. Sous cette hypothèse, les entreprises d'une strate donnée peuvent être représentées dans l'échantillon par n'importe laquelle d'entre elles. Les probabilités d'inclusion sont alors uniformes sur chaque strate, et égales au rapport entre les effectifs des strates correspondantes de l'échantillon et de la population. Les grandeurs à estimer sont ensuite extrapolées à l'aide des inverses de ces taux de sondage : plus le taux de sondage d'une strate est fort, plus nombreuses seront, en proportion, les entreprises de cette strate représentées dans l'échantillon et moins chacune d'elles devra par conséquent peser dans l'extrapolation (cf. § IV-3-2).

L'important est de bien spécifier la stratification. Les taux de fraude doivent être homogènes au sein d'une strate, et bien différenciés entre deux strates. En outre, les taux de sondage issus de la stratification doivent refléter aussi fidèlement que possible le mode de sélection de la Dgfp ; on notera qu'il s'agit bien d'une sélection et non pas d'un tirage aléatoire.

Si l'on suppose que les entreprises les plus visées ont des taux de fraude élevés - ce qui n'est pas toujours le cas, car les grandes entreprises sont fréquemment contrôlées alors qu'elle fraudent moins que les autres, en proportion de leur chiffre d'affaires - et si la probabilité d'inclusion que leur affecte le plan de sondage *a posteriori* de la stratification est plus faible que le taux de sélection réel, alors ces entreprises seront surpondérées lors de l'extrapolation.

Le mécanisme de cette surestimation est illustré schématiquement par l'exemple du tableau 2. La « stratification inadaptée » ne sépare pas les entreprises dont la propension à frauder est la plus forte (30 %) des autres, alors que les premières sont précisément surreprésentées dans l'échantillon. Le taux de fraude observé global (sur échantillon) est alors trop élevé (18,9 %) puisque les taux de fraude partiels les plus forts sont surpondérés. Le taux final extrapolé est de 18,9 %, contre 11,8 % avec une stratification plus adaptée, de sorte que l'estimation de la fraude globale est trop haute : 1 039 contre 650 pour l'estimation plus juste.

Tableau 2 : Effet de la finesse de la stratification sur l'estimation macroéconomique de la fraude

Taux de fraude	Le plan de sélection			Taux de fraude	Une stratification a posteriori inadaptée		
	Chiffres d'affaires		Fraude		Chiffres d'affaires		Fraude
	échantillon	population			échantillon	population	
30 %	1 000	1 500	450	18,9 %	1 800	5 500	1 039
5 %	800	4 000	200				
Total	1 800	5 500	650	Total	1 800	5 500	1 039
Taux de fraude après extrapolation			11,8 %	Taux de fraude après extrapolation			18,9 %

Lecture : Après extrapolation, le taux de fraude agrégé du « plan de sélection » est 11,8 % (= 650/5 500). Il s'exprime comme la moyenne des taux de fraude de chaque strate (30 % et 5 %), pondérée *par les poids en chiffre d'affaires des strates de la population* : $11,8 = (0,27 \times 30) + (0,73 \times 5)$. La stratification « inadaptée » regroupe, au sein d'une seule strate, les entreprises ayant un taux de fraude de 30 % et celles ayant un taux de fraude de 5 %. Le taux de fraude agrégé sur échantillon s'obtient en faisant la moyenne des taux élémentaires, avec les pondérations en chiffre d'affaires *des strates de l'échantillon, cette fois* : $18,9 = (0,55 \times 30) + (0,45 \times 5)$.

IV.2.1 Une première stratification, croisant le secteur d'activité, la taille et la catégorie juridique...

Depuis que les résultats des contrôles fiscaux sont exploités par l'INSEE, les extrapolations ont toujours été effectuées sur la base d'une post-stratification faisant intervenir le secteur d'activité (défini pour cette étude selon la NES¹³ en 118 postes, niveau G), la taille de l'entreprise, sa catégorie juridique (société non financière ou entreprise individuelle) et le type d'imposition. Dans un premier temps, cette méthode a été reprise dans le cadre de la présente étude, ne serait-ce que pour faciliter la comparaison des résultats avec ceux de la base précédente (*Annexe 3*)¹⁴. Bien qu'elle soit relativement robuste, elle fait l'objet de critiques récurrentes. On lui reproche de ne pas prendre en compte le « ciblage » des contrôles de manière suffisamment fidèle. Si le plan de sondage ainsi modélisé se révélait trop éloigné du schéma de tirage réel, il en résulterait des biais dans les estimations (*cf. supra*) ; or, ceci risque de se produire dans un certain nombre de cas. De fait, si la taille de l'entreprise est un critère de sélection avéré et le secteur d'activité une variable incontournable, le « ciblage » des services fiscaux s'appuie à l'évidence sur quantité d'autres informations.

IV.2.2 ...mais des variables comptables seraient pertinentes

La sélection de la Dgfi se fonde, entre autres, sur une batterie de ratios d'analyse financière, permettant de pointer des incohérences comptables. Il se trouve que certains de ces ratios devenaient accessibles dans le cadre de cette étude, du fait du rapprochement de l'échantillon de la Dgfi avec SUSE. On disposait notamment des plus simples d'entre eux : le taux de valeur ajoutée et le taux de marge¹⁵. Par principe, les entreprises fraudeuses cherchent à diminuer leurs impôts, et cela passe fréquemment par une réduction de leur bénéfice imposable. Que la fraude prenne la forme d'une dissimulation d'activité, d'une augmentation fallacieuse des charges ou de manipulations des amortissements et des provisions, elle a fréquemment pour conséquence une diminution du résultat imposable des entreprises. Même si ces dernières peuvent aussi frauder sans maquiller leurs comptes, nombre d'entreprises fraudeuses devraient avoir, selon toute vraisemblance, un taux de valeur ajoutée ou un taux de marge plutôt inférieur à la moyenne. Toutefois, a contrario, pour les entreprises dont les taux de marge sont structurellement élevés, il est peut-être plus facile de dissimuler une part d'activité. Le sens des liens entre les taux de valeur ajoutée et de marge, la propension à frauder de l'entreprise et la probabilité de sélection des services fiscaux n'est finalement pas évident *a priori*, mais l'efficacité de ces ratios dans une stratification méritait d'être testée.

Par ailleurs, il arrive que les ratios comptables envisagés ne soient pas calculables à partir de Suse. En pareil cas, soit les comptes n'ont pas été déposés, soit ils sont incomplets ; en tout état de cause, la comptabilité de l'entreprise risque d'être mal tenue ou douteuse. Or ces situations semblent être particulièrement visées par les services fiscaux. Par conséquent, même si les ratios comptables comportent des valeurs manquantes, ils demeurent intéressants et un exemple précis le démontrera (*cf. § V et annexe 2*).

Au vu de cette analyse, le taux de valeur ajoutée et le taux de marge, découpés en tranches, ont été introduits dans une deuxième stratification, en plus de l'activité, de la taille et de la catégorie juridique. Pour ne pas démultiplier à l'excès le nombre des strates, la codification du secteur d'activité a été simplifiée. De fait, les résultats montrent que la fraude se concentre sur une vingtaine de secteurs d'activité de la NES en 118 postes (niveau G). Seuls ces derniers sont isolés ; les autres sont décrits à un niveau plus agrégé de la nomenclature d'activité (niveau E de la NES). D'autres éléments de compte seraient sans

¹³ NES : nomenclature économique de synthèse, utilisée par les comptes nationaux en base 2000.

¹⁴ Toutefois, la variable de taille et le type d'imposition ne sont pas définis de la même manière et la technique d'extrapolation n'est pas non plus totalement similaire.

¹⁵ Le taux de valeur ajoutée se définit comme le rapport entre la valeur ajoutée et le chiffre d'affaires de SUSE. Dans le cadre de cette étude, le taux de marge est défini comme le rapport entre le « profit brut courant avant impôt » (PBCAI) de SUSE et le chiffre d'affaires. Le PBCAI est défini au § III-2-2, note de bas de page n°11. Ces taux sont des variables continues ; pour être utilisables dans la stratification, elles ont été découpées en tranches.

doute pertinents, notamment des données de bilan, mais du fait de la cohérence interne de la comptabilité d'entreprise, les mobiliser risquait d'engendrer des redondances statistiques. Ces dernières sont d'ailleurs mises en évidence par des régressions (cf. § IV-1, IV-2-2 et IV-2-3).

IV.2.3 Deux méthodes d'investigation : les régressions logistiques et la statistique descriptive

Il est difficile de savoir *a priori* si une stratification est optimale et aucun algorithme ne permet de la déterminer d'emblée. Deux problèmes se posent : la stratification reflète-t-elle convenablement le plan de sélection de la Dgfip, d'abord ? Ensuite, le plan de sondage ainsi défini conduit-il à des propensions à frauder homogènes au sein d'une strate et bien différenciées d'une strate à l'autre ? Ce sont là deux questions différentes, bien que la sélection des entreprises contrôlées ne se fasse pas indépendamment de leur taux de fraude présumé.

Pour répondre à la première interrogation, on dispose de méthodes économétriques (Annexe 3). Les régressions logistiques permettent de modéliser les probabilités de tirage des services fiscaux. Le rapprochement de l'échantillon avec Suse rend possible la mise en œuvre de telles méthodes puisque l'on peut alors repérer, parmi les entreprises de la population, celles qui ont été contrôlées et créer ainsi une variable binaire, égale à 1 si l'entreprise est contrôlée, à 0 sinon. C'est cette variable, représentative du tirage, que l'on cherche à expliquer. Au vu d'une batterie d'indicateurs, on détermine ainsi les facteurs les plus explicatifs des probabilités de tirage en testant simultanément un très grand nombre de variables. Les variables de stratification mentionnées ci-dessus apparaissent comme les plus significatives, au vu des tests des régressions logistiques. On a en outre testé l'appartenance à un groupe et un indicateur de productivité. Celui-ci est intéressant mais inutilisable en raison d'un nombre trop important de valeurs manquantes. Ces investigations laissent penser que la seconde stratification offre une représentation convenable du plan de tirage de l'échantillon.

Il convient en outre de vérifier le caractère suffisamment discriminant des strates ainsi définies, au regard de la propension à frauder. Bien que la variance des taux de fraude demeure sensible au sein de chaque strate, on suppose que, si les taux moyens par strate sont bien différenciés, chacun d'eux peut s'appliquer à toutes les entreprises de la strate considérée. Cela revient à admettre que la variance intra-strate est faible, comparée à la variance inter-strate et que l'on peut alors accepter l'hypothèse d'uniformité des taux de fraude sur une strate. Ces considérations reposent sur des analyses de statistique descriptive, menées sur les taux de fraude élémentaires.

On repère parfois que des liaisons fortes entre certaines variables et les probabilités de tirage sont moins nettes au vu des analyses statistiques sur les taux de fraude. En réalité, on ne teste pas exactement la même chose. Le tirage des services fiscaux, modélisé par les régressions, est effectué en fonction de l'ensemble des chefs de rectification : toutes les composantes de la fraude sont alors prises en compte. En revanche, les taux de fraude, pris au sens de cette étude, ne sont pas représentatifs de l'ensemble des comportements de fraude de l'entreprise car la fraude est alors restreinte à certains chefs de rectification (cf. § III). Dans ces conditions, il est logique que les corrélations mises en évidence pour les probabilités de tirage soient parfois moins nettes pour les taux de fraude. Concrètement, la faiblesse d'un taux de marge, par exemple, peut être la conséquence d'une fraude portant sur les amortissements, laquelle n'intervient pas dans le calcul des taux de fraude sur le résultat, pris au sens de cette étude. En outre, si les probabilités de tirage ne sont pas sans lien avec les taux de fraude présumés des entreprises, ces derniers ne sont pas les seuls éléments de sélection de la Dgfip ; celle-ci vise notamment les entreprises en fonction de leur taille.

Pour réduire les biais d'extrapolation, c'est d'abord les probabilités de tirage qu'il faut modéliser. Les régressions logistiques constituent le mode de sélection des variables le plus approprié puisqu'il ne repose pas sur une mesure partielle de la fraude observée (par

opposition à l'analyse des taux de fraude), mais uniquement sur des inférences statistiques entre l'échantillon et la population.

IV.2.4 Quelques éléments concrets d'analyse statistique

Au vu des régressions, la taille - en tranches de chiffre d'affaires - est la variable la plus explicative des probabilités d'inclusion, quel que soit le type d'imposition. Les très grandes entreprises, soumises au régime des bénéficiaires réels (BIC-BRN), ont presque neuf fois plus de chances d'être sélectionnées que les plus petites (chiffre d'affaires inférieur à 100 000 euros), les entreprises dont le chiffre d'affaires est compris entre 10 et 20 millions d'euros, presque sept fois plus de chances ; pour celles dont le chiffre d'affaires est compris entre 3 et 10 millions d'euros, ce rapport est de l'ordre de 4 (*Annexe 3, tableaux A3-1 et A3-2*). Dans le cas des entreprises assujetties au régime de la déclaration simplifiée (BIC-RSI), le rapport entre les probabilités de contrôle des entreprises les plus grandes, par référence aux plus petites, dépasse très nettement 20. La significativité de la catégorie juridique, du taux de valeur ajoutée et du taux de marge est également démontrée par les régressions.

Les analyses de statistique descriptive confirment le rôle de la taille de l'entreprise pour expliquer les taux de sondage. Toutefois, ces derniers sont à rapprocher de la stratégie du contrôle fiscal, laquelle consiste à lutter contre la fraude en couvrant de manière harmonieuse l'ensemble des contribuables tout en assurant une présence sur les enjeux. La taille de l'entreprise joue également sur les taux de fraude (*tableaux 3 et 4 et graphiques 2 et 3*). Les grandes entreprises ne sont pas celles dont les taux de fraude, calculés dans le cadre de l'étude, sont les plus forts. Au contraire, on constate qu'ils décroissent lorsque la taille augmente. Ainsi, pour les entreprises soumises au régime des BRN, le taux de fraude sur le résultat des plus petites (moins de 100 000 euros de chiffre d'affaires annuel) est en moyenne de 16,3 % ; en revanche, dès que le chiffre d'affaires de ces entreprises dépasse le million d'euros, ce taux est nettement plus faible ; toutefois, il porte sur des montants beaucoup plus élevés. Pour les entreprises assujetties au régime simplifié (RSI), ce même taux de fraude sur le résultat dépasse 65 % pour les plus petites, contre 4,9 % seulement pour les plus grandes (Annexe 3).

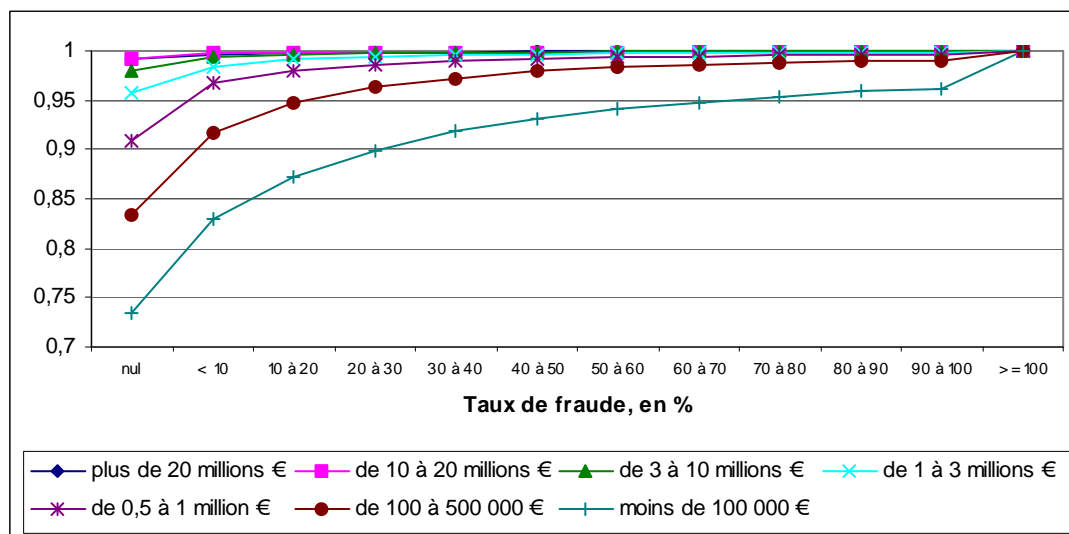
Tableau 3 : BRN - Taux de sondage et taux de fraude moyens sur le résultat, selon la taille de l'entreprise, en %

BRN		
	Taux de fraude Moyenne simple	Taux de sondage
Plus de 20 millions €	0,36	40,2
De 10 à 20 millions €	0,45	33,4
De 3 à 10 millions €	0,66	22,3
De 1 à 3 millions €	1,47	11,5
De 500 000 à 1 million €	3,07	8,9
De 100 000 à 500 000 €	6,45	6,7
Moins de 100 000 €	16,27	4,8

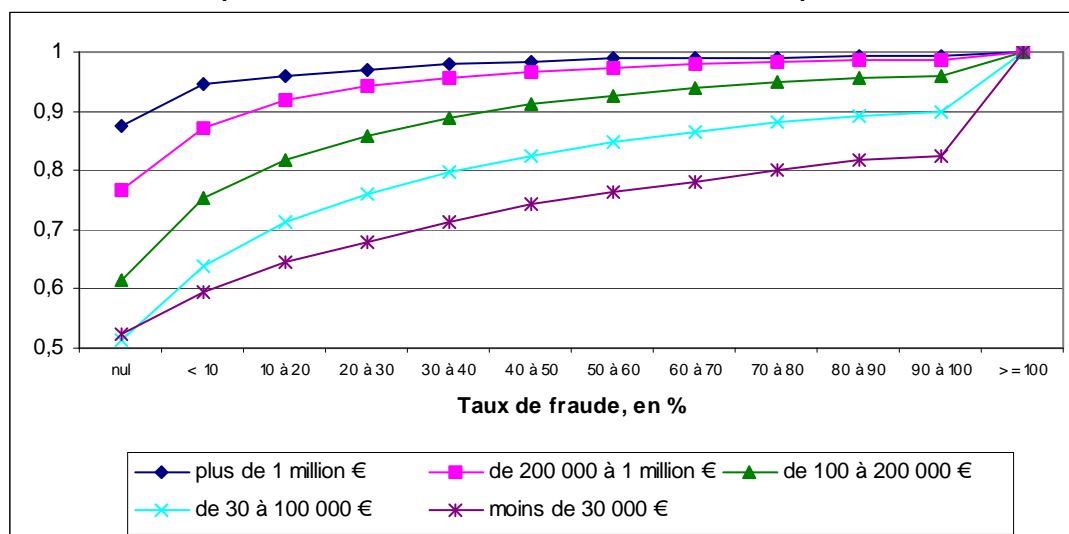
Tableau 4 : RSI - Taux de sondage et taux de fraude moyens sur le résultat, selon la taille de l'entreprise, en %

RSI		
	Taux de fraude Moyenne simple	Taux de sondage
Plus de 1 million €	4,9	29,8
De 200 000 à 1 million €	9,1	3,4
De 100 000 à 200 000 €	19,6	2,3
De 30 000 à 100 000 €	37,7	1,9
Moins de 30 000 €	65,2	1,0

Graphique 2 : BRN - Fonctions de répartition des taux de fraude sur le résultat, pour les différentes modalités de la taille des entreprises



Graphique 3 : RSI - Fonctions de répartition des taux de fraude sur le résultat, pour les différentes modalités de la taille des entreprises



Lecture : Les fonctions de répartition des graphiques 2 et 3 ainsi que celles des graphiques 4, 5 et 6 s'interprètent de la manière suivante : plus faible est la proportion des entreprises d'une strate à n'avoir fait l'objet d'aucune rectification sur le résultat (taux de fraude nul), plus la courbe est pentue et le taux de fraude moyen élevé. Les fréquences des taux de fraude les plus faibles étant en général plus fortes que celles des taux de fraude les plus élevés, la convexité des courbes est plus accentuée au départ.

Dans le cas des entreprises des BRN, celles qui réalisent moins de 100 000 € de chiffre d'affaires annuel se démarquent nettement en termes de propension à frauder : 73,5 % d'entre elles seulement ne fraudent pas, contre plus de 99 % de celles réalisant plus de 10 millions € de chiffre d'affaires, et plus de 16 % de ces entreprises ont un taux de fraude positif, mais inférieur à 30 %.

Dans le cas des RSI, la proportion des entreprises qui ne fraudent pas augmente graduellement avec la taille, passant de 51 % à 88 %. On remarque que cette proportion est globalement inférieure à celle des entreprises des BRN (toujours supérieure à 70 %). Pour les deux tranches de taille inférieures, la convexité de la courbe est moins marquée au départ ; la concentration vers les taux de fraude faibles est donc moins accentuée que pour les BRN.

Ces graphiques soulignent le rôle crucial de la proportion d'entreprises non rectifiées (taux de fraude nuls).

Les régressions mettent en évidence un effet nettement significatif du taux de valeur ajoutée et du taux de marge sur les probabilités de tirage, pour tous les types d'imposition. Si une relation inverse entre les ratios comptables et les taux de fraude est également repérable dans certains secteurs, elle n'est pas pour autant systématique. En revanche, le simple fait que ces ratios soient ou non renseignés semble porteur d'une information importante (*graphiques 5 et 6*). Ce que l'on pouvait anticiper se confirme (*cf. § IV-2-2*) : les entreprises pour lesquelles ces ratios ne sont pas disponibles, au demeurant peu fréquentes dans la population, sont surreprésentées dans l'échantillon, comme si elles étaient particulièrement recherchées par les services fiscaux ; de fait, elles se révèlent très fraudeuses. Ces entreprises seraient donc bien « ciblées ». Dès lors que des éléments de compte sont introduits dans la stratification, ce ciblage se trouve *de facto* pris en compte.

Les résultats des différentes simulations montrent en outre que ces ratios ont un impact très sensible sur les évaluations macroéconomiques, lesquelles sont, en l'absence de ces variables, surestimées globalement, et en général au niveau sectoriel. Au final, les travaux statistiques, tout comme les résultats macroéconomiques, convergent pour démontrer l'intérêt des ratios comptables dans la stratification, sans dénier le rôle majeur de la variable de taille. Pour procéder aux redressements pour fraude de la base 2005, on a donc privilégié la seconde stratification, mais les résultats de plusieurs simulations sont également présentés dans ce document (*cf. § V*). Cette modélisation enrichie du tirage n'en demeure pas moins assez fruste, au regard de la complexité du mode de sélection réel, ce que traduit la part relativement importante de l'inexpliqué dans les régressions. Elle a toutefois le mérite d'apporter quelques éléments de réponse concrets aux réserves, jusqu'alors formulées sans preuves, à l'encontre des exploitations passées des contrôles fiscaux.

Tableau 5 : BRN, RSI - Taux de sondage et taux de fraude moyens sur le résultat, selon les modalités du taux de valeur ajoutée, en %

Modalités du taux de valeur ajoutée	Moyenne simple des taux de fraude observés sur le résultat		Moyenne, pondérée par les chiffres d'affaires, des taux de fraude observés sur le résultat		Taux de sondage	
	BRN	RSI	BRN	RSI	BRN	RSI
Non renseigné	11,2	36,4	0,33	14,9	10,63 %	9,98 %
Inférieur à 0	16,1	43,7	2,13	23,5	9,48 %	1,34 %
De 0 à 50	2,3	19,6	0,23	5,9	11,91 %	1,99 %
Supérieur à 0	3,2	28,8	0,52	12,1	6,85 %	1,22 %

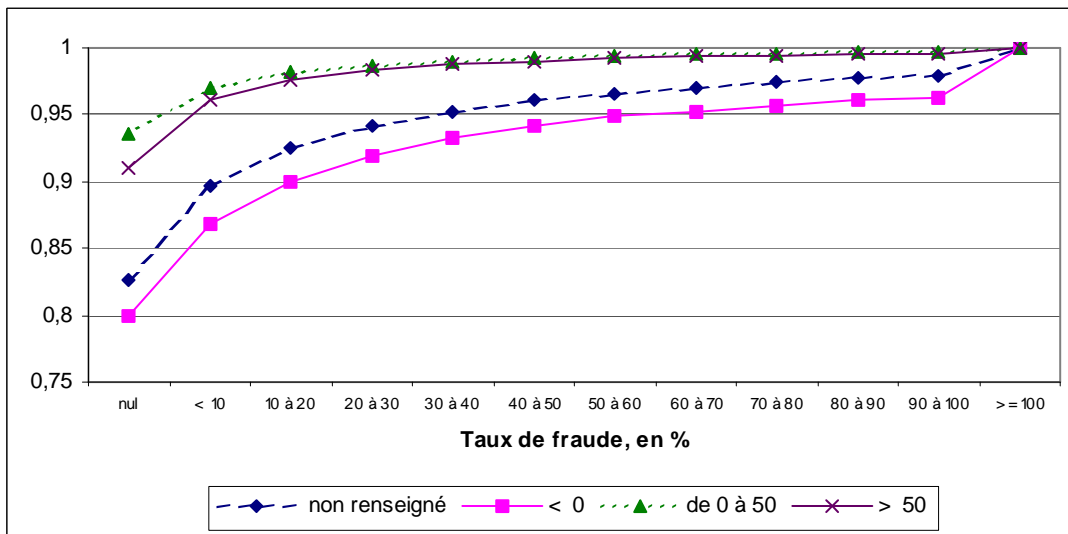
Tableau 6 : BNC - Taux de sondage et taux de fraude moyens sur le résultat, selon les modalités du taux de valeur ajoutée, en %

Modalités du taux de valeur ajoutée	BNC		Taux de sondage
	Moyenne des taux de fraude observés sur le résultat		
	Simple	Pondérée par les chiffres d'affaires	
Non renseigné	40,5	18,7	20,79
Inférieur à 25	35,0	14,1	1,53
De 25 à 50	17,3	7,7	1,71
Supérieur à 50	13,4	5,5	0,91

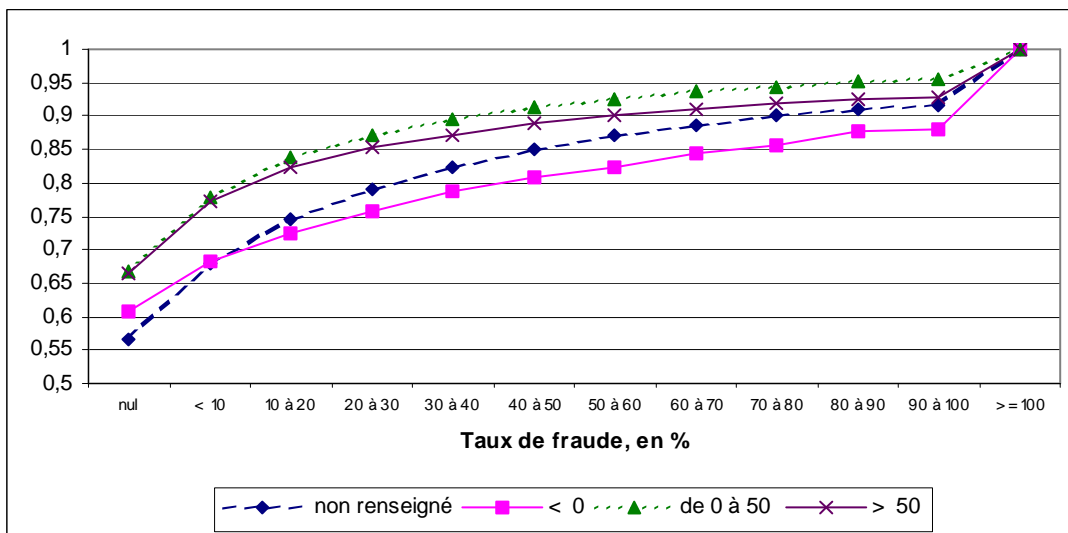
Lecture : Les tableaux 5 et 6 et les graphiques 4, 5 et 6 mettent en évidence l'effet du taux de valeur ajoutée sur les taux de fraude et sur les taux de sondage. Moins fort que celui de la taille, il est néanmoins réel. Pour les RSI et les BNC, les taux de sondage sont nettement plus forts lorsque le taux de valeur ajoutée n'est pas renseigné, ce qui traduit bien le « ciblage » des entreprises concernées.

Dans le cas des entreprises soumises au RSI, les proportions d'entreprises non rectifiées sont manifestement plus élevées pour celles dont les taux de valeur ajoutée sont positifs : 67 % contre 60 %, voire moins pour les autres (*graphique 5*). En outre, la fréquence des taux de fraude élevés (> 60 %) est plus forte pour les entreprises dont le taux de valeur ajoutée est négatif : 15,7 % contre 6,4 % pour celles dont le taux de valeur ajoutée est compris entre 0 et 50.

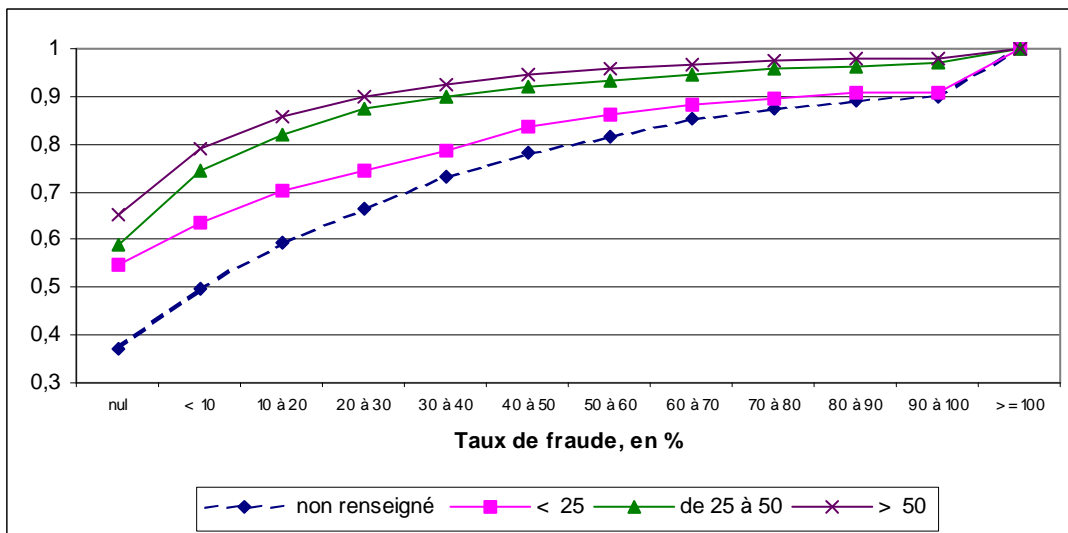
Graphique 4 : BRN - Fonctions de répartition des taux de fraude sur le résultat, pour les différentes modalités du taux de valeur ajoutée



Graphique 5 : RSI - Fonctions de répartition des taux de fraude sur le résultat, pour les différentes modalités du taux de valeur ajoutée



Graphique 6 : RSI - Fonctions de répartition des taux de fraude sur le résultat, pour les différentes modalités du taux de valeur ajoutée



IV.3 Les estimations par post-stratification

IV.3.1 Mise en œuvre de la stratification

Chaque type d'imposition (BIC-BRN, BIC-RSI, BNC) a fait l'objet d'un calcul séparé. Les variables retenues dans les deux stratifications de cette étude sont décrites dans l'annexe 2. L'enchaînement des étapes successives du calcul est formalisé dans l'encadré ci-dessous.

Encadré 2 : Calcul des montants de fraude par extrapolation

Mode opératoire

- Calcul des taux de fraude observés $tf_{n,t,e}$ sur les strates les plus fines (n,t,e) de l'échantillon

$$tf_{n,t,e} = \frac{\sum_{i \in (n,t,e)} rec_i}{\sum_{i \in (n,t,e)} ca_i} = \sum_{i \in (n,t,e)} \left(\frac{ca_i}{\sum_{i \in (n,t,e)} ca_i} \right) \times \left(\frac{rec_i}{ca_i} \right)$$

ou encore :

$$tf_{n,t,e} = \sum_{i \in (n,t,e)} \left(\frac{ca_i}{\sum_{i \in (n,t,e)} ca_i} \right) \times tf_i$$

- n désigne une première variable de stratification, l'activité par exemple, t une seconde variable de stratification, la taille par exemple ;
- e signifie que l'entreprise appartient à l'échantillon ;
- rec_i désigne la rectification (sur le chiffre d'affaires ou sur le résultat) de l'entreprise i de l'échantillon ; elle est relative à l'ensemble de la période de contrôle ;
- ca_i = chiffre d'affaires de la période de contrôle de l'entreprise i ;
- le taux de fraude de l'entreprise i , tf_i , se définit comme le rapport de ces deux grandeurs.

On remarque que les taux de fraude observés des strates fines s'expriment aussi comme des moyennes des taux de fraude élémentaires, pondérées par les chiffres d'affaires.

- Extrapolation de la fraude aux strates (n,t,P) les plus fines de la population

Les strates (n,t,P) sont définies en croisant les mêmes variables que les strates (n,t,e) .

- P signifie que l'entreprise appartient à la population ;
- les chiffres d'affaires sont cette fois les chiffres d'affaires annuels des entreprises de la population ;
- les rectifications sont également annuelles.

$$tf_{n,t,P} = \frac{\sum_{i \in (n,t,P)} rec_i}{\sum_{i \in (n,t,P)} ca_i} = \frac{\sum_{i \in (n,t,P)} ca_i \times tf_{n,t,e}}{\sum_{i \in (n,t,P)} ca_i}$$

ou encore :

$$tf_{n,t,P} = \sum_{i \in (n,t,P)} \left(\frac{ca_i}{\sum_{i \in (n,t,P)} ca_i} \right) \times tf_{n,t,e} = tf_{n,t,e}$$

Dans cette étape, on calcule la rectification de chaque entreprise de la strate (n,t,P) de la population en multipliant le taux de fraude observé de la strate correspondante de l'échantillon $tf_{n,t,e}$, calculé ci-dessus, sur la période de contrôle, par le chiffre d'affaires annuel de l'entreprise considérée, au cours de l'année de calage (en l'occurrence 2006) ; puis on agrège ces rectifications annuelles sur la strate (n,t,P) . On obtient ainsi la masse de rectifications de cette strate pour l'année de calage que l'on divise par la masse

de chiffres d'affaires correspondante pour obtenir le taux de fraude de cette strate (n,t,P) . Ce calcul revient à admettre que le taux de fraude des entreprises d'une strate élémentaire de la population au cours de l'année de calage est uniforme et égal au taux de fraude moyen de la strate correspondante de l'échantillon, mesuré sur la période de contrôle.

- Passage au niveau d'agrégation supérieur : calcul des montants de fraude et des taux de fraude $tf_{n,P}$ des strates (n,P) plus agrégées de la population

Les strates (n,P) sont définies en supprimant une variable de la stratification (n,t,P) (ou plusieurs variables si la première stratification comporte plus de deux variables). Avec les notations ci-dessus, on supprime la variable t ; c'est-à-dire qu'on agrège selon t .

$$tf_{n,P} = \frac{\sum_{t \in (n,P)} rec_{(n,t,P)}}{\sum_{t \in (n,P)} ca_{(n,t,P)}} = \sum_{t \in (n,P)} \left(\frac{ca_{(n,t,P)}}{\sum_{t \in (n,P)} ca_{n,t,P}} \right) \times tf_{n,t,P}$$

Cette étape consiste à calculer la masse de rectifications de chaque strate agrégée (n,P) de la population en sommant sur t les masses de rectifications des strates fines (n,t,P) ; on fait de même sur les masses de chiffre d'affaires et l'on obtient le taux de la strate (n,P) en rapportant la masse des rectifications à celle des chiffres d'affaires. *Ce taux agrégé s'exprime comme la moyenne des taux les plus fins $tf_{n,t,P}$, pondérée par la structure selon t des masses de chiffres d'affaires de la strate (n,P) de la population.*

- Remarque : interprétation d'une stratification moins fine

Si la stratification comporte une variable (t) en moins, le taux de fraude de la strate (n,e) de l'échantillon peut fort bien s'exprimer virtuellement comme une moyenne, selon la variable absente t , des taux de fraude des strates plus fines en (n,t,e) , effectuée, dans ce cas, avec les pondérations de l'échantillon.

$$\begin{aligned} tf'_{n,P} = tf_{n,e} &= \frac{\sum_{t \in (n,e)} \sum_{i \in (n,t,e)} rec_i}{\sum_{t \in (n,e)} \sum_{i \in (n,t,e)} ca_i} = \left(\frac{1}{ca_{n,e}} \right) \times \sum_{t \in (n,e)} (ca_{n,t,e}) \sum_{i \in (n,t,e)} \left(\frac{ca_i}{ca_{n,t,e}} \right) \times tf_i \\ &= \sum_{t \in (n,e)} \left(\frac{ca_{n,t,e}}{ca_{n,e}} \right) \times tf_{n,t,e} \end{aligned}$$

On voit alors que le taux de fraude de la strate (n,P) ci-dessus $tf'_{n,P}$ diffère du précédent $tf_{n,P}$ par la structure des pondérations en chiffres d'affaires appliquées aux taux de fraude les plus fins ($tf_{n,t,P}$ ou $tf_{n,t,e}$) pour passer aux taux plus agrégés : pour le calcul de $tf'_{n,P}$, cette structure est *celle de l'échantillon* alors qu'il s'agissait de la structure de la population pour $tf_{n,P}$. Cette remarque joue un rôle crucial dans la compréhension des mécanismes d'extrapolation, selon que la stratification est plus ou moins fine (cf tableau 2, § V et annexe 3).

La programmation des calculs a été effectuée à l'aide du logiciel SAS. On a privilégié l'expression des taux de fraude sous forme de moyennes de taux plus fins, car elle offre plus de lisibilité et conduit à une programmation un peu plus courte.

Le redressement des taux manquants

Lorsqu'une strate fine de l'échantillon est vide, ce qui est fréquent, les taux de fraude correspondants ne sont pas calculables ; on a alors procédé à des redressements avant d'agréger. Ne pas le faire reviendrait à considérer qu'aucune entreprise de cette strate ne fraude dans la population, ce qui ne paraît pas très vraisemblable, une strate pouvant fort bien ne pas être représentée dans l'échantillon, compte tenu de la finesse de la stratification. Les taux de fraude manquants ont été remplacés par des taux plus agrégés, calculés à partir des taux renseignés des strates fines, non vides, de même niveau, ce qui revient à les redresser sur la base de la moyenne des taux renseignés de même niveau. À un niveau plus agrégé, soit en (secteur d'activité * catégorie juridique), et pour un type d'imposition donné, si une strate n'est pas représentée dans l'échantillon, on peut admettre que la strate agrégée en question ne contribue pas, ou très peu, au déficit d'impôt que la Dgfiip cherche à récupérer et se dispenser de redresser. Si la Dgfiip ne procède à aucun contrôle dans ces

strates, c'est probablement que c'est inutile. Cette hypothèse est admissible à ce niveau d'agrégation.

Des calculs précédents, effectués par type d'imposition (BIC-BRN, BIC-RSI, BNC), découlent les estimations des montants de fraude, en NES, de chaque type d'imposition. Moyennant les hypothèses mentionnées au paragraphe ci-dessus, les montants de fraude, en NES, de l'ensemble des entreprises, tous types d'imposition confondus, s'obtiennent en sommant les montants de fraude de chacun des types d'imposition. Les taux de fraude, en NES, sont les rapports entre ces montants de fraude et les chiffres d'affaires correspondants.

IV.3.2 Interprétation de la post-stratification

Le calcul des agrégats fraudés par post-stratification, dont le mode opératoire est détaillé dans l'encadré ci-dessus, est assimilable à une extrapolation des montants observés sur l'échantillon, en supposant que celui-ci a été tiré selon un plan de sondage stratifié, avec tirage aléatoire simple dans chaque strate (cf. § IV-2). Les taux de fraude traduisent la propension à frauder des entreprises, supposée être la même au sein d'une strate.

Cette méthode d'extrapolation se formalise de la manière suivante.

Toutes les entreprises d'une strate ont la même probabilité d'être tirées, égale au rapport entre l'effectif de l'échantillon dans la strate et l'effectif de la population de la strate $\left(\frac{n}{N}\right)$.

Les grandeurs de l'échantillon sont extrapolées par les inverses des taux de sondage. Cela revient à faire une simple règle de trois : une entreprise pèse d'autant moins que sa probabilité de tirage est forte.

Le taux de fraude étant le même pour toutes les entreprises d'une strate, la formule d'extrapolation à la strate s de la population, des rectifications de l'échantillon s'écrit ainsi :

$$rec_{s,P} = \sum_{i \in (s,E)} rec_i \times \frac{N}{n} = \sum_{i \in (s,E)} tf \times ca_i \times \frac{N}{n} = tf \times \sum_{i \in (s,E)} ca_i \times \frac{N}{n} = tf \times ca_{extr}$$

(s,E) est la strate s de l'échantillon et (s,P) la strate s de la population ;
 ca_{extr} est le chiffre d'affaires extrapolé de la strate P.

Le calcul *a posteriori* présenté dans l'encadré 2 s'écarte toutefois légèrement de ce schéma. Le chiffre d'affaires des entreprises de la population étant connu (c'est celui de SUSE 2006), il est préférable de remplacer le chiffre d'affaires extrapolé ca_{extr} de la strate s de la population par le chiffre d'affaires réellement observé, noté ca , de cette strate : ceci revient à « caler » le chiffre d'affaires extrapolé sur le chiffre d'affaires observé.

$$rec_{s,P} = tf \times ca_{extr} \times \left(\frac{ca}{ca_{extr}}\right) = tf \times ca = \sum_{k \in (s,P)} tf \times ca_k$$

IV.3.3 Le traitement des rectifications non renseignées

Un autre problème important concerne la manière d'interpréter les rectifications qui ne sont pas renseignées dans le fichier de la Dgip. Si l'impôt correspondant n'a pas été contrôlé lors de la vérification, la rectification doit rester à valeur manquante. S'il a, au contraire, été vérifié, sans avoir été rectifié, alors la rectification doit être prise en compte, mais elle doit être mise à zéro car elle est effectivement nulle, et c'est le cas de la majeure partie des rectifications laissées « à blanc ». Plus précisément, on distingue trois types de contrôles : les contrôles généraux au cours desquelles toute la comptabilité est vérifiée et tous les impôts contrôlés, des contrôles spécifiquement dédiés à la TVA et d'autres contrôles, dits

« ciblés ». Dans les deux derniers cas, les rectifications non renseignées sont restées à valeur manquante ; en revanche, elles ont été mises à 0 pour tous les contrôles généraux, c'est-à-dire pour une très large majorité des entreprises vérifiées (environ 80 %). Ce traitement joue un rôle crucial car la proportion de rectifications laissées à valeur manquante est de nature à modifier sensiblement l'estimation des taux de fraude.

IV.3.4 Le changement de nomenclature : passage de la NES à la NAF-REV2

Le passage de la NES à la NAF-REV2 constitue l'une des principales innovations du changement de base dans lequel s'inscrit cette étude : *in fine*, les estimations de fraude devaient donc être déclinées selon la nouvelle nomenclature. Tous les calculs ont cependant été conduits en NES. Pour effectuer la conversion des taux de fraude, on a appliqué les taux extrapolés en NES aux chiffres d'affaires individuels. Ainsi, un montant de fraude a été affecté à chaque entreprise ; ces derniers ont ensuite été ré-agrégés au niveau A129 de la nouvelle nomenclature, puisque l'on disposait, individuellement, des deux nomenclatures. Les taux de fraude en A129 ont été obtenus en divisant les montants de fraude en A129 par les chiffres d'affaires correspondants. Les taux de salaires qui interviennent dans le bouclage comptable ont été convertis en A129 selon le même principe.

V - Les résultats des estimations

En l'absence de références vraiment comparables pour le passé, la discussion des valeurs obtenues est délicate. En particulier, les comparaisons sectorielles, selon la catégorie juridique ou par type d'imposition, risquent d'être affectées par l'utilisation de variables conformes aux normes de l'INSEE dans la présente étude, alors que les travaux précédents (hormis l'extrapolation) avaient été effectués dans un univers exclusivement DGI. Au-delà de ces réserves, pour apprécier l'impact des différentes hypothèses testées, il est intéressant de comparer les résultats des principales simulations.

Quelle stratification retenir ?

La sensibilité des résultats au choix de la stratification, toutes choses égales par ailleurs, est très grande (comparaison des variantes (b) et (d) du tableau 7) ; c'est ce qui frappe en premier lieu. La stratification simple (d) conduit à des montants sectoriels systématiquement supérieurs à ceux de la deuxième stratification. Globalement, l'écart $[(d) - (b) / (b)]$ atteint plus de 30 % sur la fraude sur le chiffre d'affaires des SNF, + 49,6 % sur celle des EI, soit + 43,5 % en tout. S'agissant de la fraude sur le résultat, ces écarts sont de + 19,2 % pour les SNF, + 20,6 % pour les EI, soit près de + 20 % en tout. En masse, la fraude sur le chiffre d'affaires issue de la stratification simple dépasse de 8,7 milliards d'euros celle de la seconde stratification ; la fraude sur le résultat est majorée d'un montant voisin. Ces divergences illustrent les effets sur les agrégats d'une stratification plus ou moins adaptée.

Tableau 7 : Synthèse des évaluations macroéconomiques

en milliards d'euros

	Fraude sur le chiffre d'affaires			Fraude sur le résultat		
	SNF	EI	Ensemble	SNF	EI	Ensemble
Stratification 2 et H1 (a)	7,2	11,4	18,6	18,1	19,3	37,4
Variante centrale : stratification 2 et H2 (b)	7,3	12,7	20,0	20,3	20,4	40,7
Variante centrale : stratification 2, H2, « extrapolation probas. logit » (c) (1)	7,4	12,6	20,0	19,8	20,6	40,4
Stratification simple et H2 (d)	9,7	19,0	28,7	24,2	24,6	48,8

Lecture : Le tableau ci-dessus présente quatre variantes combinant des hypothèses statistiques avec deux hypothèses de traitement des données fiscales (H1 et H2). Ces hypothèses sont les suivantes :

- la stratification simple ne croise que le secteur d'activité, la taille et la catégorie juridique ;
- la variante centrale est fondée sur la stratification 2, laquelle croise les variables de la stratification simple avec des ratios comptables (taux de valeur ajoutée et taux de marge, en tranches) et les hypothèses fiscales sont H2 ;
- hypothèses H1 : prise en compte des transferts de bénéfices et utilisation du taux officiel de 33,33 % pour transformer en base les rectifications en droits des entreprises soumises à l'IS ; pour remonter du « résultat établi par le service » au chiffre d'affaires correspondant, la part des charges est supposée être de 20 % du chiffre d'affaires ;
- hypothèses H2 : les rectifications pour transferts de bénéfices ne sont pas prises en compte et les taux d'imposition sur IS sont estimés à partir d'informations puisées dans SUSE ; pour remonter du « résultat établi par le service » au chiffre d'affaires correspondant, la part des charges est supposée être de 40 % du chiffre d'affaires.
- « extrapolation probas. logit » : les extrapolations sont fondées, non pas sur les inverses des taux de sondage, mais sur les inverses des probabilités d'inclusion estimées par les régressions logistiques.
- (1) variante du tableau A3-5

Pour les entreprises du secteur de la santé marchande, l'effet de l'introduction des ratios comptables dans la stratification a un effet baissier spectaculaire (*Annexe 2*). Ce cas est de loin le plus marquant, mais les divergences sectorielles entre les estimations issues des deux méthodes étant assez systématiquement de même sens, des raisonnements similaires à ceux qui sont développés dans l'annexe 2 pourraient sans aucun doute s'appliquer à bien d'autres activités.

S'il existe, au sein d'une classe de la stratification simple, de fortes disparités dans les comportements de fraude ; si de plus ces disparités interviennent dans la sélection des entreprises contrôlées, ce qui conduit à surreprésenter dans l'échantillon les entreprises les plus fraudeuses, on introduirait un biais important en ne cherchant pas à créer des strates plus homogènes, par subdivision des strates initiales, pour tenir compte de cette surreprésentation (§ IV-2). **La seconde stratification a donc été retenue pour construire la variante centrale des estimations, laquelle a été privilégiée pour effectuer le bouclage de la comptabilité nationale.**

L'impact des autres hypothèses

Le choix de la méthode d'extrapolation (post-stratification ou inverse des probabilités issues des régressions logistiques) n'a quasiment pas d'incidence sur les résultats macroéconomiques (*Annexe 3*). L'exercice consistant à extrapoler avec les inverses des probabilités des régressions logistiques présente un intérêt méthodologique ; mais les variables explicatives finalement retenues dans les régressions coïncidant avec celles de la stratification, les extrapolations issues des deux approches sont très voisines. Les stratifications présentent l'avantage de prendre en compte des effets croisés ; par ailleurs, le mécanisme de l'extrapolation par cette méthode étant relativement lisible, il peut être assez facilement décrypté dans les cas délicats (*Annexe 2*). En revanche, les probabilités des régressions, qui résultent d'un algorithme complexe fondé sur la méthode du maximum de vraisemblance, sont plus difficiles à interpréter que les taux de sondage d'une stratification.

Les divergences sur la fraude sur le résultat entre les variantes (a) et (b) sont dues au traitement des données de contrôle. L'élimination des transferts de bénéficiaires dans la variante (b) minore la fraude sur le résultat d'environ 3 milliards d'euros. En revanche, les rectifications sur le résultat de cette même variante sont sensiblement plus hautes du fait de l'utilisation, pour passer des droits rappelés aux bases, de taux d'imposition sur IS estimés, inférieurs au taux officiel de 33,33 %. La sensibilité des estimations à ces changements de taux d'imposition est forte : toutes choses égales d'ailleurs, ils rehaussent la fraude sur le résultat des entreprises soumises à l'IS d'environ +6 milliards d'euros dans la variante centrale (b).

La fraude sur le chiffre d'affaires est également un peu plus forte dans la variante (b), comparée à celle de la variante (a). Ce rehaussement est imputable à l'hypothèse relative à la part des charges permettant de passer du « résultat établi par le service » au chiffre d'affaires correspondant : le résultat représente 80 % du chiffre d'affaires dans la variante (a), contre 60 % dans la variante (b).

La masse de fraude sur le chiffre d'affaires

La masse totale de fraude sur le chiffre d'affaires est évaluée à 20 milliards d'euros dans la variante centrale (7,3 pour les SNF et 12,7 pour les EI). Elle se concentre sur une vingtaine de secteurs d'activité (parmi une centaine en NES 118) qui se trouvent, en général, être ceux qui réalisent les plus gros chiffres d'affaires. Dit autrement, la propension à frauder d'un secteur n'est pas le seul facteur explicatif du montant de son activité fraudée.

D'une manière générale, la propension à frauder sur le chiffre d'affaires des entreprises individuelles est nettement plus forte que celle des sociétés non financières : 7,9 % en moyenne contre 0,23 % ; c'est la raison pour laquelle les premières contribuent, à raison de 63,5 %, à la masse totale d'activité fraudée, alors que leur chiffre d'affaires est vingt fois plus faible que celui des SNF (160,5 milliards d'euros contre 3 202 pour les sociétés non financières).

La masse de fraude sur le résultat

La masse de fraude sur le résultat se monte à 40,7 milliards d'euros dans la variante centrale et se répartit à parts à peu près égales entre les SNF et les EI. Elle représente, en première approximation, la majeure partie des revenus non déclarés par les entreprises à

l'administration fiscale et qui seront réintégrés dans le calcul du PIB, l'ensemble du redressement comprenant en plus le travail au noir (cf. § VI).

La masse de fraude sur le résultat est aussi très concentrée sectoriellement : elle provient, à hauteur de 90 %, d'une vingtaine de secteurs seulement. Les secteurs dont les masses de fraude sur le résultat sont les plus importantes et ceux qui dissimulent le plus de recettes se classent à peu près selon la même hiérarchie puisque celle-ci dérive principalement de l'échelle des chiffres d'affaires.

Globalement, le montant de la fraude sur le résultat est deux fois plus élevé que celui de la fraude sur le chiffre d'affaires, la dissimulation d'activité n'étant qu'une des causes de fraude sur le résultat, parmi bien d'autres. Pour les SNF, comme pour les EI, le taux de fraude moyen sur le résultat est plus élevé que le taux de fraude moyen sur le chiffre d'affaires : 0,63 % contre 0,23 % pour les SNF, 12,7 % contre 7,9 % pour les EI. Le différentiel de taux est relativement plus important pour les SNF, ce qui pourrait être lié à la manière de frauder de ces entreprises : la fraude sur les frais généraux, qui ne se répercute que sur le résultat, aurait pris le pas sur la dissimulation de chiffre d'affaires, que des modes de gestion plus sophistiqués ont rendue plus difficile. En masse, les écarts sont aussi très sensibles : dans le cas des EI, la fraude sur le résultat dépasse de 7,7 milliards d'euros la fraude sur le chiffre d'affaires, mais cet écart atteint 13 milliards d'euros pour les SNF.

La fraude selon la taille de l'entreprise (tableaux 8 et 9)

Dans le cas des entreprises soumises au régime des BRN, la répartition des masses de fraude selon la taille des entreprises n'est pas fortement déséquilibrée. En dépit de taux de fraude extrêmement bas, les plus grandes de ces entreprises (réalisant plus de 20 millions d'euros de chiffre d'affaires annuel) contribuent, à hauteur de 16,7 %, à la fraude sur le chiffre d'affaires, et à hauteur de 31,5 % à la fraude sur le résultat des BRN, en raison de leur poids dans l'économie : elles réalisent 69,3 % du chiffre d'affaires de ce type d'imposition. D'une manière générale, les taux de fraude des entreprises des BRN sont faibles, mais plus l'entreprise est petite, plus ils sont élevés, de sorte que la moitié de la masse de fraude sur le résultat et 70 % de la fraude sur le chiffre d'affaires se concentrent sur les entreprises dont le chiffre d'affaires annuel est inférieur à 3 millions d'euros.

Les taux de fraude des entreprises assujetties au régime de la déclaration simplifié (RSI) sont en moyenne plus élevés. Ils croissent aussi en fonction inverse de la taille de l'entreprise, mais avec une plus forte dispersion. Ainsi, sur les très petites entreprises, dont le poids en chiffre d'affaires est de 2,4 % seulement, se concentrent 10,1 % de la fraude sur le résultat et 8,8 % de la fraude sur le chiffre d'affaires de ce type d'imposition. Environ 40 % de la fraude des RSI sont le fait d'entreprises réalisant entre 30 000 et 100 000 euros de chiffre d'affaires annuel, alors que leur poids (en chiffre d'affaires) n'est que de 16 %.

Tableau 8 :
La fraude des entreprises assujetties au régime des BRN selon la taille

	Fraude sur le chiffre d'affaires		Fraude sur le résultat		Chiffre d'affaires	Taux de sondage
	En % du total	Taux en %	En % du total	Taux en %	En % du total	En %
Plus de 20 millions €	16,7	0,03	31,5	0,18	69,3	40,2
De 10 à 20 millions €	3,6	0,06	4,9	0,28	7,1	33,4
De 3 à 10 millions €	9,4	0,10	10,0	0,38	10,6	22,3
De 1 à 3 millions €	20,4	0,30	17,9	0,91	7,9	11,5
De 500 000 à 1 million €	18,6	0,74	14,3	1,96	2,9	8,9
De 100 000 à 500 000 €	27,6	1,65	18,5	3,79	2,0	6,7
Moins de 100 000 €	3,6	2,66	3,0	7,56	0,2	4,8
Ensemble	100,0		100,0		100,0	

Tableau 9 : La fraude des entreprises assujetties au régime des RSI selon la taille

	Fraude sur le chiffre d'affaires		Fraude sur le résultat		Chiffre d'affaires	Taux de sondage
	En % du total	Taux en %	En % du total	Taux en %	En % du total	En %
Plus de 1 million €	0,8	0,8	1,6	2,8	7,0	29,8
De 200 000 à 1 million €	22,3	3,2	23,7	5,8	51,4	3,4
De 100 000 à 200 000 €	27,7	8,7	26,7	14,3	23,2	2,3
De 30 000 à 100 000 €	40,5	18,5	37,8	29,5	16,0	1,9
Moins de 30 000 €	8,8	27,1	10,1	53,3	2,4	1,0
Ensemble	100,0		100,0		100,0	

La fraude selon le type d'imposition

Globalement, la masse de fraude sur résultat est, à hauteur de 52,6 %, le fait des entreprises soumises au régime de la déclaration simplifiée (RSI). Les entreprises des BRN réalisent 31 % de cette fraude et celles des bénéficiaires non commerciaux, 16,4 %. En proportion de leur chiffre d'affaires, les entreprises des BRN fraudent beaucoup moins que les autres avec un taux de fraude sur le résultat de 0,4 %, contre 12,5 % pour celles des RSI et 10,8 % pour celles des BNC. Toutefois, le type d'imposition ne semble pas avoir en lui-même un effet important sur les masses de fraude ; il intervient davantage à travers ses liens avec la taille des entreprises, le secteur d'activité et la catégorie juridique.

Tableau 10 : La fraude selon le type d'imposition et la catégorie juridique*milliards d'euros*

	BRN, fraude sur le		RSI, fraude sur le		BNC, fraude sur le	
	chiffre d'affaires	résultat	chiffre d'affaires	résultat	chiffre d'affaires	résultat
SNF	3,2	11,6	3,9	8,2	0,2	0,5
EI	0,5	1,0	8,7	13,2	3,5	6,1
Ensemble	3,7	12,6	12,6	21,4	3,7	6,7

Encadré 3 La précision statistique des estimations macroéconomiques

En assimilant les estimations macroéconomiques obtenues par post-stratification à celles qui résulteraient d'un plan de sondage stratifié, il est possible d'en évaluer, en première approche, la précision statistique. Bien que ce calcul de précision soit approximatif, il permet d'apprécier la variabilité des données sur lesquelles s'appuient les estimations. Exprimée en pourcentage de valeur moyenne, la précision à 95 % des deux corrections macroéconomiques de la fraude (fraude sur le chiffre d'affaires et sur le résultat) est de l'ordre de 6 % ; on peut la considérer comme bonne, compte tenu de la nature des données et de la structure de l'échantillon (*tableau 11*). Au regard de la finalité de ces évaluations, qui est de corriger le PIB, le résultat peut même être jugé satisfaisant. Toutefois, à un niveau d'agrégation plus fin, la précision est moindre, par définition. La méthode de ce calcul est décrite ci-dessous.

La variance de l'estimateur du total d'une grandeur Y s'exprime comme la somme, sur toutes les strates, des estimations des variances des estimations des moyennes des strates, chacune d'elles étant multipliée par le carré du nombre d'entreprises de la strate dans la population.

Formellement, la formule de calcul est la suivante :

- estimateur de la variance de la grandeur Y sur une strate h de l'échantillon :

$$s_h^2 = \left(\frac{1}{n_h - 1} \right) \sum_{i_h=1}^{n_h} (y_{i_h} - \bar{y}_h)^2$$

- estimation de la variance de la moyenne de la grandeur Y sur une strate h de l'échantillon :

$$\hat{V}(\hat{Y}_h) = \frac{1}{n_h} \left(1 - \frac{n_h}{N_h} \right) s_h^2$$

- estimation de la variance de l'estimateur du total de la grandeur Y :

$$\hat{V}(\hat{T}(Y)) = \sum_{h=1}^k N_h^2 \hat{V}(\hat{Y}_h)$$

- avec les notations ci-dessous :

- $\hat{T}(Y)$: estimation du total de la grandeur Y ;
- N_h : nombre d'entreprises de la strate h de la population ;
- n_h : nombre d'entreprises de la strate h de l'échantillon ;
- y_{i_h} : montant observé de la grandeur y pour l'entreprise i de la strate h

Les estimations de la fraude ayant été effectuées sur la base de stratifications séparées pour chaque type d'imposition, cette formule a été appliquée par type d'imposition, et dans le cadre de la deuxième stratification (variante centrale), pour évaluer la variance de la fraude sur le chiffre d'affaires (omissions de recettes) d'une part, celle de la fraude sur le résultat (rectifications du résultat) d'autre part. L'estimation de la variance totale s'obtient en sommant les estimations de la variance des estimateurs relatifs à chaque type d'imposition. Cela revient à considérer que les échantillons sont tirés indépendamment dans chaque type d'imposition ; cette hypothèse paraît assez réaliste.

Les résultats sont résumés dans les tableaux 11 et 12 ci-dessous.

Tableau 11 : Écart-type des grandeurs estimées

milliards d'euros

	Fraude sur le chiffre d'affaires	Fraude sur le résultat
SNF	0,44	0,76
EI	0,42	0,90
Ensemble	0,61	1,18

Tableau 12 : Intervalle de confiance à 95 % des grandeurs moyennes

milliards d'euros

	Fraude sur le chiffre d'affaires		Fraude sur le résultat	
	Borne inférieure	Borne supérieure	Borne inférieure	Borne supérieure
SNF	6,4	8,2	18,7	21,9
EI	11,9	13,5	18,6	22,2
Ensemble	18,8	21,2	38,3	43,1

de 5,8 % pour la fraude sur le résultat. Pour les sociétés non financières, la précision à 95 % de la fraude

sur le chiffre d'affaires est de l'ordre de 12,2 %, celle de la fraude sur le résultat est de 7,6 %. Pour les entreprises individuelles, les chiffres correspondants sont de 6,6 % et de 8,9 %, respectivement

Un calcul par type d'imposition permet de localiser les estimations les plus précises (tableaux 13 et 14).

Tableau 13 : Écart-type de la fraude selon la catégorie juridique, par type d'imposition

milliards d'euros

	BRN, fraude sur le		RSI, fraude sur le		BNC, fraude sur le	
	chiffre d'affaires	résultat	chiffre d'affaires	résultat	chiffre d'affaires	résultat
SNF	0,26	0,01	0,36	0,76	0,03	0,07
EI	0,05	0,00	0,37	0,85	0,20	0,30
Ensemble	0,27	0,01	0,51	1,14	0,20	0,31

Tableau 14 : Précision à 95 % de l'estimation de la fraude, par type d'imposition

en % du montant moyen estimé

	BRN, fraude sur le		RSI, fraude sur le		BNC, fraude sur le	
	chiffre d'affaires	résultat	chiffre d'affaires	résultat	chiffre d'affaires	résultat
SNF	16,4	0,2	18,2	18,6	Non significatif	
EI	Non significatif		8,5	12,9	11,1	9,9
Ensemble	14,7	0,2	8,1	10,7	10,6	9,4

Pour les sociétés non financières soumises au régime des BRN, la précision de la fraude sur le résultat est particulièrement bonne (0,2 %). Les sociétés non financières étant majoritairement assujetties aux BRN, ceci explique que la précision de la fraude sur le résultat soit légèrement meilleure globalement pour cette catégorie juridique que pour les entreprises individuelles. La faiblesse de l'écart-type de la fraude sur le résultat des BRN est probablement une conséquence de celle de la variance intra-strate de ce type d'imposition. Certes la taille de l'échantillon des BRN est plus importante, ce qui constitue un facteur de gain de précision ; toutefois, en dépit de cela, la précision de la fraude sur le chiffre d'affaires est moins bonne, probablement en raison d'une plus forte variabilité de cette grandeur au sein des strates.

La faible taille des échantillons d'entreprises assujetties au régime simplifié (RSI) ou au régime des BNC a pour effet de réduire la précision des estimations pour ces deux types d'imposition. Dans ces sous-populations, on a probablement aussi plus de variabilité intra-strate.

VI - Simulation du bouclage de la comptabilité nationale

VI.1 Le bouclage théorique

Pour redresser le compte des entreprises de la comptabilité nationale, on dispose, à l'issue de ces calculs, de deux rectifications macroéconomiques : une rectification portant sur le chiffre d'affaires et une rectification relative à l'EBE, laquelle s'applique à la valeur ajoutée puisqu'il n'y a pas d'interventions sur les salaires lors des contrôles fiscaux. **Ainsi, à la production des comptes nationaux, on ajoute le chiffre d'affaires dissimulé, c'est-à-dire la rectification du chiffre d'affaires pour omission de recettes : $\Delta (P) = \text{rect} (CA)$. À la valeur ajoutée des comptes, on ajoute la rectification de l'EBE : $\Delta (VA) = \text{rect} (EBE)$.** Pour procéder au bouclage comptable, il n'a pas été nécessaire de mobiliser les rectifications spécifiquement opérées sur les charges lors des contrôles fiscaux. En effet, compte tenu de la nature particulière des rectifications retenues¹⁶, il est légitime d'admettre que la correction des consommations intermédiaires de la comptabilité nationale s'obtient en faisant la différence entre celle de la production et celle de la valeur ajoutée.

Pour évaluer la correction pour fraude de l'EBE de la comptabilité nationale, il faut en outre estimer la masse salariale dissimulée. Cette estimation repose sur plusieurs hypothèses. En premier lieu, la production dissimulée est, selon toute vraisemblance, le fruit d'un travail, lui aussi partiellement dissimulé (même s'il est également possible que des entreprises fraudent sur l'activité sans dissimuler des heures travaillées, auquel cas la productivité réelle des salariés est plus élevée que la productivité apparente de l'entreprise). Pour simplifier, on a donc considéré que la production dissimulée des entreprises était le fait du « travail au noir », et réciproquement. On a en outre supposé que la productivité du « travail au noir » était identique à la productivité moyenne de l'entreprise. Ainsi, il devient possible de calculer la masse salariale dissimulée - c'est-à-dire le travail au noir des entreprises - en vertu des hypothèses ci-dessus, en appliquant des taux de salaire¹⁷ sectoriels à la valeur ajoutée dissimulée de chaque secteur.

Reste à déterminer cette valeur ajoutée dissimulée. Pour l'évaluer, on a fait l'hypothèse qu'un chiffre d'affaires dissimulé ne s'accompagnait d'aucune dissimulation d'achats. La valeur ajoutée correspondant à la production fraudée est alors égale à cette production. Sans doute cette hypothèse est-elle un peu excessive. Il n'est pas exclu que des entreprises dissimulant des recettes cherchent à en atténuer partiellement l'impact sur le résultat en masquant aussi certains des achats afférents ; mais de telles situations sont probablement assez rares, car, ce faisant, les entreprises concernées se privent d'une partie de la TVA déductible et augmentent l'assiette de leur impôt sur les bénéfices.

Il se trouve, d'autre part, que l'hypothèse consistant à supposer qu'il n'y a pas de réduction des achats pour les entreprises qui dissimulent des recettes n'est pas démentie par les résultats des estimations macroéconomiques. De fait, la rectification macroéconomique de la production est à peu près deux fois plus faible que celle du résultat. La différence correspond donc, pour l'essentiel, à des rectifications pour rejet de charges, dues à une majoration injustifiée des frais généraux, des frais de personnel ou des consommations intermédiaires. Si les entreprises qui dissimulent des recettes faisaient de même pour les charges, le montant des rectifications pour rejet de charges serait encore plus élevé, car il devrait excéder l'écart macroéconomique entre la rectification sur le résultat et la rectification sur le chiffre d'affaires.

Par ailleurs, au niveau de l'entreprise, il est peu probable que celle-ci se permette de frauder, à la fois, par omission de recettes et augmentation induite de ses charges, même si une telle situation ne peut pas être totalement éliminée. Finalement, d'un point de vue macroéconomique, on peut considérer qu'il y a schématiquement deux grandes catégories

¹⁶ Les rectifications de nature purement fiscale, celles qui touchent aux amortissements ou aux provisions et les transferts de bénéfices ont été éliminés ; celles qui sont conservées portent donc principalement sur le chiffre d'affaires (omissions de recettes) et sur les charges.

¹⁷ Le taux de salaire se définit comme la part des salaires dans la valeur ajoutée.

d'entreprises fraudeuses, au sens des comptes nationaux : celles qui fraudent sur le chiffre d'affaires et celles qui ne le font pas, mais fraudent en majorant leurs charges. La masse salariale dissimulée ne serait le fait que des premières, pour lesquelles les consommations intermédiaires ne sont pas corrigées ; la correction macroéconomique des consommations intermédiaires, négative, ne concernerait que les entreprises fraudant par majoration induite de charges. Sur la base de cette dichotomie, la manière dont seront utilisées les rectifications macroéconomiques sur la production et sur le résultat en comptabilité nationale est résumée dans le tableau 15.

Tableau 15 : Le redressement du compte des entreprises

Entreprises fraudant sur le chiffre d'affaires		Entreprises fraudant sur le résultat		Ensemble des entreprises	
	$\Delta (P_1) = \text{rect.}(CA)$		$\Delta (P_2) = 0$		$\Delta (P) = \text{rect.}(CA)$
$\Delta (CI_1) = 0$		$\Delta (CI_2) < 0$		$\Delta(CI) = \Delta (CI_2)$	
	$\Delta(VA_1)=\Delta(P_1)$		$\Delta(VA_2)=-\Delta(CI_2)$		$\Delta(VA) = \text{rect.}(\text{res.})$
$\Delta (M_1) > 0$		$\Delta (M_2) = 0$		$\Delta (M) = \Delta (M1)$	
	$\Delta (EBE_1) = \Delta(VA_1) - \Delta (M1)$		$\Delta (EBE_2) = \Delta(VA_2)$		$\Delta (EBE) = \Delta(VA) - \Delta (M1)$

Note :

$\Delta (P) = \text{rect.}(CA)$: rectification macroéconomique pour omissions de recettes, issue de la présente étude.

$\Delta(VA) = \text{rect.}(\text{res.})$: rectification macroéconomique sur le résultat, issue de la présente étude.

$\Delta(CI_2) = \Delta(P) - \Delta(VA) = \text{rect.}(CA) - \text{rect.}(\text{res.})$

VI.2 Les simulations et les retraitements spécifiques à la comptabilité nationale

Dans le cadre de cette étude, on a effectué une simulation du bouclage comptable, à partir des chiffres d'affaires annuels de SUSE 2006. Les évaluations de la base 2005 proprement dites seront calées sur les chiffres d'affaires sectoriels des comptes nationaux. Ces derniers dérivent de ceux de SUSE, mais subissent en outre différents traitements visant essentiellement à compléter le champ de SUSE, la comptabilité nationale devant fournir une représentation de l'économie aussi exhaustive que possible. Ce calage comptable n'est pas de nature à modifier les résultats fondamentaux de cette étude et n'aura qu'un impact limité sur la mesure de la fraude, celle-ci étant finalement très faible, en proportion du chiffre d'affaires.

Le résultat des simulations du redressement du compte des entreprises, effectuées conformément au schéma théorique décrit ci-dessus, figure dans le tableau 16.

Tableau 16 : Le redressement du compte des entreprises, hors travail clandestin, avant retraitements comptables

	Sociétés non financières		Entreprises individuelles		Ensemble	
Production		7,3		12,7		20,0
Cons. Intermédiaires	- 12,9		- 7,7		- 20,6	
Valeur ajoutée		20,3		20,4		40,7
Masse salariale	5,1		3,7		8,8	
EBE		15,2		16,7		31,8

Note : De petits écarts d'équilibrage des totaux sont liés au fait que les chiffres ont été arrondis en milliards

Dans les faits, les redressements pour fraude de la comptabilité nationale ne sont pas organisés comme dans le tableau 16 ; ils font l'objet d'un retraitement. Le but visé est d'éviter de corriger la masse salariale pour ne pas risquer un désaccord avec les effectifs de SUSE, sur lesquels s'appuie le calcul de cette masse salariale. On considère que le travail

au noir est assimilable à un service, produit par des entreprises individuelles « virtuelles » et consommé intermédiairement par les SNF et par les EI. Dans ces conditions, la masse salariale n'a plus à être corrigée ; le travail au noir s'ajoute à la correction (négative) des consommations intermédiaires, pour les SNF et les EI, respectivement. Ce travail au noir, consommé par les SNF et les EI, étant produit par des EI « virtuelles », il vient également abonder la production des EI, et uniquement celle-ci. Ainsi, le travail au noir des entreprises n'a d'impact globalement, ni sur la valeur ajoutée, ni sur l'EBE, mais il gonfle la production et modifie le partage de la valeur ajoutée et celui de l'EBE, à l'avantage des entreprises individuelles. L'organisation des redressements est alors celle du tableau 17.

Tableau 17 : Le redressement du compte des entreprises, hors travail clandestin, après retraitements comptables

	Sociétés non financières		Entreprises individuelles		Ensemble	
Production		7,3		21,5		28,8
Cons. Intermédiaires	- 7,9		- 4,0		- 11,9	
Valeur ajoutée		15,2		25,5		40,7
Masse salariale	0		0		0	
EBE		15,2		25,5		40,7

milliards d'euros

VI.3 Le travail au noir

On distingue en général deux formes de travail au noir : les heures travaillées non déclarées - mentionnées ci-dessus - des entreprises ayant une existence juridique, et le travail totalement clandestin que l'on attribue, conventionnellement, à des entreprises individuelles n'ayant pas d'existence juridique. L'une des hypothèses retenues pour effectuer le bouclage comptable consiste à considérer que la production dissimulée n'est autre que le produit des heures de travail non déclarées. La masse salariale estimée ci-dessus fournit donc *de facto* une mesure du travail dissimulé des entreprises (cf. § VI - 2). Le travail clandestin (ateliers clandestins, petits travaux domestiques, notamment pour l'entretien du logement, etc.) doit aussi être intégré dans le calcul du PIB, mais il est estimé indépendamment. Son évaluation se heurte, depuis toujours, à la difficulté de réunir de l'information. En l'absence de nouvelle source, les taux de travail clandestin n'ont pas été actualisés dans le cadre de la présente étude. Deux pistes ont été explorées, mais elles se sont révélées, l'une et l'autre, inexploitable.

Une première piste aurait consisté à évaluer le travail au noir en s'appuyant sur la mesure de la fraude sociale, laquelle relève de l'activité de l'ACOSS¹⁸. Des taux de fraude sur les cotisations ont bien été publiés par cet organisme, mais ils sont inutilisables dans le cadre de la présente étude, pour plusieurs raisons : ils ne sont pas différenciés selon la nature du travail dissimulé (travail clandestin ou travail au noir des entreprises), ils ne sont pas déclinés selon la nomenclature des secteurs d'activité des comptes et surtout, ils portent sur la fraude aux cotisations et non pas sur la fraude sur les salaires, ce qui pose un délicat problème de remontée de l'une à l'autre.

Une autre piste aurait été de déduire le travail clandestin de la confrontation entre les effectifs employés issus d'une source administrative d'une part, d'une source déclarative d'autre part : les « travailleurs au noir » devraient, en théorie, être comptabilisés dans la seconde mais pas dans la première. Une esquisse de comparaison entre les effectifs employés mesurés par les DADS et ceux de l'enquête Emploi de l'INSEE a été tentée. L'exercice s'est révélé décevant, des problèmes de champ et de définition de l'emploi venant occulter les autres causes de divergence.

En conservant les taux de travail au noir clandestin de la base 2000, le redressement de la valeur ajoutée passerait de 40,7 milliards d'euros à 55,1 milliards d'euros. Le bouclage complet, simulé sur la base des chiffres d'affaires de Suse 2006, est alors celui du

¹⁸ Agence centrale des organismes de sécurité sociale

tableau 18 ci-dessous. L'introduction dans le calcul du PIB de l'activité non déclarée au fisc et du travail au noir clandestin rehausse ce dernier d'environ 3,0 %. Près des trois quarts de ce relèvement sont imputés sur les entreprises individuelles. Les entreprises individuelles ont de fait une propension à frauder plus forte que les sociétés non financières ; de plus, les retraitements comptables ont pour effet de leur attribuer en totalité la production liée au travail au noir.

En base 2000, la correction globale pour fraude et travail au noir de la valeur ajoutée des comptes nationaux était à peu près du même ordre, bien que son évaluation ait été fondée sur des méthodes notablement différentes de celle de la présente étude, tant du point de vue statistique qu'en ce qui concerne la nature des données de contrôle et la manière de les interpréter. Mettre point par point en parallèle deux calculs dont les cheminements sont parfois divergents était un exercice périlleux et fragile qui n'a pas pu être détaillé dans ce document.

Tableau 18 : Le redressement du compte des entreprises, y compris travail clandestin, après retraitements comptables

milliards d'euros

	Sociétés non financières		Entreprises individuelles		Ensemble	
Production		7,3		37,8		45,1
Cons. Intermédiaires	- 7,9		- 2,1		- 10,0	
Valeur ajoutée		15,2		39,9		55,1
Masse salariale	0		0		0	
EBE		15,2		39,9		55,1

Conclusion

Compte tenu des aléas qui entouraient la faisabilité de cette étude au démarrage, elle doit être considérée comme expérimentale. Dans un tel contexte, il est normal de procéder à des simulations pour tester des hypothèses et des méthodes. Les écarts entre les mesures de la fraude ainsi obtenues ne doivent pas être perçus comme inquiétants. Ces divergences alimentent au contraire la réflexion.

Concernant la manière de traiter les résultats des contrôles fiscaux, les simulations révèlent la très grande sensibilité des estimations de la fraude sur le résultat à l'hypothèse retenue pour passer des rectifications en « droits rappelés » aux rectifications « en base » correspondantes. L'idéal serait de ne mobiliser que des rectifications en base, mais il semble que ce ne soit pas possible.

D'un point de vue statistique, il ne faut pas perdre de vue le sens et les limites de l'exercice effectué. Le but ultime est une extrapolation dont le principe est d'étendre à toute une population les comportements de fraude observés sur un échantillon dont on aurait souhaité qu'il soit représentatif. On peut croire aux résultats macroéconomiques ainsi obtenus si cet échantillon est effectivement représentatif ou si la manière de l'utiliser permet de le considérer comme tel. L'organisation des données mobilisées pour effectuer la présente étude a permis d'approfondir les traitements statistiques, notamment d'améliorer la modélisation du « ciblage » des entreprises contrôlées. Pour autant, les précautions prises ne suffisent probablement pas à corriger toutes les conséquences sur les estimations d'un échantillonnage qui n'a pas été effectué en fonction de l'objectif visé.

Bibliographie

Blades D. et Roberts. D (2003), « Mesurer l'économie non observée », *Cahiers statistiques de l'OCDE*, N°5, janvier 2003

Cornu-Pauchet M. (2003), « L'évaluation de l'économie souterraine - Un recensement des études », *ACOSS, Stat Étude N°8*, septembre 2003

Joubert N. (2009), « Processus de détection et évaluation de la fraude sociale », *Revue Économique 2009/5*, Volume 60, p. 1235-1256

Ho Ta Khanh J. (2003), « Les redressements pour fraude et travail au noir dans la base 2000 des comptes nationaux : sources, méthodes et résultats » rapport de mission, Insee Département des Comptes Nationaux

Magnien F. (2006), « L'évaluation statistique de la fraude fiscale et sociale en France - État des lieux, essai de mesure, perspectives », *Rapport pour le Conseil des prélèvements obligatoires - Octobre 2006*

Marbot C. (2008), « En France, qui recourt aux services à domicile ? », *France portait social, Édition 2008*

Muller P. (1988), « Les comptes d'entreprises par sous-secteurs d'activité » *Note de méthode de la base 1980, août 1988*

OCDE (2003), *Manuel sur la mesure de l'économie non observée*

Schneider F. avec la collaboration de Enste D. (2002), « Sous la protection de l'ombre - la croissance de l'économie souterraine », *Dossiers économiques 30 - Fonds monétaire international*

Smith Ph. (1994), « Évaluation de la dimension de l'économie souterraine : le point de vue de Statistique Canada », *Document de recherche - N°13-604-MIB au catalogue - N°28*

United Nations Publication (2008), « *Non-observed Economy in National Accounts* », *Survey of Country Practices*

Wagner C. (1995), « La prise en compte de l'économie au noir : l'exemple de la méthode italienne », *Économie et statistique - N°285-286*, 1995

Willard J.-C (1989), « L'économie souterraine dans les comptes nationaux », *Économie et statistique - N°226*, novembre 1989

Glossaire

ACOSS : agence centrale des organismes de sécurité sociale

BIC : bénéfices industriels et commerciaux

BNC : bénéfices non commerciaux

BRN : bénéfice réel normal

CCF : consommation de capital fixe

CGI : code général des impôts

DADS : déclaration annuelle de données sociales

Dgfiip : direction générale des finances publiques

DGI : direction générale des impôts

EBE : excédent brut d'exploitation

EI : entreprise individuelle

IR : impôt sur le revenu

IS : impôt sur les sociétés

NAF : nomenclature d'activité française

NAF-REV2 : nomenclature d'activité française - révision 2

NES : nomenclature économique de synthèse

PBCAI : profit brut courant avant impôt

PIB : produit intérieur brut

RDB : revenu disponible brut

RNB : revenu national brut

RSI : régime simplifié d'imposition

SIRENE : système d'identification du répertoire des entreprises

SNF : société non financière

SUSE : système unifié de statistiques d'entreprise

TVA : taxe sur la valeur ajoutée

VA : valeur ajoutée

ANNEXE 1

Annexe descriptive : échantillon, contrôles, taux d'imposition

Le fichier Suse-Dgfip

Le fichier-échantillon compte 117 300 entreprises sur le champ de l'étude, ce qui représente un taux de sondage apparent global de 4,5 %. Pour chaque entreprise, et pour chaque exercice susceptible d'avoir été contrôlé¹⁹ (2000 à 2006), les principales variables de ce fichier sont les suivantes :

- d'une part, des variables de SUSE, les principales étant le code APE, tous les codes d'activité que l'on peut en déduire, notamment la NES et la NAF-REV2, la catégorie juridique (société non financière ou entreprise individuelle), le type d'imposition (BIC-BRN, BIC-RSI, BNC) et bien sûr les variables comptables relatives aux exercices 2000 à 2006 (chiffre d'affaires, valeur ajoutée, profit brut courant avant impôt, excédent brut d'exploitation, résultat après impôt, impôt sur les bénéfices, etc.) ;
- d'autre part, les variables du fichier des contrôles fiscaux, les principales étant le type d'imposition selon la nomenclature de la Dgfip, le chiffre d'affaires le plus élevé de la période de contrôle, la date et durée du contrôle, les rectifications pour omissions de recettes en base, les rectifications sur amortissements et provisions en base, les transferts de bénéfices en base, le « résultat établi par le service », les droits rappelés sur l'impôt sur les bénéfices pour les entreprises soumises à l'impôt sur les sociétés (IS) d'une part, à l'impôt sur le revenu (IR) d'autre part, les rectifications en base du déficit des entreprises concernées, les rappels de TVA (brute et totale).
- on dispose ainsi, pour chaque exercice, d'une trentaine de variables dans le fichier-échantillon, auxquelles il faut ajouter celles du fichier de calage ; au total, le fichier complet compte donc plus de 200 variables.

L'échantillon par type d'imposition, secteur d'activité et catégorie juridique

Les entreprises se répartissent entre les trois types d'imposition habituellement utilisés à l'INSEE ; le type d'imposition a été codifié à partir des informations de SUSE, après vérification de la cohérence avec le type d'imposition de la Dgfip, laquelle s'avère plutôt satisfaisante.

La catégorie juridique distingue deux types d'entreprises : les entreprises individuelles (EI) qui se caractérisent par le fait que le propriétaire est une personne physique et la catégorie des sociétés non financières (SNF), regroupant toutes les autres entreprises. La catégorie juridique a été attribuée sur la base de celle de SUSE, et donc du répertoire SIRENE.

Les entreprises de chacun des types d'imposition présentent des caractéristiques assez différentes et ne sont pas également représentées dans l'échantillon (*tableau A1-1*). Celles qui sont assujetties au régime des ***bénéfices réels des bénéfices industriels et commerciaux (BIC-BRN)*** sont en moyenne de plus grande taille ; ce sont essentiellement des sociétés non financières (SNF) plutôt que des entreprises individuelles (EI). Elles sont en outre très bien représentées dans l'échantillon de la Dgfip (le taux de sondage apparent est de 10,7 %) qui couvre presque tous les postes de la NES et, cela, de manière relativement équilibrée. En revanche, les entreprises des bénéfices industriels et commerciaux soumises au ***régime de la déclaration simplifiée (BIC-RSI)*** sont plus petites.

¹⁹ Certains contrôles démarrent avant 2000, mais ils sont peu nombreux. Pour alléger les traitements, on a remplacé les variables de Suse correspondant à ces exercices, antérieurs à 2000, par celles de l'année 2000.

La moitié d'entre elles sont des entreprises individuelles ; elles sont beaucoup moins nombreuses dans l'échantillon (le taux de sondage apparent est de 2,2 %) et un plus grand nombre de postes de la NES n'est pas représenté. Toutefois, l'échantillon demeure assez équilibré en NES. À l'opposé, les entreprises soumises au **régime des bénéficiaires non commerciaux (BNC)** sont presque toutes des entreprises individuelles, essentiellement concentrées dans les secteurs des services et elles sont de petite taille. Ce type d'imposition est nettement sous-représenté dans l'échantillon, avec un taux de sondage apparent de 1,7 % ; toutefois, ce sous-échantillon est lui aussi relativement équilibré entre les secteurs d'activité concernés.

Tableau A1 - 1 : La population et l'échantillon selon le type d'imposition et la catégorie juridique

	Nombre d'entreprises dans la population		Nombre d'entreprises dans l'échantillon		Taux de sondage
	Effectif	%	Effectif	%	%
BIC-BRN					
SNF	728 154	91,8	76 282	95,5	10,47
EI	64 749	8,2	3 623	4,5	5,59
Ensemble	792 902	100,0	79 905	100,0	10,07
BIC-RSI					
SNF	609 745	45,7	18 751	65,1	3,07
EI	724 315	54,3	10 057	34,9	1,39
Ensemble	1 334 060	100,0	28 808	100,0	2,16
BNC					
SNF	26 862	5,3	957	11,2	3,56
EI	477 128	94,7	7 599	88,8	1,59
Ensemble	503 990	100,0	8 556	100,0	1,70
Ensemble					
SNF	1 364 761	51,9	95 990	81,9	7,03
EI	1 266 192	48,1	21 279	18,1	1,68
Ensemble	2 630 953	100,0	117 269 (1)	100,0	4,46

Tableau A1 - 2 : Répartition des contrôles fiscaux selon la nature contrôle

	en %		
	BIC-BRN	BIC-RSI	BNC
Contrôle général	85,3	88,8	92,7
Contrôle ciblé TVA	11,0	12,8	5,2
Autre contrôle ciblé	3,7	1,4	2,1
Ensemble des contrôles	100	100	100

Tableau A1 - 3 : Répartition des contrôles fiscaux selon l'année de démarrage du contrôle

	en % du total		
	BIC-BRN	BIC-RSI	BNC
2000 et avant	24,0	18,7	12,8
2001	30,6	27,9	30,9
2002	26,8	29,3	32,5
2003	13,6	17,5	20,2
2004 et après	5,0	6,6	3,6
Total	100	100	100

Tableau A1 - 4 : Répartition des contrôles fiscaux selon l'année de fin de contrôle

en % du total

	BIC-BRN	BIC-RSI	BNC
Avant 2001	2,8	1,8	2,5
2002	15,1	16,2	22,1
2003	29,7	30,4	32,1
2004	30,5	32,7	32,0
2005	18,3	16,0	9,6
2006	3,6	2,9	1,7
Total	100	100	100

Tableau A1 - 5 : Répartition des contrôles fiscaux selon la durée

en % du total

	BIC-BRN	BIC-RSI	BNC
Moins de 1 an	2,8	3,9	0,9
De 1 an à 1,5 an	4,3	8,4	7,7
De 1,5 an à 2 ans	1,6	3,4	0,8
De 2 ans à 2,5 ans	20,7	26,7	35,1
De 2,5 ans à 3 ans	6,4	5,7	2,1
3 ans et plus	64,2	51,9	53,4
Total	100,0	100,0	100,0

Tableau A1 - 6 : Indicateurs statistiques sur les taux d'imposition sur IR des entreprises contrôlées

en % du résultat

	BIC-BRN	BIC-RSI	BNC
Moyenne arithmétique simple	31,6	19,4	27,9
Médiane	30,8	19,1	26,4
Moyenne harmonique	31,1	19,0	26,4
Écart-type	4,05	2,60	5,84
Quantile à 2 %	23,9	8,9	15,1
Quantile à 95 %	38,7	21,6	36,3

Tableau A1-7 : Indicateurs statistiques sur les taux d'imposition sur IS des entreprises contrôlées

en % du résultat

	BIC-BRN	BIC-RSI	BNC (3)
Moyenne arithmétique simple (1)	28,3	23,7	28
Médiane (1)	28,6	24,0	
Moyenne harmonique (1)	27,4	23,2	
Moyenne pondérée (2)	25,9	22,6	
Écart-type	4,39	3,45	
Quantile à 2 %	15,0	15,0	
Quantile à 95 %	35,0	35,0	

- (1) Statistiques sur les taux d'imposition individuels des entreprises rectifiées
(2) Moyenne des taux d'imposition, pondérée par les chiffres d'affaires, sur le champ des entreprises contrôlées.
(3) Il s'agit du taux finalement retenu pour passer des rectifications en droits aux rectifications en base.

ANNEXE 2

Description des stratifications

Pour tenir compte aux mieux de la diversité des entreprises dans les sous-populations correspondant à chaque type d'imposition, les deux stratifications utilisées ont été nuancées par type d'imposition. Le découpage en tranches de taille se fonde sur le chiffre d'affaires. On a en principe cherché, sans toujours le faire de manière systématique, à réduire la variance du chiffre d'affaires au sein des strates, mais elle demeure forte dans les tranches extrêmes.

La stratification simple croise, au sein de chaque type d'imposition, et pour chaque catégorie juridique, la NES 118 avec la taille définie selon des modalités précisées ci-dessous.

Tableau A2-1 : La stratification simple selon la taille, pour les entreprises assujetties aux BIC-BRN

Chiffre d'affaires	Nombre d'entreprises		Taux de sondage	Moyenne simple des taux de fraude élémentaires	
	Population	Échantillon	en %	Sur CA, en %	Sur résultat, en %
SNF					
Plus de 10 millions €	34 051	13 263	38,95	0,09	0,41
De 3 à 10 millions €	62 997	14 302	22,70	0,16	0,62
De 1 à 3 millions €	135 155	16 875	12,49	0,56	1,63
De 300 000 à 1 million €	199 017	16 468	8,27	1,44	3,46
Moins de 300 000 €	222 382	14 617	6,57	5,18	11,07
Ensemble des SNF(1)	728 154	76 282	10,48		
EI					
Plus de 1 million €	12 613	666	5,28	0,67	2,52
De 500 000 à 1 million €	11 318	666	5,88	2,15	4,27
De 100 000 à 500 000 €	23 758	1 404	5,91	6,48	11,22
Moins de 100 000 €	13 398	860	6,42	12,15	20,26
Ensemble des EI (1)	64 749	3 623	5,59		

Tableau A2 - 2 : La stratification simple selon la taille, pour les entreprises assujetties aux BIC-RSI

Chiffre d'affaires	Nombre d'entreprises		Taux de sondage	Moyenne simple des taux de fraude élémentaires	
	Population	Échantillon	en %	Sur CA, en %	Sur résultat, en %
SNF					
Plus de 1,5 million €	1 543	743	48,15	0,80	2,46
De 200 000 à 1,5 million €	159 085	8 432	5,30	5,16	7,95
De 30 000 à 200 000 €	296 091	7 320	2,47	14,42	22,49
Moins de 30 000 €	100 713	1 927	1,91	22,68	54,92
Ensemble des SNF(1)	609 745	18 751	3,08		
EI					
Plus de 800 000 €	1 653	233	14,10	6,34	7,23
De 100 000 à 800 000 €	238 751	5 047	2,11	15,32	20,11
De 50 000 à 100 000 €	179 658	2 233	1,24	28,61	33,24
Moins de 50 000 €	281 767	2499	0,89	47,38	58,4
Ensemble des EI (1)	724 315	10 057	1,39		

(1) y compris les entreprises pour lesquelles la taille est à valeur manquante

Tableau A2-3 : La stratification simple selon la taille, pour les entreprises des BNC

Chiffre d'affaires	Nombre d'entreprises		Taux de sondage	Moyenne simple des taux de fraude élémentaires	
	Population	Échantillon	En %	Sur CA, en %	Sur résultat, en %
SNF					
Plus de 500 000 €	7 195	339	4,71	0,87	3,03
De 150 000 à 500 000 €	9 556	353	3,69	7,19	13,51
Moins de 150 000 €	8 712	255	2,93	24,58	43,32
Ensemble des SNF(1)	26 862	957	3,56		
EI					
Plus de 300 000 €	25 395	1 097	4,32	5,05	10,71
De 50 000 à 300 000 €	268 740	4 398	1,64	18,85	26,73
Moins de 50 000 €	170 603	2 076	1,22	32,99	42,45
Ensemble des EI (1)	477 128	7 599	1,59		

(1) y compris les entreprises pour lesquelles la taille est à valeur manquante.

La deuxième stratification

Dans cette seconde stratification, un code d'activité combiné a été défini en conservant la NES 118 pour les secteurs les plus fraudeurs (une vingtaine) et en retenant le niveau E de la NES pour les autres. On trouvera le détail des modalités de ce code dans les tableaux A3-1 et A3-2 de l'annexe 3. Ainsi, le nombre des modalités du code d'activité est ramené de 118 à une trentaine.

Tableau A2-4 : Stratification selon la taille des entreprises des BIC-BRN

Taille en chiffre d'affaires	Entreprises de la population		Entreprises de l'échantillon		Taux de sondage
	Effectifs	%	Effectifs	%	en %
Non renseigné	785	0,1	785	1,0	
Plus de 20 millions €	17 080	2,2	6 861	8,6	40,2
De 10 à 20 millions €	15 944	2,0	5 321	6,7	33,4
De 3 à 10 millions €	63 563	8,0	14 153	17,7	22,3
De 1 à 3 millions €	146 333	18,5	16 875	21,1	11,5
De 500 000 à 1 million €	128 534	16,2	11 424	14,3	8,9
De 100 000 à 500 000 €	226 397	28,6	15 137	18,9	6,7
Moins de 100 000 €	194 267	24,5	9 349	11,7	4,8
Ensemble	792 903	100,0	79 905	100,0	10,1

Tableau A2-5 : Stratification selon la taille des entreprises des BIC-RSI

Taille en chiffre d'affaires	Entreprises de la population		Entreprises de l'échantillon		Taux de sondage
	Effectifs	%	Effectifs	%	en %
Non renseigné	380	0,0	380	1,3	
Plus de 1 million €	4 637	0,3	1 382	4,8	29,8
De 200 000 à 1 million €	245 985	18,4	8 388	29,1	3,41
De 100 000 à 200 000 €	281 631	21,1	6 613	23,0	2,35
De 30 000 à 100 000 €	452 329	33,9	8 580	29,8	1,90
Moins de 30 000 €	349	26,2	3 465	12,0	0,99
Ensemble	1 334 060	100	28 808	100	2,16

Tableau A2-6 : Stratification selon la taille des entreprises soumises au régime des BNC

Taille en chiffre d'affaires	Entreprises de la population		Entreprises de l'échantillon		Taux de sondage
	Effectifs	%	Effectifs	%	en %
Non renseigné	14 067	2,8	316	3,7	
Plus de 500 000 €	15 634	3,1	857	10,0	5,48
De 200 000 à 500 000 €	52 523	10,4	1 597	18,7	3,04
De 100 000 à 200 000 €	104 064	20,7	1 905	22,3	1,83
De 50 000 à 100 000 €	142 180	28,2	2 019	23,6	1,42
Moins de 50 000 €	175 522	34,8	1 862	21,8	1,06
Ensemble	503 990	100	8 556	100	1,7

Pour simplifier les calculs, au sein d'un type d'imposition, les répartitions selon le code d'activité combiné ont été directement croisées avec des codes de taille détaillés (un par type d'imposition, comportant de 5 à 7 modalités) et avec la catégorie juridique (2 modalités), alors que, dans la première stratification, chaque catégorie juridique avait fait l'objet de ventilations séparées selon des codes de taille spécifiques à chacune d'elles, moins détaillés que les précédents.

Dans cette seconde stratification, les strates en (type d'imposition, activité, taille, catégorie juridique) ont été éclatées selon les modalités du taux de valeur ajoutée et du taux de marge, au nombre de trois. Le taux de valeur ajoutée individuel de chaque entreprise a été défini en divisant sa valeur ajoutée hors taxes (celle de SUSE) par son chiffre d'affaires ; le taux de marge rapporte le PBCAI (*cf. note n°11 du § III-2-1*) au chiffre d'affaires également. On a calculé des ratios moyens sur la base des deux derniers exercices comptables de la période de contrôle. Ainsi définis, ces ratios moyens sont des variables continues ; pour être introduits dans la stratification, ils ont été découpés en tranches, déterminées empiriquement au vu des histogrammes des distributions de ces taux et après analyse des moyennes et des variances.

Ces deux stratifications sont l'une et l'autre extrêmement fines. Pour les BRN, la première compte un peu moins de 1 000 postes, la seconde, environ 270 postes avant éclatement des strates selon les tranches des taux de valeur ajoutée et de marge, mais 2 450 à l'issue de cette opération. C'est la raison pour laquelle, il existe un nombre important de strates vides dans l'échantillon, au niveau le plus fin, ce qui a nécessité de procéder à des redressements sur les taux de fraude (*cf. § IV-3-1*).

Effet des ratios comptables sur l'extrapolation, analysé à travers l'exemple de la santé marchande

Il apparaît, à l'issue du calcul, que l'introduction des ratios comptables dans la stratification a un effet baissier quasi systématique sur les montants de fraude extrapolés. Cet effet est assez diffus, sauf dans quelques secteurs d'activités où il est particulièrement spectaculaire, notamment pour les entreprises du secteur de la santé marchande, lesquelles sont majoritairement des entreprises individuelles assujetties au régime des bénéficiaires non commerciaux (BNC). On se bornera à montrer, sur ce seul exemple, comment interviennent les ratios comptables dans le mécanisme d'extrapolation (*tableau A2-7 et encadré 2*).

Dans l'exemple du tableau A2-7, la relation entre les taux de fraude et les taux de sondage est conforme à l'intuition : les taux de fraude observés diminuent lorsque le taux de valeur ajoutée augmente et les taux de sondage varient aussi en fonction inverse de ces taux de valeur ajoutée. Par ailleurs, on remarque, d'une part que les taux de fraude sur le chiffre d'affaires sont très forts pour les entreprises dont le taux de valeur ajoutée n'est pas renseigné (39,1 % pour la modalité 2 de la taille et 59,5 % pour la modalité 3), d'autre part que ces mêmes entreprises sont nettement surreprésentées dans l'échantillon. Ces entreprises, qui ont une forte propension à frauder, semblent donc « ciblées ».

Lorsque les modalités du taux de valeur ajoutée interviennent dans la stratification, le montant de fraude d'une strate s'obtient en multipliant le taux de fraude de cette strate (en taille, taux de valeur ajoutée) par le chiffre d'affaire correspondant : le montant total de la fraude sur le chiffre d'affaires est alors de 2 247 millions d'euros. Le taux de fraude d'une strate de taille donnée s'exprime comme la moyenne des taux de fraude fins (selon le taux de valeur ajoutée) de cette strate de taille, pondérée par la structure des chiffres d'affaires de la population selon les modalités du taux de valeur ajoutée (colonne (a) du tableau A2-7) : il est égal à 8,0 % pour la modalité 2 de la taille, par exemple. On remarque que, dans cette moyenne, le taux de fraude le plus élevé (39,1 %), est très faiblement pondéré (0,2 %), alors que le taux de fraude le plus faible (7,9%) est affecté d'un poids de 96,4 %.

Si l'on extrapole selon la taille, sans tenir compte des modalités du taux de valeur ajoutée dans la stratification, on utilise les taux de fraude par taille calculés sur l'échantillon, lesquels s'expriment implicitement²⁰ comme des moyennes des taux de fraude fins (taille * taux de valeur ajoutée), pondérées, cette fois, par la structure des chiffres d'affaires de l'échantillon selon les modalités du taux de valeur ajoutée (colonne (b) du tableau A2-7). Pour la modalité 2 de la taille, le taux élevé, de 39,1 %, mentionné ci-dessus est alors affecté d'un poids de 40,2 %, et le taux le plus faible, soit 7,9 %, d'un poids de 51 %. Le taux de fraude de la modalité 2 de la taille est alors égal 20,3 % ; il est nettement plus fort que le taux de 8,0 % mentionné ci-dessus et le montant de fraude qui en dérive est multiplié par 2,5 environ. La forte distorsion entre les structures (a) et (b) du tableau A2-7 est liée à la surreprésentation dans l'échantillon des entreprises les plus fraudeuses, parmi lesquelles se trouvent, entre autres, celles dont les taux de valeur ajoutée ne sont pas renseignés.

La manière dont les ratios comptables interviennent dans le mécanisme d'extrapolation est, au vu de cet exemple, conforme à ce que l'on pouvait entrevoir. Il existe un biais de sélection de l'échantillon qui, s'il n'est pas pris en compte par une stratification *ad hoc*, se traduit, lors de l'extrapolation, par une surpondération de la fraude des entreprises « ciblées », particulièrement fraudeuses, au détriment d'entreprises moins fraudeuses ; d'où une surestimation du montant de fraude extrapolé. Ce biais d'extrapolation est corrigé si la stratification reflète plus fidèlement les écarts entre les probabilités de tirage, et notamment le fait que ces dernières sont plus élevées pour les entreprises les plus fraudeuses (cf. § IV-2 et encadré 2).

Finalement, deux conditions sont nécessaires pour que l'introduction des ratios comptables dans la stratification ait un impact sensible sur les estimations. À la base, les taux de fraude des strates en (taille * ratios) doivent être bien différenciés ; mais il faut en outre qu'il existe, en face, une forte déformation entre la structure des chiffres d'affaires de l'échantillon (pondération implicite de la stratification sans ratios) et celle de la population (pondération de la stratification avec ratios). Si l'un des ratios comptables n'est pas réellement discriminant, ce qui est fréquent, il n'y a pas d'inconvénient à le conserver dans la stratification. Les taux de fraude moyens des différentes modalités de ce ratio sont alors voisins et donc sans conséquences sur les extrapolations, en cas de distorsions des pondérations.

S'agissant de cet exemple, on se trouve typiquement dans une situation où la première stratification ne modélise pas convenablement le ciblage des services fiscaux et il faut lui préférer la seconde. Dans le cas particulier des professions de santé, il convient en outre de signaler que l'échantillon est peu important (moins d'un professionnel sur cent a été contrôlé au cours des trois campagnes de contrôles mobilisées), ce qui accroît la variance des taux observés et fragilise par conséquent les estimations.

²⁰ Sur ce point, on trouvera la formalisation dans l'encadré 2 du texte.

Tableau A2-7 : Évaluation de la fraude sur le chiffre d'affaires selon les deux stratifications, pour le secteur de la santé marchande (entreprises des BNC uniquement)

Taille	Taux de valeur ajoutée	nombre d'entreprises	nombre de contrôles	Chiffre d'affaires SUSE 2006		Stratification selon (taille * taux de valeur ajoutée)				Stratification selon la taille uniquement		
						Taux de fraude sur le chiffre d'affaires (%)		Fraude sur le chiffre d'affaire (millions €)		poids CA échantillon	Taux de fraude agrégés, avec pond. (b)	Fraude sur le chiffre d'affaires (millions €)
						millions €	% (a)	taille * taux de VA	agrégés, pond. population			
1	NR	97	97	61,0	0,6	4,5	2,2	2,8	1,3	26,5	2,9	1,8
1	1	40	8	25,2	0,2	6,2	2,2	1,6	0,5	1,4	2,9	0,7
1	2	458	25	286,5	2,8	5,5	2,2	15,7	6,2	4,6	2,9	8,3
1	3	16 730	265	9 957,7	96,4	2,1	2,2	203,1	215,1	67,6	2,9	289,0
sous-total		17 325	395	10 330,5	100,0		2,2	223,1	223,1	100,0	2,9	299,8
2	NR	749	749	55,5	0,2	39,1	8,0	21,9	4,4	40,2	20,3	11,2
2	1	510	35	48,4	0,2	11,0	8,0	5,3	3,9	1,6	20,3	9,8
2	2	6 393	108	720,2	3,1	8,2	8,0	59,2	57,4	7,1	20,3	146,1
2	3	186 863	734	22 180,8	96,4	7,9	8,0	1 747,0	1 767,8	51,0	20,3	4 500,0
sous-total		194 515	1 626	23 004,9	100,0		8,0	1 833,5	1 833,5	100,0	20,3	4 667,1
3	NR	226	226	3,6	0,2	59,5	8,9	2,2	0,3	46,2	31,5	1,1
3	1	6 229	43	76,0	3,6	10,7	8,9	8,2	6,8	6,7	31,5	24,0
3	2	9 224	43	231,2	10,8	6,2	8,9	14,3	20,6	8,5	31,5	72,9
3	3	53 567	113	1 828,8	85,5	9,1	8,9	166,0	162,9	38,6	31,5	576,6
sous-total		69 246	425	2 139,7	100,0		8,9	190,6	190,6	100,0	31,5	674,6
Ensemble		281 086	2 446	35 475,1				2 247,2	2 247,2			5 641,5

Colonne (1) - Modalités de la taille : de 1 à 3, par ordre de taille décroissant.

Colonne (2) - Modalités du taux de valeur ajoutée : 1, pour les ratios les plus faibles, à 3, pour les plus forts.

Poids CA échantillon (b) : répartition, au sein d'une tranche de taille, des chiffres d'affaires des entreprises de l'échantillon selon les modalités du taux de valeur ajoutée.

Note : Ce tableau présente un calcul volontairement simplifié : il ne porte que sur une partie du champ, ne fait pas intervenir l'activité et certains redressements sont négligés. Les estimations de la fraude de ce secteur ne sont donc pas exactement les mêmes que celles qui interviennent dans les évaluations macroéconomiques du tableau 7 du texte.

On pouvait se demander si certaines professions de santé n'étaient pas plus visées que d'autres. C'est la raison pour laquelle, pour le secteur de la santé marchande, l'activité a été définie dans la post-stratification à un niveau plus fin que celui de la NES 118 (par les codes APE) ; le ciblage de la profession ainsi modélisé s'est révélé ne pas avoir de répercussions spectaculaires sur les estimations.

ANNEXE 3

Estimation économétrique des probabilités d'inclusion

Principe des régressions logistiques

Le problème est d'estimer la probabilité de tirage d'une entreprise au vu de la réalisation d'un échantillonnage. L'échantillon étant inséré dans la population, on peut comparer les caractéristiques des entreprises sélectionnées à celles qui ne l'ont pas été. Techniquement, on dispose, pour cela, d'une variable binaire, égale à 1 pour les entreprises qui ont été effectivement tirées et à 0 pour les autres : cette variable représente en quelque sorte le tirage et c'est elle qu'il s'agit d'expliquer. On a recours pour cela à des régressions, dites « logistiques », dont les variables explicatives sont, d'une part, les différentes modalités de variables qualitatives décrivant les caractéristiques des entreprises (secteur d'activité, taille, catégorie juridique, etc.), d'autre part des variables quantitatives, telles que des grandeurs comptables (taux de valeur ajoutée, de marge, etc.), découpées, on non, en tranches. Les régressions linéaires classiques, fondées sur la méthode des moindres carrés ordinaires ou généralisés, ne permettraient pas de modéliser une variable expliquée dichotomique, telle que la variable de tirage décrite ci-dessus.

Le principe de ces régressions logistiques est de passer par l'intermédiaire d'une transformation consistant à exprimer la probabilité à expliquer comme la fonction de répartition d'une variable latente, fonction comprise par définition entre 0 et 1 et appelée fonction de lien. La fonction de lien la plus utilisée est celle de la loi logistique (modèle logit) ; on utilise aussi celle de la loi normale (modèle probit). Les paramètres sont estimés par la méthode du maximum de vraisemblance. La vraisemblance s'exprime de la manière suivante :

$$L = \prod_{i \in pop} P(x_i)^{y_i} (1 - P(x_i))^{1 - y_i}$$

(pop : lire population)

$P(x_i)$ est la probabilité d'inclusion de l'entreprise i ; y_i est égale à 1 si cette entreprise appartient à l'échantillon et à 0 sinon. Maximiser cette vraisemblance revient à rechercher le jeu de probabilités qui rend le plus « probable » la réalisation du tirage tel qu'on l'observe.

De telles régressions ont été mises en œuvre pour chaque type d'imposition. On aboutit ainsi à la spécification d'un plan de sondage par type d'imposition. Pour chaque entreprise de la population, l'estimateur de la régression n'est autre que la probabilité d'inclusion régissant le tirage de cette entreprise (probabilité de la loi de Bernoulli). Ces probabilités s'expriment en fonction des déterminants mis en évidence par les régressions. Les variables les plus explicatives sont celles qui ont été retenues dans la deuxième stratification et qui ont également été validées par les travaux de statistique descriptive effectués sur les taux de sondage. Parmi les autres variables envisagées (dimension géographique, productivité, appartenance à un groupe) aucune n'a pu être conservée, soit par manque de significativité, soit en raison d'effets de colinéarité ou à cause d'un nombre excessif d'enregistrements non renseignés.

Quelques résultats issus des régressions logistiques

Entreprises assujetties au régime des bénéfiques réels - BIC-BRN

La procédure de régression ne prenant en compte que les enregistrements pour lesquels toutes les variables explicatives sont renseignées, la régression logistique n'a été effectuée que sur 88 % de l'échantillon : 70 653 entreprises sur 79 900. Les modalités de toutes les

variables retenues sont significatives, y compris les différentes modalités du code d'activité agrégé défini dans l'annexe 2, alors que les modalités de la NES ne l'étaient pas. On trouvera dans le tableau A3-1 les résultats détaillés de la régression logistique, pour les entreprises des BRN.

Parmi les indicateurs calculés par la procédure logistique, les plus parlants sont les « rapports de cote », fréquemment désignés par le terme anglais « d'odds ratios ». Ces rapports sont définis pour chaque modalité d'une variable qualitative disjonctée de la régression, autre que la modalité de référence. Si les probabilités de tirage étaient faibles, les odds ratios donneraient une bonne approximation du rapport entre la probabilité de tirage des entreprises présentant cette modalité et la probabilité de tirage des entreprises présentant la modalité de référence.

Dans le cas des BRN, la probabilité de tirage des entreprises dont le chiffre d'affaires est supérieur à 20 millions d'euros est élevée (40,2 %) ; celle des entreprises dont le chiffre d'affaires est inférieur à 100 000 euros est de 4,8 %, soit nettement plus basse. Le rapport entre ces deux probabilités d'inclusion est de 8,4. Il est très voisin de l'odds ratio, égal à 8,5. Bien que les conditions de l'approximation de la probabilité de tirage relative par l'odds ratio ne soient pas réunies, on est frappé par l'étonnante cohérence entre ces deux chiffres. Pour les entreprises dont le chiffre d'affaires est compris entre 10 et 20 millions d'euros, le rapport des taux de sondage est de 7,0, contre 6,5 pour le rapport de cote. L'examen des odds-ratios confirme l'intérêt des variables retenues. Les effets des taux de valeur ajoutée et de marge sur les probabilités de tirage sont moins forts que celui de la taille, entre autres parce que ces variables ne sont vraiment pertinentes que dans certains secteurs d'activité.

Entreprises soumises au régime d'imposition simplifié - BIC-RSI

Pour les entreprises assujetties au régime de la déclaration simplifiée, la proportion d'entreprises de l'échantillon qui a pu être traitée par la régression est plus faible : 71 %, soit 20 529 sur 28 800. La significativité du code agrégé d'activité est moins nette que dans le cas des BRN (tableau A3-2). Toutefois, elle reste bonne, en général, pour les modalités correspondant aux secteurs d'activité les plus fraudeurs. La variable la plus significative est de très loin la tranche de chiffre d'affaires. Les ratios comptables conservent un intérêt certain.

Entreprises soumises au régime des bénéficiaires non commerciaux BNC

Un peu plus de la moitié des entreprises de l'échantillon seulement a été prise en compte dans la régression en raison du nombre important des valeurs non renseignées des variables explicatives. Les modalités du code d'activité agrégé sont toutes significatives. Celles des autres variables également. Les odds ratio figurent dans le tableau A3-3.

Tableau A3 - 1 : Résultats de la régression logistique sur la variable de tirage des entreprises assujetties au régime des BRN

Variables explicatives	Coefficient	Écart-type	Odds ratio
Constante	-2,6951	0,0497	
Secteur d'activité combiné - niveau E de la NES			
Industries agricoles et agroalimentaires	-0,7086	0,0600	0,492
Biens de consommation	-0,2800	0,0501	0,756
Automobile	-0,1788 ***	0,0830	0,836
Biens d'équipement	-0,2686	0,0500	0,764
Biens intermédiaires	-0,3224	0,0483	0,724
Énergie	-0,4805	0,0950	0,618
Services aux entreprises	-0,2582	0,0587	0,772
Services aux particuliers	-0,1912 **	0,0621	0,826
Éducation, santé, action sociale	-0,5876	0,0851	0,556
Administration	-0,8987	0,0874	0,407
Secteur d'activité combiné - niveau G de la NES			
Transport (NES- niveau F)	-0,4533	0,0561	0,636
IAA - Divers	-0,7545	0,0624	0,470
Bâtiment	-0,4794	0,0476	0,619
Travaux publics	-0,5482	0,0557	0,578
Commerce et réparation	-0,8476	0,0499	0,428
Commerce de gros, intermédiaire	-0,5308	0,0473	0,588
Commerce de détail et réparation	-1,2886	0,0482	0,276
Transport routier de voyageurs	-0,7205	0,0772	0,487
Transport routier de marchandises	-0,4999	0,0525	0,607
Auxiliaires financiers et d'assurance	-0,4599	0,0629	0,631
Promotion, gestion immobilière	-0,2723	0,0486	0,762
Location immobilière	-0,4527	0,0513	0,636
Activités informatiques	-0,1553 **	0,0519	0,856
Services professionnels	-0,6170	0,0507	0,540
Administration d'entreprise	-0,1433 **	0,0513	0,867
Publicité et études de marché	-0,2099	0,0560	0,811
Architecture, ingénierie, contrôle	-0,3923	0,0514	0,675
Location sans opérateur	-0,3468	0,0616	0,707
Sécurité, nettoyage, services divers	-0,2370	0,0509	0,789
Hôtel, café, restaurant	-0,6489	0,0496	0,523
Activités récréatives, culture, sport	-0,2228	0,0577	0,800
Services personnels	-1,0327	0,0663	0,356
Éducation marchande	-0,5721	0,0681	0,564
Santé marchande	-1,0489	-	Référence
Taux de valeur ajoutée en tranches			
Inférieur à 50	0,4817	0,0214	1,619
De 0 à 50	0,3941	0,0112	1,483
Supérieur à 50	0	-	Référence
Taux de valeur ajoutée en tranches			
Inférieur à 0	0,3058	0,0200	1,358
De 0 à 30	0,3035	0,0187	1,355
Supérieur à 30	0	-	Référence
Taille de l'entreprise			
Plus de 20 millions €	2,1453	0,0220	8,544
De 10 à 20 millions d'euros	1,8672	0,0228	6,470
De 3 à 10 millions d'euros	1,2876	0,0177	3,624
De 1 à 3 millions d'euros	0,5917	0,0168	1,807
De 500 000 à 1 million €	0,2959 **	0,0176	1,344
De 100 000 à 500 000 €	-0,0379	0,0163	0,963
Moins de 100 000 €	0	-	Référence
Catégorie juridique			
Entreprises individuelles	-0,2667	0,0200	0,766
Sociétés non financières	0	-	Référence

Tableau A3-2 : Résultats de la régression logistique sur la variable de tirage des entreprises assujetties au régime des RSI

Variables explicatives	Coefficient	Écart-type	Odds ratio
Constante	-5,2597	0,1487	-
Secteur d'activité combiné - niveau E de la NES			
Industries agricoles et agroalimentaires	-0,2584***	0,1822	0,772
Biens de consommation	0,3025***	0,1560	1,353
Automobile	0,3398***	0,3184	1,405
Biens d'équipement	0,2054***	0,1592	1,228
Biens intermédiaires	0,3397***	0,1548	1,404
Énergie	0,9489*	0,2717	2,583
Services aux entreprises	0,4510**	0,1585	1,570
Services aux particuliers	0,5285**	0,1884	1,696
Éducation, santé, action sociale	0,3572**	0,1679	1,429
Administration	-0,2577***	0,5256	0,773
Transport (NES- niveau F)	0,5631**	0,1808	1,756
Secteur d'activité combiné - niveau G de la NES			
IAA - Divers	-0,5709*	0,1589	0,565
Industrie du meuble	-0,2179***	0,1786	0,804
Matériel médico-chirurgical	-0,3831**	0,2178	0,682
Bâtiment	0,4097**	0,1472	1,506
Travaux publics	0,4140**	0,1600	1,513
Commerce et réparation	0,1351***	0,1507	1,145
Commerce de gros, intermédiaire	0,5693	0,1487	1,767
Commerce de détail et réparation	-0,3186**	0,1479	0,727
Transport routier de voyageurs	-0,5030*	0,1700	0,605
Transport routier de marchandises	0,5386*	0,1548	1,714
Auxiliaires financiers et d'assurance	0,5049*	0,1737	1,657
Promotion, gestion immobilière	1,0718	0,1496	2,921
Location immobilière	0,3501**	0,1501	1,419
Activités informatiques	0,2642***	0,1579	1,302
Services professionnels	0,5954	0,1527	1,814
Administration d'entreprise	1,0628	0,1553	2,895
Architecture, ingénierie, contrôle	0,4276*	0,1563	1,533
Location sans opérateur	0,5159*	0,1633	1,675
Sécurité, nettoyage, services divers	0,7858	0,1510	2,194
Hôtel, café, restaurant	0,4998*	0,1474	1,648
Activités récréatives, culture, sport	0,5295*	0,1568	1,698
Services personnels	-0,4338**	0,1549	0,648
Santé marchande	0,2792***	0,1807	1,322
Administration	0	-	référence
Taux de valeur ajoutée en tranches			
Inférieur à 50	0,1976	0,0382	1,218
De 0 à 50	0,2648	0,0179	1,303
Supérieur à 50	0	-	référence
Taux de marge en tranches			
Inférieur à 0	0,3446	0,0244	1,411
De 0 à 15	0,4889	0,0190	1,631
Supérieur à 15	0	-	référence
Taille de l'entreprise			
Plus de 1 million d'€	3,1437	0,0451	23,191
De 200 000 à 1 million €	0,9980	0,0292	2,713
De 100 000 à 200 000 €	0,6353	0,0292	1,888
De 30 000 à 100 000 €	0,4443	0,0273	1,559
Moins de 30 000 €	0	-	référence
Catégorie juridique			
Entreprises individuelles	-0,3436	0,0174	0,709
Sociétés non financières	0	-	référence

Note : Les coefficients sont les coefficients bruts de la régression logistique dont la variable expliquée est la variable de tirage des entreprises (égale à 1 ou 0). Tous ces coefficients sont significatifs au seuil de 5 %, sauf ceux qui sont marqués d'un ou plusieurs signes (*) ; plus nombreux sont ces signes moins la variable est significative. Dans la colonne de droite figurent les odds ratios. Si les probabilités de tirage sont faibles, ces odds ratios fournissent une approximation de la probabilité de tirage relative des entreprises présentant la modalité par rapport à la celle des entreprises de la modalité de référence.

Tableau A3 - 6 : BNC - Les principaux rapports de cote (odds ratios)

Taille en chiffres d'affaires	Odds ratios	Taux de valeur ajoutée	Odds ratios
plus de 500 000 €	12,0	inférieur à 25	1,6
de 200 à 500 000 €	8,0	De 25 à 50	1,8
de 100 à 200 000 €	4,9	supérieur à 50	Référence
de 50 à 100 000 €	2,8	Taux de marge	
moins de 50 000 €	Référence	inférieur à 20	2,4
		de 20 à 40	1,6
		supérieur à 40	Référence

Extrapolation de la fraude à l'aide des probabilités estimées par les régressions

En toute logique, les probabilités estimées par les régressions devraient pouvoir être utilisées pour extrapoler à l'ensemble de la population les montants de fraude observés sur l'échantillon ; en effet, elles simulent un plan de sondage, à l'instar des taux de sondage de la stratification.

Pour procéder à cette seconde extrapolation, il suffisait de remplacer les coefficients d'extrapolation de la deuxième stratification (inverses des taux de sondage, soit N/n) par les inverses des probabilités d'inclusion de la régression. En pratique, cela revenait à multiplier les montants de fraude individuels obtenus par la post-stratification (deuxième stratification) par un coefficient correctif égal au rapport entre le taux de sondage de la strate à laquelle appartient l'entreprise considérée et l'estimateur de sa probabilité d'inclusion, issu de la régression. Ainsi conduit, le calcul comporte deux étapes bien distinctes :

- d'abord, l'estimation de la propension à frauder des entreprises de chaque strate de l'échantillon, c'est-à-dire des taux de fraude de ces strates ;
- ensuite, l'extrapolation, à l'ensemble de la population, de la fraude dérivée de ces taux, à l'aide des inverses des probabilités des régressions.

On remarquera que ces deux étapes existent implicitement dans l'extrapolation simple par post-stratification, dont le principe est présenté au § IV-3 du texte.

L'extrapolation à partir des probabilités des régressions posait malgré tout un problème, lié à l'existence de probabilités d'inclusion manquantes, au demeurant assez nombreuses sur le champ des RSI et des BNC. En extrapolant ces probabilités sans redresser les valeurs manquantes, on peut regretter de se priver d'une partie de l'information disponible ; toutefois, cette manière de procéder respecte l'optimisation du plan de sondage par la méthode du maximum de vraisemblance. On a néanmoins testé une variante d'extrapolation comportant un redressement des probabilités non renseignées. Ce redressement est le plus simple qui soit : les probabilités manquantes sont remplacées par les taux de sondage de la deuxième stratification, laquelle croise les variables explicatives des régressions.

Les résultats des trois méthodes d'extrapolation - par post stratification, à l'aide des probabilités des régressions, avec et sans redressements - (*tableaux A3-4 et A3-5 ci-dessous*) sont très voisins, y compris pour les types d'imposition où les probabilités manquantes sont les plus nombreuses (BNC et dans une moindre mesure RSI). Il est naturel que les deux plans de sondage, celui de la régression logistique et celui de la deuxième stratification, conduisent à des estimations très proches de la fraude puisque les variables explicatives retenues dans les régressions sont précisément celles de la deuxième stratification. Les extrapolations obtenues par la post-stratification étant les plus lisibles, elles ont été finalement privilégiées.

Tableau A3 - 4 : Extrapolation de la fraude sur le chiffre d'affaires*Millions d'euros*

Extrapolation		Deuxième stratification (1)	Probabilité des régressions sans redressement	Probas. des régressions avec redressement par les taux de sondage de la stratification
SNF	BRN	3215,5	3155,2	3301,1
	RSI	3903,6	4054,7	4189,9
	BNC	177,6	216,6	246,6
EI	BRN	452,4	414,2	510,0
	RSI	8 663,4	8 754,5	8 914,3
	BNC	3 543,7	3 405,1	3 485,5
Ensemble SNF		7 296,7	7 426,5	7 737,7
Ensemble EI		12 659,5	12 573,8	12 909,8
Ensemble		19 956,2	20 000,3	20 647,5

Tableau A3 - 5 : Extrapolation de la fraude sur le résultat*Millions d'euros*

Extrapolation		Deuxième stratification (1)	Probabilités des régressions sans redressement	Probas. des régressions avec redressement par les taux de sondage de la stratification
SNF	BRN	11 551,6	11 370,4	11 885,4
	RSI	8 415,8	8 648,3	8 951,4
	BNC	516,9	571,3	644,2
EI	BRN	1 033,8	802,7	1 004,7
	RSI	13 223,7	13 124,6	13 379,7
	BNC	6 133,0	5 878,8	5 982,8
Ensemble SNF		20 484,3	20 590,0	21 481,0
Ensemble EI		20 390,5	19 806,1	20 367,2
Ensemble		40 874,8	40 396,1	40 848,2

(1) : Les hypothèses de traitement des données fiscales de cette simulation diffèrent à la marge de celles la variante centrale de ce rapport, ce qui explique de légers écarts entre les estimations de la variante centrale et celles de la colonne de gauche de ce tableau. En revanche, les trois variantes des tableaux ci-dessus ne diffèrent que par la technique d'extrapolation.

G 9001	J. FAYOLLE et M. FLEURBAEY Accumulation, profitabilité et endettement des entreprises		Macro-economic import functions with imperfect competition - An application to the E.C. Trade		françaises : une évaluation empirique des théories de la structure optimale du capital	G 9412	J. BOURDIEU - B. CŒURÉ - B. COLIN-SEDILLOT Investissement, incertitude et irréversibilité Quelques développements récents de la théorie de l'investissement
G 9002	H. ROUSSE Détection et effets de la multicollinéarité dans les modèles linéaires ordinaires - Un prolongement de la réflexion de BELSLEY, KUH et WELSCH	G 9203	I. STAPIC Les échanges internationaux de services de la France dans le cadre des négociations multilatérales du GATT Juin 1992 (1ère version) Novembre 1992 (version finale)	G 9312	L. BLOCH - B. CŒURÉ Q de Tobin marginal et transmission des chocs financiers	G 9413	B. DORMONT - M. PAUCHET L'évaluation de l'élasticité emploi-salaire dépend-elle des structures de qualification ?
G 9003	P. RALLE et J. TOUJAS-BERNATE Indexation des salaires : la rupture de 1983	G 9204	P. SEVESTRE L'économétrie sur données individuelles-temporelles. Une note introductive	G 9313	Equipes Amadeus (INSEE), Banque de France, Métric (DP) Présentation des propriétés des principaux modèles macroéconomiques du Service Public	G 9414	I. KABLA Le Choix de breveter une invention
G 9004	D. GUELLEC et P. RALLE Compétitivité, croissance et innovation de produit	G 9205	H. ERKEL-ROUSSE Le commerce extérieur et l'environnement international dans le modèle AMADEUS (réestimation 1992)	G 9314	B. CREPON - E. DUGUET Research & Development, competition and innovation	G 9501	J. BOURDIEU - B. CŒURÉ - B. SEDILLOT Irreversible Investment and Uncertainty : When is there a Value of Waiting ?
G 9005	P. RALLE et J. TOUJAS-BERNATE Les conséquences de la désindexation. Analyse dans une maquette prix-salaires	G 9206	N. GREENAN et D. GUELLEC Coordination within the firm and endogenous growth	G 9315	B. DORMONT Quelle est l'influence du coût du travail sur l'emploi ?	G 9502	L. BLOCH - B. CŒURÉ Imperfections du marché du crédit, investissement des entreprises et cycle économique
G 9101	Equipe AMADEUS Le modèle AMADEUS - Première partie - Présentation générale	G 9207	A. MAGNIER et J. TOUJAS-BERNATE Technology and trade : empirical evidences for the major five industrialized countries	G 9316	D. BLANCHET - C. BROUSSE Deux études sur l'âge de la retraite	G 9503	D. GOUX - E. MAURIN Les transformations de la demande de travail par qualification en France Une étude sur la période 1970-1993
G 9102	J.L. BRILLET Le modèle AMADEUS - Deuxième partie - Propriétés variantielles	G 9208	B. CREPON, E. DUGUET, D. ENCAOUA et P. MOHNEN Cooperative, non cooperative R & D and optimal patent life	G 9317	D. BLANCHET Répartition du travail dans une population hétérogène : deux notes	G 9504	N. GREENAN Technologie, changement organisationnel, qualifications et emploi : une étude empirique sur l'industrie manufacturière
G 9103	D. GUELLEC et P. RALLE Endogenous growth and product innovation	G 9209	B. CREPON et E. DUGUET Research and development, competition and innovation : an application of pseudo maximum likelihood methods to Poisson models with heterogeneity	G 9318	D. EYSSARTIER - N. PONTY AMADEUS - an annual macro-economic model for the medium and long term	G 9505	D. GOUX - E. MAURIN Persistence des hiérarchies sectorielles de salaires: un réexamen sur données françaises
G 9104	H. ROUSSE Le modèle AMADEUS - Troisième partie - Le commerce extérieur et l'environnement international	G 9301	J. TOUJAS-BERNATE Commerce international et concurrence imparfaite : développements récents et implications pour la politique commerciale	G 9319	G. CETTE - Ph. CUNÉO - D. EYSSARTIER - J. GAUTIÉ Les effets sur l'emploi d'un abaissement du coût du travail des jeunes	G 9505	D. GOUX - E. MAURIN Bis Persistence of inter-industry wages differentials: a reexamination on matched worker-firm panel data
G 9105	H. ROUSSE Effets de demande et d'offre dans les résultats du commerce extérieur manufacturé de la France au cours des deux dernières décennies	G 9302	Ch. CASES Durées de chômage et comportements d'offre de travail : une revue de la littérature	G 9401	D. BLANCHET Les structures par âge importent-elles ?	G 9506	S. JACOBZONE Les liens entre RMI et chômage, une mise en perspective <i>NON PARU - article sorti dans Economie et Prévision n°122 (1996) - pages 95 à 113</i>
G 9106	B. CREPON Innovation, taille et concentration : causalités et dynamiques	G 9303	H. ERKEL-ROUSSE Union économique et monétaire : le débat économique	G 9402	J. GAUTIÉ Le chômage des jeunes en France : problème de formation ou phénomène de file d'attente ? Quelques éléments du débat	G 9507	G. CETTE - S. MAHFOUZ Le partage primaire du revenu Constat descriptif sur longue période
G 9107	B. AMABLE et D. GUELLEC Un panorama des théories de la croissance endogène	G 9304	N. GREENAN - D. GUELLEC / G. BROUSSAUDIER - L. MIOTTI Innovation organisationnelle, dynamisme technologique et performances des entreprises	G 9403	P. QUIRION Les déchets en France : éléments statistiques et économiques	G 9601	Banque de France - CEPREMAP - Direction de la Prévision - Erasme - INSEE - OFCE Structures et propriétés de cinq modèles macro-économiques français
G 9108	M. GLAUDE et M. MOUTARDIER Une évaluation du coût direct de l'enfant de 1979 à 1989	G 9305	P. JAILLARD Le traité de Maastricht : présentation juridique et historique	G 9404	D. LADIRAY - M. GRUN-REHOMME Lissage par moyennes mobiles - Le problème des extrémités de série	G 9602	Rapport d'activité de la DESE de l'année 1995
G 9109	P. RALLE et alii France - Allemagne : performances économiques comparées	G 9306	J.L. BRILLET Micro-DMS : présentation et propriétés	G 9405	V. MAILLARD Théorie et pratique de la correction des effets de jours ouvrables	G 9603	J. BOURDIEU - A. DRAZNIKES L'octroi de crédit aux PME : une analyse à partir d'informations bancaires
G 9110	J.L. BRILLET Micro-DMS NON PARU	G 9307	J.L. BRILLET Micro-DMS - variantes : les tableaux	G 9406	F. ROSENWALD La décision d'investir	G 9604	A. TOPIOL-BENSAÏD Les implantations japonaises en France
G 9111	A. MAGNIER Effets accélérateur et multiplicateur en France depuis 1970 : quelques résultats empiriques	G 9308	S. JACOBZONE Les grands réseaux publics français dans une perspective européenne	G 9407	S. JACOBZONE Les apports de l'économie industrielle pour définir la stratégie économique de l'hôpital public	G 9605	P. GENIER - S. JACOBZONE Comportements de prévention, consommation d'alcool et tabagie : peut-on parler d'une gestion globale du capital santé ? <i>Une modélisation microéconométrique empirique</i>
G 9112	B. CREPON et G. DUREAU Investissement en recherche-développement : analyse de causalités dans un modèle d'accélérateur généralisé	G 9309	L. BLOCH - B. CŒURE Profitabilité de l'investissement productif et transmission des chocs financiers	G 9408	L. BLOCH, J. BOURDIEU, B. COLIN-SEDILLOT, G. LONGUEVILLE Du défaut de paiement au dépôt de bilan : les banquiers face aux PME en difficulté	G 9606	C. DOZ - F. LENGART Factor analysis and unobserved component models: an application to the study of French business surveys
G 9113	J.L. BRILLET, H. ERKEL-ROUSSE, J. TOUJAS-BERNATE "France-Allemagne Couplées" - Deux économies vues par une maquette macro-économétrique	G 9310	J. BOURDIEU - B. COLIN-SEDILLOT Les théories sur la structure optimal du capital : quelques points de repère	G 9409	D. EYSSARTIER, P. MAIRE Impacts macro-économiques de mesures d'aide au logement - quelques éléments d'évaluation	G 9607	N. GREENAN - D. GUELLEC La théorie coopérative de la firme
G 9201	W.J. ADAMS, B. CREPON, D. ENCAOUA Choix technologiques et stratégies de dissuasion d'entrée	G 9311	J. BOURDIEU - B. COLIN-SEDILLOT Les décisions de financement des entreprises	G 9410	F. ROSENWALD Suivi conjoncturel de l'investissement		
G 9202	J. OLIVEIRA-MARTINS, J. TOUJAS-BERNATE			G 9411	C. DEFEUILLEY - Ph. QUIRION Les déchets d'emballages ménagers : une analyse économique des politiques française et allemande		

G 9608	N. GREENAN - D. GUELLEC Technological innovation and employment reallocation
G 9609	Ph. COUR - F. RUPPRECHT L'intégration asymétrique au sein du continent américain : un essai de modélisation
G 9610	S. DUCHENE - G. FORGEOT - A. JACQUOT Analyse des évolutions récentes de la productivité apparente du travail
G 9611	X. BONNET - S. MAHFOUZ The influence of different specifications of wages-prices spirals on the measure of the NAIRU : the case of France
G 9612	PH. COUR - E. DUBOIS, S. MAHFOUZ, J. PISANI-FERRY The cost of fiscal retrenchment revisited: how strong is the evidence ?
G 9613	A. JACQUOT Les flexions des taux d'activité sont-elles seulement conjoncturelles ?
G 9614	ZHANG Yingxiang - SONG Xueqing Lexique macroéconomique Français-Chinois
G 9701	J.L. SCHNEIDER La taxe professionnelle : éléments de cadrage économique
G 9702	J.L. SCHNEIDER Transition et stabilité politique d'un système redistributif
G 9703	D. GOUX - E. MAURIN Train or Pay: Does it Reduce Inequalities to Encourage Firms to Train their Workers?
G 9704	P. GENIER Deux contributions sur dépendance et équité
G 9705	E. DUGUET - N. IUNG R & D Investment, Patent Life and Patent Value An Econometric Analysis at the Firm Level
G 9706	M. HOUEBINE - A. TOPIOL-BENSAÏD Les entreprises internationales en France : une analyse à partir de données individuelles
G 9707	M. HOUEBINE Polarisation des activités et spécialisation des départements en France
G 9708	E. DUGUET - N. GREENAN Le biais technologique : une analyse sur données individuelles
G 9709	J.L. BRILLET Analyzing a small French ECM Model
G 9710	J.L. BRILLET Formalizing the transition process : scenarios for capital accumulation
G 9711	G. FORGEOT - J. GAUTIÉ Insertion professionnelle des jeunes et processus de déclassement
G 9712	E. DUBOIS High Real Interest Rates: the Consequence of a Saving Investment Disequilibrium or of an insufficient Credibility of Monetary Authorities?
G 9713	Bilan des activités de la Direction des Etudes et Synthèses Economiques - 1996
G 9714	F. LEQUILLER Does the French Consumer Price Index Overstate Inflation?
G 9715	X. BONNET Peut-on mettre en évidence les rigidités à la baisse des salaires nominaux ? Une étude sur quelques grands pays de l'OCDE
G 9716	N. IUNG - F. RUPPRECHT Productivité de la recherche et rendements d'échelle dans le secteur pharmaceutique français
G 9717	E. DUGUET - I. KABLA Appropriation strategy and the motivations to use the patent system in France - An econometric analysis at the firm level
G 9718	L.P. PELÉ - P. RALLE Âge de la retraite : les aspects incitatifs du régime général
G 9719	ZHANG Yingxiang - SONG Xueqing Lexique macroéconomique français-chinois, chinois-français
G 9720	M. HOUEBINE - J.L. SCHNEIDER Mesurer l'influence de la fiscalité sur la localisation des entreprises
G 9721	A. MOURougANE Crédibilité, indépendance et politique monétaire Une revue de la littérature
G 9722	P. AUGERAUD - L. BRIOT Les données comptables d'entreprises Le système intermédiaire d'entreprises Passage des données individuelles aux données sectorielles
G 9723	P. AUGERAUD - J.E. CHAPRON Using Business Accounts for Compiling National Accounts: the French Experience
G 9724	P. AUGERAUD Les comptes d'entreprise par activités - Le passage aux comptes - De la comptabilité d'entreprise à la comptabilité nationale - A <i>paraître</i>
G 9801	H. MICHAUDON - C. PRIGENT Présentation du modèle AMADEUS
G 9802	J. ACCARDO Une étude de comptabilité générationnelle pour la France en 1996
G 9803	X. BONNET - S. DUCHÊNE Apports et limites de la modélisation « Real Business Cycles »
G 9804	C. BARLET - C. DUGUET - D. ENCAOUA - J. PRADEL The Commercial Success of Innovations An econometric analysis at the firm level in French manufacturing
G 9805	P. CAHUC - Ch. GIANELLA - D. GOUX - A. ZILBERBERG Equalizing Wage Differences and Bargaining Power - Evidence from a Panel of French Firms
G 9806	J. ACCARDO - M. JLASSI La productivité globale des facteurs entre 1975 et 1996
G 9807	Bilan des activités de la Direction des Etudes et Synthèses Economiques - 1997

G 9808	A. MOURougANE Can a Conservative Governor Conduct an Accommodative Monetary Policy ?
G 9809	X. BONNET - E. DUBOIS - L. FAUVET Asymétrie des inflations relatives et menus costs : tests sur l'inflation française
G 9810	E. DUGUET - N. IUNG Sales and Advertising with Spillovers at the firm level: Estimation of a Dynamic Structural Model on Panel Data
G 9811	J.P. BERTHIER Congestion urbaine : un modèle de trafic de pointe à courbe débit-vitesse et demande élastique
G 9812	C. PRIGENT La part des salaires dans la valeur ajoutée : une approche macroéconomique
G 9813	A.Th. AERTS L'évolution de la part des salaires dans la valeur ajoutée en France reflète-t-elle les évolutions individuelles sur la période 1979-1994 ?
G 9814	B. SALANIÉ Guide pratique des séries non-stationnaires
G 9901	S. DUCHÊNE - A. JACQUOT Une croissance plus riche en emplois depuis le début de la décennie ? Une analyse en comparaison internationale
G 9902	Ch. COLIN Modélisation des carrières dans Destinie
G 9903	Ch. COLIN Evolution de la dispersion des salaires : un essai de prospective par microsimulation
G 9904	B. CREPON - N. IUNG Innovation, emploi et performances
G 9905	B. CREPON - Ch. GIANELLA Wages inequalities in France 1969-1992 An application of quantile regression techniques
G 9906	C. BONNET - R. MAHIEU Microsimulation techniques applied to inter-generational transfers - Pensions in a dynamic framework: the case of France
G 9907	F. ROSENWALD L'impact des contraintes financières dans la décision d'investissement
G 9908	Bilan des activités de la DESE - 1998
G 9909	J.P. ZOYEM Contrat d'insertion et sortie du RMI Evaluation des effets d'une politique sociale
G 9910	Ch. COLIN - FI. LEGROS - R. MAHIEU Bilans contributifs comparés des régimes de retraite du secteur privé et de la fonction publique
G 9911	G. LAROQUE - B. SALANIÉ Une décomposition du non-emploi en France
G 9912	B. SALANIÉ Une maquette analytique de long terme du marché du travail
G 9912 Bis	Ch. GIANELLA Une estimation de l'élasticité de l'emploi peu qualifié à son coût
G 9913	Division « Redistribution et Politiques Sociales » Le modèle de microsimulation dynamique DESTINIE
G 9914	E. DUGUET Macro-commandes SAS pour l'économétrie des panels et des variables qualitatives
G 9915	R. DUHAUTOIS Evolution des flux d'emplois en France entre 1990 et 1996 : une étude empirique à partir du fichier des bénéficiaires réels normaux (BRN)
G 9916	J.Y. FOURNIER Extraction du cycle des affaires : la méthode de Baxter et King
G 9917	B. CRÉPON - R. DESPLATZ - J. MAIRESSE Estimating price cost margins, scale economies and workers' bargaining power at the firm level
G 9918	Ch. GIANELLA - Ph. LAGARDE Productivity of hours in the aggregate production function: an evaluation on a panel of French firms from the manufacturing sector
G 9919	S. AUDRIC - P. GIVORD - C. PROST Evolution de l'emploi et des coûts par qualification entre 1982 et 1996
G 2000/01	R. MAHIEU Les déterminants des dépenses de santé : une approche macroéconomique
G 2000/02	C. ALLARD-PRIGENT - H. GUILMEAU - A. QUINET The real exchange rate as the relative price of nontradables in terms of tradables: theoretical investigation and empirical study on French data
G 2000/03	J.-Y. FOURNIER L'approximation du filtre passe-bande proposée par Christiano et Fitzgerald
G 2000/04	Bilan des activités de la DESE - 1999
G 2000/05	B. CREPON - F. ROSENWALD Investissement et contraintes de financement : le poids du cycle Une estimation sur données françaises
G 2000/06	A. FLIPO Les comportements matrimoniaux de fait
G 2000/07	R. MAHIEU - B. SÉDILLOT Microsimulations of the retirement decision: a supply side approach
G 2000/08	C. AUDENIS - C. PROST Déficit conjoncturel : une prise en compte des conjonctures passées
G 2000/09	R. MAHIEU - B. SÉDILLOT Equivalent patrimonial de la rente et souscription de retraite complémentaire
G 2000/10	R. DUHAUTOIS Ralentissement de l'investissement : petites ou grandes entreprises ? industrie ou tertiaire ?
G 2000/11	G. LAROQUE - B. SALANIÉ Temps partiel féminin et incitations financières à l'emploi
G2000/12	Ch. GIANELLA Local unemployment and wages
G2000/13	B. CREPON - Th. HECKEL - Informatisation en France : une évaluation à partir de données individuelles

	- Computerization in France: an evaluation based on individual company data	G2002/01	F. MAGNIEN - J.-L. TAVERNIER - D. THESMAR Les statistiques internationales de PIB par habitant en standard de pouvoir d'achat : une analyse des résultats	G2002/16	F. MAUREL - S. GREGOIR Les indices de compétitivité des pays : interprétation et limites	G2004/06	M. DUÉE L'impact du chômage des parents sur le devenir scolaire des enfants
G2001/01	F. LEQUILLER - La nouvelle économie et la mesure de la croissance du PIB - The new economy and the measurement of GDP growth	G2002/02	Bilan des activités de la DESE - 2001	G2003/01	N. RIEDINGER - E. HAUVY Le coût de dépollution atmosphérique pour les entreprises françaises : Une estimation à partir de données individuelles	G2004/07	P. AUBERT - E. CAROLI - M. ROGER New Technologies, Workplace Organisation and the Age Structure of the Workforce: Firm-Level Evidence
G2001/02	S. AUDRIC La reprise de la croissance de l'emploi profite-t-elle aussi aux non-diplômés ?	G2002/03	B. SÉDILLOT - E. WALRAET La cessation d'activité au sein des couples : y a-t-il interdépendance des choix ?	G2003/02	P. BISCOURP et F. KRAMARZ Création d'emplois, destruction d'emplois et internationalisation des entreprises industrielles françaises : une analyse sur la période 1986-1992	G2004/08	E. DUGUET - C. LELARGE Les brevets accroissent-ils les incitations privées à innover ? Un examen microéconométrique
G2001/03	I. BRAUN-LEMAIRE Evolution et répartition du surplus de productivité	G2002/04	G. BRILHAULT - Rétropolation des séries de FBCF et calcul du capital fixe en SEC-95 dans les comptes nationaux français - Retropolation of the investment series (GFCF) and estimation of fixed capital stocks on the ESA-95 basis for the French balance sheets	G2003/03	Bilan des activités de la DESE - 2002	G2004/09	S. RASPILLER - P. SILLARD Affiliating versus Subcontracting: the Case of Multinationals
G2001/04	A. BEAUDU - Th. HECKEL Le canal du crédit fonctionne-t-il en Europe ? Une étude de l'hétérogénéité des comportements d'investissement à partir de données de bilan agrégées	G2002/05	P. BISCOURP - B. CRÉPON - T. HECKEL - N. RIEDINGER How do firms respond to cheaper computers? Microeconomic evidence for France based on a production function approach	G2003/04	P.-O. BEFFY - J. DERUYON - N. FOURCADE - S. GREGOIR - N. LAÏB - B. MONFORT Évolutions démographiques et croissance : une projection macro-économique à l'horizon 2020	G2004/10	J. BOISSINOT - C. L'ANGEVIN - B. MONFORT Public Debt Sustainability: Some Results on the French Case
G2001/05	C. AUDENIS - P. BISCOURP - N. FOURCADE - O. LOISEL Testing the augmented Solow growth model : An empirical reassessment using panel data	G2002/06	C. AUDENIS - J. DERUYON - N. FOURCADE L'impact des nouvelles technologies de l'information et de la communication sur l'économie française - un bouclage macro-économique	G2003/05	P. AUBERT La situation des salariés de plus de cinquante ans dans le secteur privé	G2004/11	S. ANANIAN - P. AUBERT Travailleurs âgés, nouvelles technologies et changements organisationnels : un réexamen à partir de l'enquête « REPONSE »
G2001/06	R. MAHIEU - B. SÉDILLOT Départ à la retraite, irréversibilité et incertitude	G2002/07	J. BARDAJI - B. SÉDILLOT - E. WALRAET Évaluation de trois réformes du Régime Général d'assurance vieillesse à l'aide du modèle de microsimulation DESTINIE	G2003/06	P. AUBERT - B. CRÉPON Age, salaire et productivité La productivité des salariés décline-t-elle en fin de carrière ?	G2004/12	X. BONNET - H. PONCET Structures de revenus et propensions différentes à consommer - Vers une équation de consommation des ménages plus robuste en prévision pour la France
G2001/07	Bilan des activités de la DESE - 2000	G2002/08	J.-P. BERTHIER Réflexions sur les différentes notions de volume dans les comptes nationaux : comptes aux prix d'une année fixe ou aux prix de l'année précédente, séries chaînées	G2003/07	H. BARON - P.O. BEFFY - N. FOURCADE - R. MAHIEU Le ralentissement de la productivité du travail au cours des années 1990	G2004/13	C. PICART Évaluer la rentabilité des sociétés non financières
G2001/08	J. Ph. GAUDEMET Les dispositifs d'acquisition à titre facultatif d'annuités viagères de retraite	G2002/09	F. HILD Les soldes d'opinion résumant-ils au mieux les réponses des entreprises aux enquêtes de conjoncture ?	G2003/08	P.-O. BEFFY - B. MONFORT Patrimoine des ménages, dynamique d'allocation et comportement de consommation	G2004/14	J. BARDAJI - B. SÉDILLOT - E. WALRAET Les retraites du secteur public : projections à l'horizon 2040 à l'aide du modèle de microsimulation DESTINIE
G2001/09	B. CRÉPON - Ch. GIANELLA Fiscalité, coût d'usage du capital et demande de facteurs : une analyse sur données individuelles	G2002/10	I. ROBERT-BOBÉE Les comportements démographiques dans le modèle de microsimulation Destinie - Une comparaison des estimations issues des enquêtes Jeunes et Carrières 1997 et Histoire Familiale 1999	G2003/09	P. BISCOURP - N. FOURCADE Peut-on mettre en évidence l'existence de rigidités à la baisse des salaires à partir de données individuelles ? Le cas de la France à la fin des années 90	G2005/01	S. BUFFETEAU - P. GODEFROY Conditions de départ en retraite selon l'âge de fin d'études : analyse prospective pour les générations 1945 à 1974
G2001/10	B. CRÉPON - R. DESPLATZ Evaluation des effets des dispositifs d'allègements de charges sociales sur les bas salaires	G2002/11	J.-P. ZOYEM La dynamique des bas revenus : une analyse des entrées-sorties de pauvreté	G2003/10	M. LECLAIR - P. PETIT Présence syndicale dans les firmes : quel impact sur les inégalités salariales entre les hommes et les femmes ?	G2005/02	C. AFSA - S. BUFFETEAU L'évolution de l'activité féminine en France : une approche par pseudo-panel
G2001/11	J.-Y. FOURNIER Comparaison des salaires des secteurs public et privé	G2002/12	F. HILD Prévisions d'inflation pour la France	G2003/11	P.-O. BEFFY - X. BONNET - M. DARRACQ-PARIES - B. MONFORT MZE: a small macro-model for the euro area	G2005/03	P. AUBERT - P. SILLARD Délocalisations et réductions d'effectifs dans l'industrie française
G2001/12	J.-P. BERTHIER - C. JAULENT R. CONVENEVOLE - S. PISANI Une méthodologie de comparaison entre consommations intermédiaires de source fiscale et de comptabilité nationale	G2002/13	M. LECLAIR Réduction du temps de travail et tensions sur les facteurs de production	G2004/01	P. AUBERT - M. LECLAIR La compétitivité exprimée dans les enquêtes trimestrielles sur la situation et les perspectives dans l'industrie	G2005/04	M. LECLAIR - S. ROUX Mesure et utilisation des emplois instables dans les entreprises
G2001/13	P. BISCOURP - Ch. GIANELLA Substitution and complementarity between capital, skilled and less skilled workers: an analysis at the firm level in the French manufacturing industry	G2002/14	E. WALRAET - A. VINCENT - Analyse de la redistribution intragénérationnelle dans le système de retraite des salariés du privé - Une approche par microsimulation - Intragenerational distributional analysis in the french private sector pension scheme - A microsimulation approach	G2004/02	M. DUÉE - C. REBILLARD La dépendance des personnes âgées : une projection à long terme	G2005/05	C. L'ANGEVIN - S. SERRAVALLE Performances à l'exportation de la France et de l'Allemagne - Une analyse par secteur et destination géographique
G2001/14	I. ROBERT-BOBÉE Modelling demographic behaviours in the French microsimulation model Destinie: An analysis of future change in completed fertility	G2002/15	P. CHONE - D. LE BLANC - I. ROBERT-BOBÉE Offre de travail féminine et garde des jeunes enfants	G2004/03	S. RASPILLER - N. RIEDINGER Régulation environnementale et choix de localisation des groupes français	G2005/06	Bilan des activités de la Direction des Études et Synthèses Économiques - 2004
G2001/15	J.-P. ZOYEM Diagnostic sur la pauvreté et calendrier de revenus : le cas du "Panel européen des ménages"			G2004/04	A. NABOULET - S. RASPILLER Les déterminants de la décision d'investir : une approche par les perceptions subjectives des firmes	G2005/07	S. RASPILLER La concurrence fiscale : principaux enseignements de l'analyse économique
G2001/16	J.-Y. FOURNIER - P. GIVORD La réduction des taux d'activité aux âges extrêmes, une spécificité française ?			G2004/05	N. RAGACHE La déclaration des enfants par les couples non mariés est-elle fiscalement optimale ?	G2005/08	C. L'ANGEVIN - N. LAÏB Éducation et croissance en France et dans un panel de 21 pays de l'OCDE
G2001/17	C. AUDENIS - P. BISCOURP - N. RIEDINGER Existe-t-il une asymétrie dans la transmission du prix du brut aux prix des carburants ?					G2005/09	N. FERRARI Prévoir l'investissement des entreprises Un indicateur des révisions dans l'enquête de conjoncture sur les investissements dans l'industrie.

G2005/10	P.-O. BEFFY - C. L'ANGEVIN Chômage et boucle prix-salaires : apport d'un modèle « qualifiés/peu qualifiés »
G2005/11	B. HEITZ A two-states Markov-switching model of inflation in France and the USA: credible target VS inflation spiral
G2005/12	O. BIAU - H. ERKEL-ROUSSE - N. FERRARI Réponses individuelles aux enquêtes de conjoncture et prévision macroéconomiques : Exemple de la prévision de la production manufacturière
G2005/13	P. AUBERT - D. BLANCHET - D. BLAU The labour market after age 50: some elements of a Franco-American comparison
G2005/14	D. BLANCHET - T. DEBRAND - P. DOURGNON - P. POLLET L'enquête SHARE : présentation et premiers résultats de l'édition française
G2005/15	M. DUÉE La modélisation des comportements démogra- phiques dans le modèle de microsimulation DESTINIE
G2005/16	H. RAOUI - S. ROUX Étude de simulation sur la participation versée aux salariés par les entreprises
G2006/01	C. BONNET - S. BUFFETEAU - P. GODEFROY Disparités de retraite de droit direct entre hommes et femmes : quelles évolutions ?
G2006/02	C. PICART Les gazelles en France
G2006/03	P. AUBERT - B. CRÉPON - P. ZAMORA Le rendement apparent de la formation continue dans les entreprises : effets sur la productivité et les salaires
G2006/04	J.-F. OUVRARD - R. RATHELOT Demographic change and unemployment: what do macroeconomic models predict?
G2006/05	D. BLANCHET - J.-F. OUVRARD Indicateurs d'engagements implicites des systèmes de retraite : chiffrages, propriétés analytiques et réactions à des chocs démographiques types
G2006/06	G. BIAU - O. BIAU - L. ROUVIERE Nonparametric Forecasting of the Manufacturing Output Growth with Firm-level Survey Data
G2006/07	C. AFSA - P. GIVORD Le rôle des conditions de travail dans les absences pour maladie
G2006/08	P. SILLARD - C. L'ANGEVIN - S. SERRAVALLE Performances comparées à l'exportation de la France et de ses principaux partenaires Une analyse structurelle sur 12 ans
G2006/09	X. BOUTIN - S. QUANTIN Une méthodologie d'évaluation comptable du coût du capital des entreprises françaises : 1984- 2002
G2006/10	C. AFSA L'estimation d'un coût implicite de la pénibilité du travail chez les travailleurs âgés
G2006/11	C. LELARGE Les entreprises (industrielles) françaises sont- elles à la frontière technologique ?
G2006/12	O. BIAU - N. FERRARI Théorie de l'opinion Faut-il pondérer les réponses individuelles ?
G2006/13	A. KOUBI - S. ROUX Une réinterprétation de la relation entre productivité et inégalités salariales dans les entreprises
G2006/14	R. RATHELOT - P. SILLARD The impact of local taxes on plants location decision
G2006/15	L. GONZALEZ - C. PICART Diversification, recentrage et poids des activités de support dans les groupes (1993-2000)
G2007/01	D. SRAER Allègements de cotisations patronales et dynamique salariale
G2007/02	V. ALBOUY - L. LEQUIEN Les rendements non monétaires de l'éducation : le cas de la santé
G2007/03	D. BLANCHET - T. DEBRAND Aspiration à la retraite, santé et satisfaction au travail : une comparaison européenne
G2007/04	M. BARLET - L. CRUSSON Quel impact des variations du prix du pétrole sur la croissance française ?
G2007/05	C. PICART Flux d'emploi et de main-d'œuvre en France : un réexamen
G2007/06	V. ALBOUY - C. TAVAN Massification et démocratisation de l'enseignement supérieur en France
G2007/07	T. LE BARBANCHON The Changing response to oil price shocks in France : a DSGE type approach
G2007/08	T. CHANEY - D. SRAER - D. THESMAR Collateral Value and Corporate Investment Evidence from the French Real Estate Market
G2007/09	J. BOISSINOT Consumption over the Life Cycle: Facts for France
G2007/10	C. AFSA Interpréter les variables de satisfaction : l'exemple de la durée du travail
G2007/11	R. RATHELOT - P. SILLARD Zones Franches Urbaines : quels effets sur l'emploi salarié et les créations d'établissements ?
G2007/12	V. ALBOUY - B. CRÉPON Aléa moral en santé : une évaluation dans le cadre du modèle causal de Rubin
G2008/01	C. PICART Les PME françaises : rentables mais peu dynamiques
G2008/02	P. BISCOURP - X. BOUTIN - T. VERGÉ The Effects of Retail Regulations on Prices Evidence from the Loi Galland
G2008/03	Y. BARBESOL - A. BRIANT Economies d'agglomération et productivité des

G2009/09	G. LALANNE - E. POULIQUEN - O. SIMON Prix du pétrole et croissance potentielle à long terme
G2009/10	D. BLANCHET - J. LE CACHEUX - V. MARCUS Adjusted net savings and other approaches to sustainability: some theoretical background
G2009/11	V. BELLAMY - G. CONSALES - M. FESSEAU - S. LE LAIDIER - É. RAYNAUD Une décomposition du compte des ménages de la comptabilité nationale par catégorie de ménage en 2003
G2009/12	J. BARDAJI - F. TALLET Detecting Economic Regimes in France: a Qualitative Markov-Switching Indicator Using Mixed Frequency Data
G2009/13	R. AEBERHARDT - D. FOUGÈRE - R. RATHELOT Discrimination à l'embauche : comment exploiter les procédures de <i>testing</i> ?
G2009/14	Y. BARBESOL - P. GIVORD - S. QUANTIN Partage de la valeur ajoutée, approche par données microéconomiques
G2009/15	I. BUONO - G. LALANNE The Effect of the Uruguay round on the Intensive and Extensive Margins of Trade
G2010/01	C. MINODIER Avantages comparés des séries des premières valeurs publiées et des séries des valeurs révisées - Un exercice de prévision en temps réel de la croissance trimestrielle du PIB en France
G2010/02	V. ALBOUY - L. DAVEZIES - T. DEBRAND Health Expenditure Models: a Comparison of Five Specifications using Panel Data
G2010/03	C. KLEIN - O. SIMON Le modèle MÉSANGE réestimé en base 2000 Tome 1 – Version avec volumes à prix constants
G2010/04	M.-É. CLERC - É. COUDIN L'IPC, miroir de l'évolution du coût de la vie en France ? Ce qu'apporte l'analyse des courbes d'Engel
G2010/05	N. CECI-RENAUD - P.-A. CHEVALIER Les seuils de 10, 20 et 50 salariés : impact sur la taille des entreprises françaises
G2010/06	R. AEBERHARDT - J. POUGET National Origin Differences in Wages and Hierarchical Positions - Evidence on French Full- Time Male Workers from a matched Employer- Employee Dataset
G2010/07	S. BLASCO - P. GIVORD Les trajectoires professionnelles en début de vie active : quel impact des contrats temporaires ?
G2010/08	P. GIVORD Méthodes économétriques pour l'évaluation de politiques publiques
G2010/09	P.-Y. CABANNES - V. LAPÈGUE - E. POULIQUEN - M. BEFFY - M. GAINI Quelle croissance de moyen terme après la crise ?
G2010/10	I. BUONO - G. LALANNE La réaction des entreprises françaises à la baisse des tarifs douaniers étrangers
G2008/04	D. BLANCHET - F. LE GALLO Les projections démographiques : principaux mécanismes et retour sur l'expérience française
G2008/05	D. BLANCHET - F. TOUTLEMONDE Évolutions démographiques et déformation du cycle de vie active : quelles relations ?
G2008/06	M. BARLET - D. BLANCHET - L. CRUSSON Internationalisation et flux d'emplois : que dit une approche comptable ?
G2008/07	C. LELARGE - D. SRAER - D. THESMAR Entrepreneurship and Credit Constraints - Evidence from a French Loan Guarantee Program
G2008/08	X. BOUTIN - L. JANIN Are Prices Really Affected by Mergers?
G2008/09	M. BARLET - A. BRIANT - L. CRUSSON Concentration géographique dans l'industrie manufacturière et dans les services en France : une approche par un indicateur en continu
G2008/10	M. BEFFY - É. COUDIN - R. RATHELOT Who is confronted to insecure labor market histories? Some evidence based on the French labor market transition
G2008/11	M. ROGER - E. WALRAET Social Security and Well-Being of the Elderly: the Case of France
G2008/12	C. AFSA Analyser les composantes du bien-être et de son évolution Une approche empirique sur données individuelles
G2008/13	M. BARLET - D. BLANCHET - T. LE BARBANCHON Microsimuler le marché du travail : un prototype
G2009/01	P.-A. PIONNIER Le partage de la valeur ajoutée en France, 1949-2007
G2009/02	Laurent CLAVEL - Christelle MINODIER A Monthly Indicator of the French Business Climate
G2009/03	H. ERKEL-ROUSSE - C. MINODIER Do Business Tendency Surveys in Industry and Services Help in Forecasting GDP Growth? A Real-Time Analysis on French Data
G2009/04	P. GIVORD - L. WILNER Les contrats temporaires : trappe ou marchepied vers l'emploi stable ?
G2009/05	G. LALANNE - P.-A. PIONNIER - O. SIMON Le partage des fruits de la croissance de 1950 à 2008 : une approche par les comptes de surplus
G2009/06	L. DAVEZIES - X. D'HAULTFOEUILLE Faut-il pondérer?... Ou l'éternelle question de l'économètre confronté à des données d'enquête
G2009/07	S. QUANTIN - S. RASPILLER - S. SERRAVALLE Commerce intragroupe, fiscalité et prix de transferts : une analyse sur données françaises
G2009/08	M. CLERC - V. MARCUS Élasticités-prix des consommations énergétiques des ménages

G2010/11	R. RATHELOT - P. SILLARD L'apport des méthodes à noyaux pour mesurer la concentration géographique - Application à la concentration des immigrés en France de 1968 à 1999	prises sur la base des contrôles fiscaux et son insertion dans les comptes nationaux
G2010/12	M. BARATON - M. BEFFY - D. FOUGÈRE Une évaluation de l'effet de la réforme de 2003 sur les départs en retraite - Le cas des enseignants du second degré public	
G2010/13	D. BLANCHET - S. BUFFETEAU - E. CRENNER S. LE MINEZ Le modèle de microsimulation Destinie 2 : principales caractéristiques et premiers résultats	
G2010/14	D. BLANCHET - E. CRENNER Le bloc retraites du modèle Destinie 2 : guide de l'utilisateur	
G2010/15	M. BARLET - L. CRUSSON - S. DUPUCH - F. PUECH Des services échangés aux services échangeables : une application sur données françaises	
G2010/16	M. BEFFY - T. KAMIONKA Public-private wage gaps: is civil-servant human capital sector-specific?	
G2010/17	P.-Y. CABANNES - H. ERKEL-ROUSSE - G. LALANNE - O. MONSO - E. POULIQUEN Le modèle Mésange réestimé en base 2000 Tome 2 - Version avec volumes à prix chaînés	
G2010/18	R. AEBERHARDT - L. DAVEZIES Conditional Logit with one Binary Covariate: Link between the Static and Dynamic Cases	
G2011/01	T. LE BARBANCHON - B. OURLIAC - O. SIMON Les marchés du travail français et américain face aux chocs conjoncturels des années 1986 à 2007 : une modélisation DSGE	
G2011/02	C. MARBOT Une évaluation de la réduction d'impôt pour l'emploi de salariés à domicile	
G2011/03	L. DAVEZIES Modèles à effets fixes, à effets aléatoires, modèles mixtes ou multi-niveaux : propriétés et mises en œuvre des modélisations de l'hétérogénéité dans le cas de données groupées	
G2011/04	M. ROGER - M. WASMER Heterogeneity matters: labour productivity differentiated by age and skills	
G2011/05	J.-C. BRICONGNE - J.-M. FOURNIER V. LAPÈGUE - O. MONSO De la crise financière à la crise économique L'impact des perturbations financières de 2007 et 2008 sur la croissance de sept pays industrialisés	
G2011/06	P. CHARNOZ - É. COUDIN - M. GAINI Wage inequalities in France 1976-2004: a quantile regression analysis	
G2011/07	M. CLERC - M. GAINI - D. BLANCHET Recommendations of the Stiglitz-Sen-Fitoussi Report: A few illustrations	
G2011/08	M. BACHELET - M. BEFFY - D. BLANCHET Projeter l'impact des réformes des retraites sur l'activité des 55 ans et plus : une comparaison de trois modèles	
G2011/09	C. LOUVOT-RUNAVOT L'évaluation de l'activité dissimulée des entre-	