

CHAPITRE 2
LES PRINCIPES GÉNÉRAUX D'ÉLABORATION DES
COMPTES TRIMESTRIELS

2 - LES PRINCIPES GÉNÉRAUX D'ÉLABORATION DES COMPTES TRIMESTRIELS

SEC 95 : « Les méthodes appliquées pour élaborer les comptes trimestriels « peuvent être subdivisées en deux grandes catégories : les procédures directes et les procédures indirectes. Le recours à des procédures directes dépend de la disponibilité à intervalles trimestriels, moyennant naturellement les simplifications qui s'imposent, des mêmes sources de données que celles servant à l'établissement des comptes annuels. Par contre, les procédures indirectes reposent sur la désagrégation des données annuelles à l'aide de techniques mathématiques ou statistiques faisant appel à des indicateurs de référence permettant de procéder à des extrapolations pour l'année courante. »

La construction des comptes trimestriels français repose sur une méthode indirecte² et s'appuie sur deux types de séries, les comptes nationaux annuels et des données conjoncturelles provenant de multiples sources.

Les comptes annuels rassemblent des informations économiques exhaustives, ou du moins les plus complètes possibles, et les synthétisent dans le cadre de la comptabilité nationale. La qualité de ces données a néanmoins comme contrepartie leur délai. Ainsi, la situation économique d'une année N est décrite pour la première fois par la comptabilité annuelle au mois de mai qui suit la fin de l'année N ; c'est le compte provisoire, qui est élaboré en grande partie avec des données conjoncturelles partielles³. Un an plus tard, c'est-à-dire en mai de l'année N+2, le compte de l'année N est révisé et devient « semi-définitif » ; il ne repose que partiellement sur des données exhaustives. La version définitive des comptes de l'année N est finalement publiée en mai de l'année N+3.

À l'inverse, de nombreuses données conjoncturelles, de périodicités mensuelles ou trimestrielles, sont disponibles rapidement et fournissent des informations sur la situation économique récente. Ces données sont publiées par l'Insee ou proviennent de divers organismes assurant la collecte d'informations statistiques (Direction générale des finances publiques, Banque de France...). Elles résultent souvent d'un arbitrage entre rapidité de publication et qualité statistique, ce qui peut impliquer par exemple que leur construction repose sur des échantillons. Dans ce cas, si ce sont des données d'entreprises, elles ne couvrent pas les établissements récemment créés et cela peut être à l'origine de divergences cycliques avec les informations annuelles exhaustives, si la démographie des entreprises est liée au cycle.

Ces données conjoncturelles diffèrent souvent, en niveau et en évolution, des comptes nationaux annuels, notamment pour des raisons de définition et de champ couvert. Ainsi, les indices de chiffre d'affaires utilisés pour mesurer la production de certaines branches ne recoupent pas exactement la notion de production en comptabilité nationale, car ils comptabilisent les seules ventes et excluent la production stockée. Les variations de stocks peuvent donc expliquer en partie les différences d'évolution entre le compte annuel de production et les données annualisées des indices de chiffres d'affaires. Il arrive également que les données conjoncturelles ne correspondent pas exactement aux différents champs du cadre comptable. Ainsi, les effectifs salariés, publiés par l'Insee tous les trimestres, sont répartis en secteurs d'activité alors que l'emploi de la comptabilité nationale est ventilé en branches d'activité⁴. Les évolutions des effectifs salariés ne retranscrivent donc que partiellement celles de l'emploi par branche car elles ne distinguent pas les diverses activités pouvant être effectuées au sein d'une même société.

² Plusieurs autres pays de l'Union européenne privilégient également la méthode indirecte, mais dans une grande majorité de cas les comptes trimestriels sont élaborés selon une méthode directe.

³ La confection du compte provisoire est d'ailleurs réalisée en commun par les comptables trimestriels et annuels.

⁴ Les effectifs d'une branche d'activité correspondent aux travailleurs contribuant à une même activité, quelle que soit leur entreprise, tandis que les effectifs d'un secteur d'activité regroupent les travailleurs des entreprises ayant pour activité principale celle du secteur.

Les comptes trimestriels mettent en cohérence l'information exhaustive annuelle et celle disponible plus rapidement mais partielle des données conjoncturelles. La méthodologie utilisée pour cette synthèse s'appuie principalement sur la méthode de l'étalonnage-calage :

- l'étalonnage transforme les données conjoncturelles, pour les rapprocher de la définition et du champ de la série correspondante des comptes annuels ;
- le calage assure ensuite une cohérence totale entre les séries des comptes trimestriels et les comptes annuels.

2.1. La méthode de base : l'étalonnage-calage

2.1.1. L'étalonnage

Les comptes trimestriels associent à chaque poste de la comptabilité nationale **un indicateur** mensuel ou trimestriel, disponible rapidement et correspondant au mieux au concept et au champ du poste comptable. Cette association est effectuée à un niveau d'agrégation intermédiaire : par exemple la production en automobile, les exportations de produits agricoles, etc. L'idée de base des comptes trimestriels est d'« adapter » les indicateurs aux comptes annuels : en estimant la relation statistique qui relie sur le passé l'indicateur annualisé au compte correspondant ; en postulant que cette relation observée sur données annuelles demeure pertinente lorsqu'elle est appliquée aux données trimestrielles.

La méthode statistique des comptes trimestriels cherche donc à corriger les différences systématiques entre l'information contenue dans les indicateurs conjoncturels et celle des comptes annuels. Ainsi, par exemple, pour estimer la production manufacturière, les indicateurs utilisés sont dans la plupart des branches les indices de production industrielle (IPI). Or certaines branches ne sont pas couvertes en totalité à un niveau fin de la nomenclature (des critères de taille notamment excluent les plus petites entreprises). À titre d'exemple, si on suppose que dans une branche donnée, la production des entreprises non enquêtées (par exemple les très petites entreprises), évolue en moyenne plus vite que celle des autres entreprises, la tendance donnée par l'IPI sera plus faible que celle de la production annuelle de cette branche. Un modèle statistique peut alors permettre de corriger ce biais en redressant l'évolution retracée par l'IPI.

La relation d'étalonnage est une équation linéaire simple entre le compte annuel et l'indicateur annualisé, qui s'écrit pour toutes les années a :

$$C_a = \alpha + \beta \times I_a + u_a$$

où C_a est le compte annuel de l'année a , I_a l'indicateur annualisé, c'est-à-dire la somme annuelle de l'indicateur trimestriel : $I_a = \sum_{t=1}^4 I_{a,t}$, et u_a la cale de la relation d'étalonnage, qui représente les évolutions du compte qui ne sont pas retranscrites par celles de l'indicateur étalonné.

Les coefficients α et β sont estimés sur plusieurs années communes du compte et de l'indicateur, la période devant être suffisamment longue pour que l'estimation soit statistiquement précise. Pour optimiser l'estimation, trois types de modèles sont utilisés, selon les caractères statistiques de la cale⁵. On peut alors définir le résidu \mathcal{E}_a de ces modèles, qui, lui, sera un bruit blanc lorsque le modèle a été bien choisi⁶.

⁵ Sur l'importance de prendre en compte la nature statistique de la cale, voir Bournay J. et Laroque G. [5].

⁶ Conceptuellement, cette méthode repose donc sur le caractère statistique de la relation annuelle entre le compte et l'indicateur. Cette démarche est différente des méthodes de désagrégation temporelle qui sont souvent utilisées, du type Chow G. et Lin A. G. [8]. Ces dernières modélisent d'abord le caractère statistique de la relation trimestrielle recherchée, qui est ensuite estimée à partir des données annuelles.

- L'estimation est effectuée sur le modèle en niveau, lorsque la cale est stationnaire et non auto-corrélée. L'équation suivante est estimée par les moindres carrés ordinaires (MCO) :

$$C_a = \alpha + \beta \times I_a + \varepsilon_a$$

La cale u_a est alors égale au résidu de l'estimation, ε_a .

- L'estimation porte sur le modèle en niveau, lorsque la cale présente de l'auto-corrélation. L'équation est estimée par les moindres carrés quasi-généralisés (MCQG) :

$$C_a = \alpha + \beta \times I_a + u_a$$

avec l'hypothèse que $u_a = \rho \times u_{a-1} + \varepsilon_a$

Ce modèle présente un double intérêt :

- du point de vue de l'estimation des coefficients, s'il s'agit du « vrai modèle », l'estimation est plus efficace par la méthode des MCQG que par la méthode des MCO ;
 - du point de vue de la prévision du compte annuel, la modélisation du comportement de la cale permet de prendre en compte sa persistance pour l'estimation du compte sur l'année en cours.
- Le modèle est estimé en différences, lorsque la cale est non stationnaire. Ce modèle est un cas limite du modèle précédent avec $\rho = 1$. L'équation suivante est estimée par les moindres carrés ordinaires :

$$\Delta C_a = \gamma + \beta \times \Delta I_a + \varepsilon_a$$

La cale u_a est alors telle que : $u_a = u_{a-1} + \varepsilon_a$.

Dans ce cas, si le coefficient γ est significativement différent de zéro, la relation d'étalonnage reliant l'indicateur au compte est en fait :

$$C_a = \beta \times I_a + \gamma \times T_a + u_a$$

où T_a est une tendance linéaire et où la cale u_a n'est pas forcément de moyenne nulle.

Ce modèle présente de nouveau un double intérêt :

- si le compte annuel et l'indicateur sont des séries intégrées d'ordre 1 et non cointégrées, leur différenciation permet de travailler sur des séries stationnaires et de réaliser des tests standards ;
- comme pour le modèle précédent, la prise en compte de la persistance de la cale permet d'améliorer l'estimation du compte sur l'année en cours. Ce cas est très plausible lorsque les comptes et les indicateurs sont construits en évolution et qu'une divergence de l'indicateur une année donnée crée un écart permanent sur les niveaux.

Un exemple de choix de modèle d'étalonnage est développé en *annexe 1*. Il se fait grâce à différents critères statistiques. Le test de portmanteau (*annexe 2*) permet de rejeter le ou les modèles pour lesquels le résidu de l'équation estimée est auto-corrélé. Les écarts-types des résidus permettent de comparer le pouvoir explicatif des trois méthodes sur toute la période, tandis que les écarts-types des résidus récursifs centrent la comparaison sur le pouvoir prédictif.

Cette relation entre le compte et l'indicateur annualisé est supposée être stable pour que, estimée sur le passé, elle reste valable sur le passé très récent, et permette les meilleures prévisions possibles des années où le compte annuel n'est pas connu. En outre, le biais de l'indicateur, corrigé par la relation d'étalonnage, est supposé être invariant au sein de l'année. De cette façon, la même relation peut être utilisée pour « redresser » l'indicateur trimestriel. Les coefficients estimés α et β sont alors appliqués à cet indicateur et la première estimation (avant calage) du compte trimestriel correspond à :

$$C_{a,t} = \frac{\alpha}{4} + \beta \times I_{a,t}$$

où $C_{a,t}$ est le compte du trimestre t de l'année a ⁷.

2.1.2. Le calage

Reproduire uniquement la relation estimée entre l'indicateur et le compte annuel ne permet pas de conserver au niveau trimestriel toute l'information contenue dans le compte annuel. Il subsiste la cale, qui contient une information que l'indicateur, même redressé, ne peut pas donner. Il s'agit alors de conserver cette information en « calant » les comptes trimestriels sur les comptes annuels sur tout le passé disponible. Pour cela, la cale annuelle doit être répartie sur chacun des trimestres.

En théorie, la méthode de « trimestrialisation » des cales devrait prendre en compte le caractère statistique de la cale et être différente pour les cales non auto-corrélées, auto-corrélées et non-stationnaires. Une simple division par quatre des cales annuelles, en théorie plus correcte lorsque la cale n'est pas auto-corrélée, n'est cependant pas la solution retenue. Elle peut en effet engendrer de brusques changements de niveaux sur certains premiers trimestres, si la cale s'avère ces années-là non négligeable par rapport au compte annuel. En pratique, la méthode de « trimestrialisation » utilisée, qui répartit la cale annuelle sur chacun des trimestres, a été choisie de façon pragmatique, et permet de minimiser les variations d'un trimestre à l'autre, de sorte que la contribution de la cale à la volatilité du compte trimestriel soit la plus réduite possible : cette méthode de lissage conduit à ce que le profil trimestriel des cales soit le plus régulier possible ; pour ce faire, les cales trimestrielles résultent d'un programme de minimisation de la somme des carrés des écarts entre les cales successives (cf partie 1.4 sur le lissage).

Le compte trimestriel devient donc :

$$C_{a,t} = \frac{\alpha}{4} + \beta \times I_{a,t} + u_{a,t}$$

où $u_{a,t}$ est la cale trimestrielle du trimestre t de l'année a obtenue par le lissage de la cale annuelle u_a .

⁷ Si le modèle a été estimé en différences et si la constante est significative, la relation d'étalonnage comporte une tendance linéaire, qu'il faut conserver en trimestriel, c'est-à-dire que l'équation devient :

$$C_{a,t} = \beta \times I_{a,t} + \gamma \times T_{a,t}$$

où $T_{a,t}$ est une tendance linéaire telle que $\sum_{t=1}^4 T_{a,t} = T_a$ pour toutes les années a . Par exemple, la tendance $T_{a,t}$ peut être définie par : $T_{a,t} = (4 \times a + t) / 16$ (pour l'année a et le trimestre t compris entre 1 et 4) et la tendance annuelle T_a doit alors être définie par : $T_a = (16 \times a + 10) / 16$.

La cale trimestrielle est déterminée de sorte qu'elle respecte chaque année l'égalité $\sum_{t=1}^4 u_{a,t} = u_a$ ⁸.

Pour chaque année jusqu'à l'année du compte annuel provisoire, la somme des comptes trimestriels sur l'année est alors égale au compte annuel. Pour l'année en cours, qui n'est pas « calée », la cale annuelle est extrapolée. L'extrapolation de la cale utilise le fait que les résidus estimés par les modèles sont des bruits blancs ; ce résidu est considéré comme nul l'année considérée. Par conséquent, le mode d'extrapolation de la cale dépend du modèle retenu lors de l'estimation.

Pour le modèle en niveau, la cale n'est pas auto-corrélée donc :

$$u_{P+1} = 0$$

où P est l'année du compte provisoire.

Si la cale est auto-corrélée, mais stationnaire,

$$u_{P+1} = \hat{\rho}u_P$$

Si la cale est non-stationnaire,

$$u_{P+1} = u_P$$

La procédure de lissage est effectuée après cette extrapolation des cales annuelles, ce qui permet de ne pas créer de rupture entre les trimestres calés et les trimestres correspondant aux comptes provisoires.

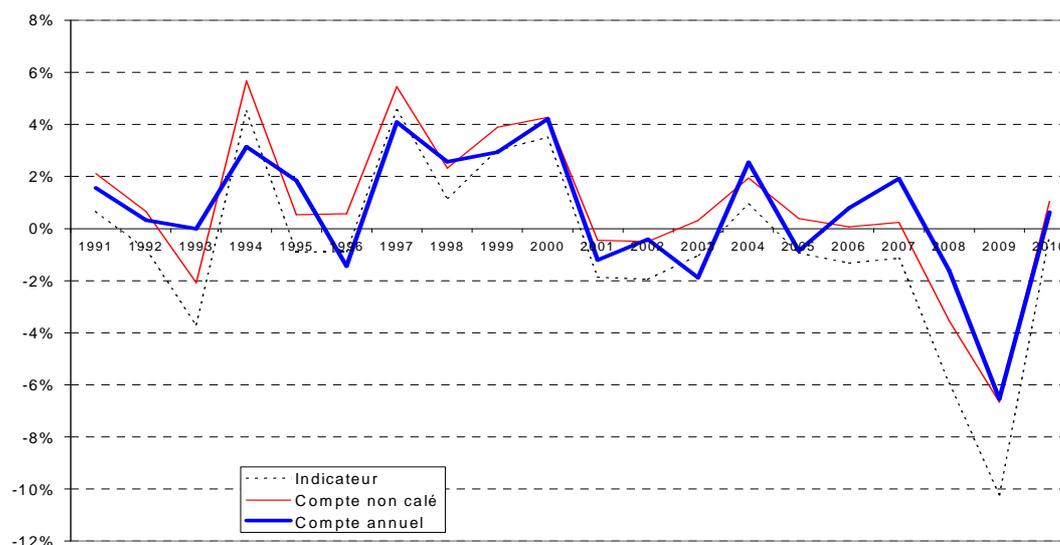
⁸ Si le modèle a été estimé en différences et si la constante est significative, la relation devient :

$$C_{a,t} = \beta \times I_{a,t} + \gamma \times T_{a,t} + u_{a,t}$$

2.1.3. Exemple

Pour illustrer la méthode d'étalonnage sur données annuelles et de calage par lissage trimestriel des résidus, l'exemple de la production en bois-papier est étudié. Le *graphique 1* compare les taux de croissance du compte annuel (C_a), de l'indicateur (I_a) et du compte à résidus nuls. Ce dernier correspond au résultat de l'étalonnage, tel qu'il aurait été obtenu si le compte annuel de l'année donnée n'était pas connu, c'est-à-dire en prolongeant la cale chaque année de façon à annuler le résidu.

Graphique 1 : Effet de l'étalonnage-calage, illustration avec le compte de production de la branche bois-papier (indicateur : IPI, période d'estimation : 1990-2009)

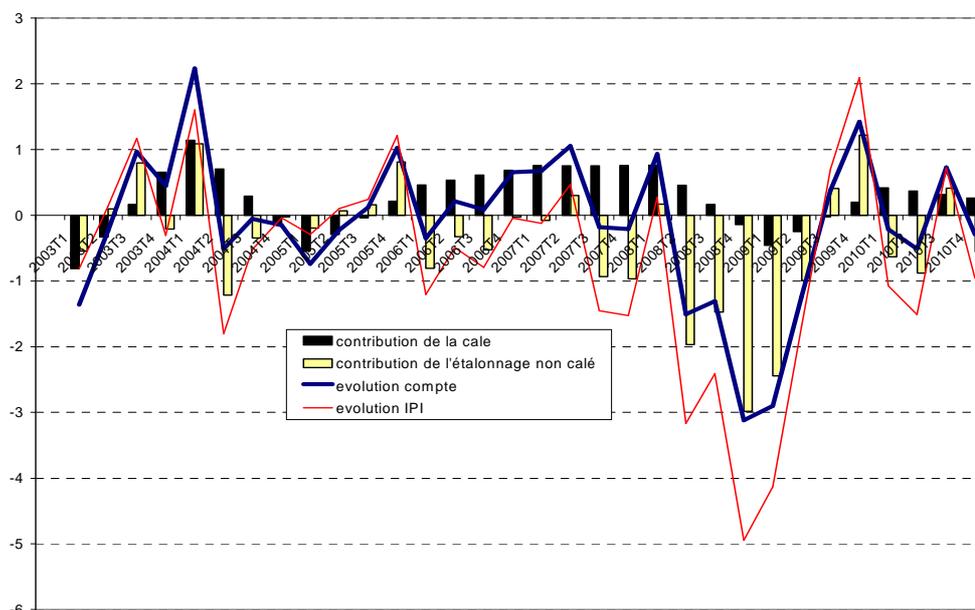


Source : comptes nationaux trimestriels, base 2005

La production de la branche bois-papier est étalonnée sur l'indicateur de production industrielle (IPI). Les évolutions de l'indicateur sont relativement en phase avec celles du compte, mais semblent en moyenne sur la période moins dynamiques. L'étalonnage modifie alors l'indicateur en corrigeant ce biais : il lisse quelque peu les évolutions mais conserve les inflexions de l'indicateur. Ainsi, le compte à résidu nul affiche une allure proche de celle de l'indicateur, tout en étant plus dynamique en moyenne et plus proche du compte annuel. Les résidus sont relativement faibles par rapport à l'amplitude des fluctuations du compte, ce qui tend à montrer que l'IPI est un assez bon indicateur de la production.

Sur une période courte pour plus de lisibilité, le *graphique 2* illustre ensuite les propriétés du calage en montrant pour la même opération (production de bois-papier), le profil trimestriel du compte obtenu après calage, avec les contributions respectives du compte « non calé » (ou « à résidus nuls ») et de la cale. On constate que la contribution de la cale à la volatilité du compte est la plus faible possible, compte tenu de la contrainte de calage des comptes trimestriels sur les comptes annuels.

Graphique 2 : Effet de l'étalonnage-calage en termes trimestriels, illustration avec le compte de production de la branche bois-papier



Source : comptes nationaux trimestriels, base 2005

2.1.4. Le lissage

Le lissage est utilisé dans deux cas principaux. Il sert comme cela a été évoqué dans la partie 1.2 à répartir les cales annuelles issues des étalonnages pour assurer la cohérence entre les comptes trimestriels et les comptes annuels. Il permet également de trimestrialiser certaines séries pour lesquelles aucune donnée conjoncturelle n'est disponible : les comptes annuels sont prolongés par une extrapolation de l'année en cours et les comptes trimestriels découlent alors directement du lissage de la série annuelle.

Cela nécessite donc de prolonger par une extrapolation le compte annuel non disponible sur l'année en cours. C'est le plus souvent sur la base d'expertises *ad hoc* (telles que fournies par les commissions des comptes ou des experts du domaine concerné au sein de l'Insee) ou à défaut d'une prolongation « raisonnée » des tendances passées.

La procédure de lissage d'une série annuelle C_a , utilisée dans les comptes trimestriels, consiste à estimer une série trimestrielle $C_{a,t}$ résultant de la minimisation du carré des écarts d'un trimestre sur l'autre sous la contrainte que, chaque année a , la somme des trimestres soit égale au montant annuel C_a :

$$\min \sum_t (C_{a,t} - C_{a,t-1})^2$$

sous contrainte : $\sum_{t=1}^4 C_{a,t} = C_a$ pour toutes les années a

Le développement de ce programme de minimisation est explicité en *annexe 3*.

2.1.5. Une méthode intermédiaire entre le lissage et l'étalonnage : le lissage en taux

L'étalonnage-calage est utilisé pour évaluer un compte trimestriel dès qu'il existe un indicateur suffisamment proche du concept et du champ du compte. S'il n'en existe pas, le compte trimestriel est souvent obtenu par lissage du compte annuel. Dans certains cas toutefois, il existe un indicateur indirect, qui est lié au compte considéré, notamment par un taux d'imposition ou de cotisation. Ainsi, par exemple, aucune information conjoncturelle n'est disponible concernant les impôts sur les importations. Néanmoins, ces impôts sont liés aux importations par le taux d'imposition. L'étalonnage n'est pas une méthode appropriée dans ce cas car cela supposerait une relation linéaire entre les deux comptes, c'est-à-dire un taux d'imposition apparent constant. Notamment, s'agissant de l'année en cours, on ne peut supposer la constance du taux apparent en cas d'évolution réglementaire du taux.

Une méthode plus appropriée est utilisée, qui consiste à lisser le ratio des comptes, c'est-à-dire le taux d'imposition apparent. Comme pour le lissage simple, cette méthode nécessite d'effectuer une extrapolation l'année courante, ce qui permet notamment de tenir compte des évolutions réglementaires.

Cette méthode de lissage en taux est également adaptée pour évaluer la valeur d'un compte à partir du volume, lorsqu'il n'existe pas d'indicateur de prix : cela revient alors à lisser le prix annuel.

Prenons l'exemple d'un compte annuel T_a (les droits de douanes dans l'exemple précédent) relié économiquement à un autre compte annuel C_a (les importations dans l'exemple précédent). Autrement dit ces via un taux tx_a . S'il n'existe pas d'information trimestrielle disponible sur le taux ou même directement sur le compte T , la procédure de « lissage en taux » revient à estimer le taux trimestriel $tx_{a,t}$ résultant de la minimisation du carré des écarts d'un trimestre sur l'autre, sous la contrainte que, chaque année a , la somme des trimestres de $T_{a,t}$ soit égale à son montant annuel :

$$\min \sum_t (tx_{a,t} - tx_{a,t-1})^2$$

sous contrainte : $\sum_{t=1}^4 tx_{a,t} \times C_{a,t} = tx_a \times C_a = T_a$ pour toutes les années a

Le compte trimestriel est alors défini par :

$$T_{a,t} = tx_{a,t} \times C_{a,t}$$

Le développement de ce programme de minimisation est explicité en *annexe 4*.

2.2. La correction de l'effet des jours ouvrables (CJO) et la correction des variations saisonnières (CVS)

L'objectif des comptes trimestriels est de retracer la dynamique des grandeurs économiques selon un rythme trimestriel. Cependant, les variations d'un trimestre sur l'autre des comptes bruts sont difficilement interprétables du fait de la présence d'effets saisonniers et d'effets de composition trimestrielle des jours dans le calendrier.

Pour faciliter la lecture des évolutions des comptes trimestriels, les séries sont corrigées de l'effet des jours ouvrables (CJO) et de celui des variations saisonnières (CVS). Ces corrections sont appliquées aux indicateurs, puis l'étalonnage-calage permet le calcul du compte « CVS-CJO »⁹.

En pratique, les coefficients de l'équation d'étalonnage et les cales sont estimés à partir des indicateurs bruts et des comptes annuels. Le compte trimestriel brut est obtenu en appliquant la relation d'étalonnage-calage à l'indicateur brut :

$$C_{a,t}^{brut} = \frac{\alpha}{4} + \beta \times I_{a,t}^{brut} + u_{a,t}$$

tandis que le compte trimestriel CVS-CJO se déduit de l'indicateur CVS-CJO en lui appliquant la relation ainsi estimée sur données brutes :

$$C_{a,t}^{cjo-cvs} = \frac{\alpha}{4} + \beta \times I_{a,t}^{cjo-cvs} + u_{a,t}$$

Une autre méthode consisterait à calculer un compte brut par étalonnage-calage, puis à appliquer les traitements statistiques directement sur le compte brut, et non pas sur les indicateurs. Cette méthode ne permettrait pas d'effectuer certaines désaisonnalisations à des niveaux plus fins que ceux de l'étalonnage. Or des désaisonnalisations à des niveaux fins peut présenter un intérêt lorsque le compte est étalonné sur une somme d'indicateurs correspondant à des produits hétérogènes du point de vue de leur saisonnalité.

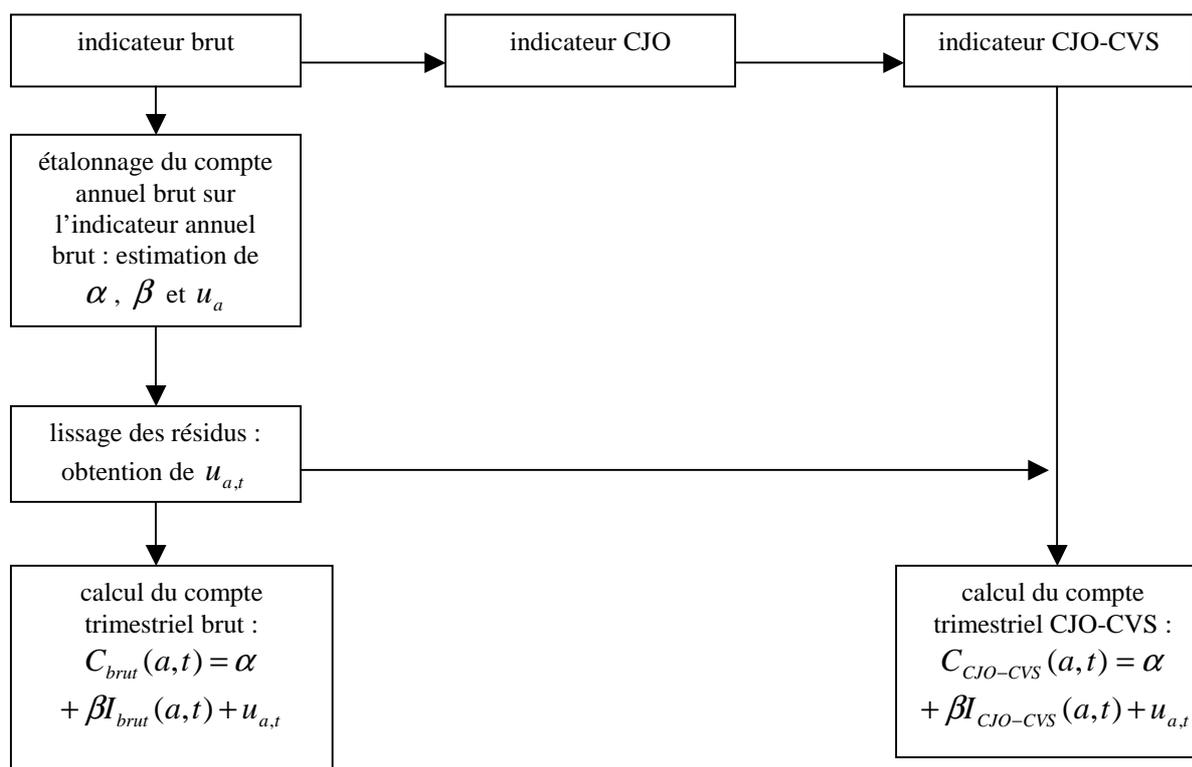
Il faut noter que la méthode utilisée par les comptes trimestriels impose que l'effet de la saisonnalité soit neutre sur l'ensemble de l'année. C'est une contrainte usuelle qui facilite la lecture des séries annuelles et qui ne déforme pas les séries dont la saisonnalité est stable d'une année sur l'autre.

En revanche, la correction de l'effet des jours ouvrables n'est pas neutre sur l'année ; le nombre de jours fériés, par exemple, est différent d'une année sur l'autre, ce qui a un impact sur les évolutions annuelles. La somme sur l'année d'un indicateur CVS-CJO est donc différente de celle de l'indicateur brut : la différence correspond à l'effet annuel de la correction de l'effet des jours ouvrables. Or, pour obtenir les comptes trimestriels bruts et CVS-CJO, les mêmes coefficients d'étalonnage et les mêmes cales sont appliquées. La somme sur l'année du compte trimestriel CVS-CJO est alors différente de celle du compte brut, c'est-à-dire du compte annuel : elle est égale au compte annuel corrigé des effets des jours ouvrables.

La méthode globale suivie pour passer d'un indicateur brut au compte trimestriel CVS-CJO est retracée par le schéma 1.

⁹ Le terme « CVS-CJO » est le plus utilisé, dans les différentes publications et tableaux de diffusion. Il ne correspond toutefois pas à l'ordre chronologique des traitements effectués dans les comptes, puisque les indicateurs sont dans un premier temps corrigés des jours ouvrables et dans un deuxième temps corrigés des variations saisonnières.

Schéma 1 : Méthode suivie pour passer d'un indicateur brut au compte trimestriel CVS-CJO.



Source : comptes nationaux trimestriels

2.2.1. La correction de l'effet des jours ouvrables

L'effet des jours ouvrables sur les séries économiques est potentiellement très important. Ainsi, si on fait l'hypothèse extrême que l'activité de production est entièrement proportionnelle au nombre d'heures travaillées, la production d'un trimestre donné avec un jour férié sera inférieure à celle du trimestre précédent, si celui-ci n'a aucun jour férié, de l'ordre de -1,5 % (soit 1/65, 65 étant le nombre moyen de jours ouvrables d'un trimestre sans jour férié). Il est important de distinguer cette baisse d'un effet propre au cycle économique. Le but de la correction de l'effet des jours ouvrables est alors de construire des séries "à jours ouvrables identiques", pour lesquelles l'analyse des évolutions n'est pas perturbée par les différences de jours ouvrables d'un trimestre sur l'autre.

De fait, les effets des jours ouvrables sont loin d'être proportionnels. Par exemple, les processus de production sont adaptables et peuvent permettre de rattraper en partie la perte due à un jour férié. Par ailleurs les effets peuvent être différents selon les jours de la semaine. D'une part, la productivité des salariés n'est pas la même tous les jours de la semaine. D'autre part, de nombreux postes de consommation sont plus importants le samedi que les autres jours de la semaine.

Tout ceci conduit à adopter des méthodes statistiques qui évaluent sur chacune des séries l'effet des jours ouvrables et permettent de différencier les jours de la semaine. Ces méthodes sont en général plus efficaces sur des séries mensuelles que trimestrielles. En effet, d'une année sur l'autre, le nombre de jours fériés comme le nombre des différents jours de la semaine que compte un trimestre donné ne varie pas beaucoup¹⁰, ce qui rend difficile l'estimation de leur effet. En revanche, la variabilité du nombre de jours fériés est plus forte en termes

¹⁰ Par exemple, le 2^e trimestre, qui compte exactement 13 semaines, comporte toujours 13 lundis, 13 mardis, ..., 13 dimanches.

mensuels. Ainsi, dans les comptes trimestriels, la correction des jours ouvrables est effectuée autant que possible sur les indicateurs mensuels.

La correction de l'effet des jours ouvrables part de l'hypothèse selon laquelle l'évolution d'un indicateur peut être décomposée en deux composantes orthogonales (non corrélées) : une composante liée uniquement aux effets de jours ouvrables et une composante corrigée des jours ouvrables (CJO), qui contient en particulier la composante saisonnière de la série. La composante qui correspond à l'effet des jours ouvrables doit alors être estimée indépendamment de la saisonnalité ; l'effet moyen de jours fériés attachés à un mois particulier (Noël...) est intégré à la composante saisonnière.

La méthode de base utilisée pour les comptes trimestriels est simple : il s'agit d'une régression de la variable mensuelle brute sur des variables représentant respectivement le nombre de lundis ouvrables (c'est-à-dire non fériés)..., samedis ouvrables, et de dimanches, fériés ou non, de chaque mois. Pour ne pas intégrer d'effet propre à la saisonnalité, ces variables sont désaisonnalisées, en ne conservant que l'écart à la moyenne sur chaque mois de ces nombres de jours. La somme des jours ouvrables, des dimanches, et des jours fériés ne tombant pas un dimanche, est égal au nombre de jours du mois, toujours fixe, sauf pour le mois de février, qui possède la particularité d'avoir un nombre de jours fluctuants (28 ou 29). Ainsi, il convient d'ajouter à la régression une indicatrice d'année bissextile (cf. *annexe 5* pour plus de détails).

L'indicateur brut est donc décomposé selon :

$$I_t = \alpha_1 N_{lun}^{ouv} + \alpha_2 N_{mar}^{ouv} + \dots + \alpha_6 N_{sam}^{ouv} + \alpha_7 N_{dim} + \beta I_{biss} + I_t^{cjo}$$

où N_{lun}^{ouv} est le nombre de lundis ouvrables du mois t , désaisonnalisé, I_{biss} l'indicatrice d'année bissextile ; I_t^{cjo} , résidu de cette équation, est l'indicateur corrigé des jours ouvrables.

Dans ce modèle, l'effet des jours fériés n'est pas identifiable. Comme la somme des jours ouvrables, des dimanches (fériés ou non) et des jours fériés (exceptés dimanches fériés) est constante d'une année sur l'autre pour chaque mois (à l'exception des mois de février à cause des années bissextiles), les coefficients estimés pour la correction de l'effet des jours ouvrables doivent alors être interprétés en référence aux jours fériés. Ainsi, le coefficient des lundis ouvrables ne correspond pas à l'effet brut d'un lundi ouvrable, mais à l'effet de ce jour relativement à celui d'un jour férié. Par exemple, la consommation des ménages en hôtels, cafés, restaurants, liée au tourisme, est plus forte lors d'un jour férié que lors d'un jour ouvrable. Les coefficients devant le nombre de jours ouvrables sont alors négatifs.

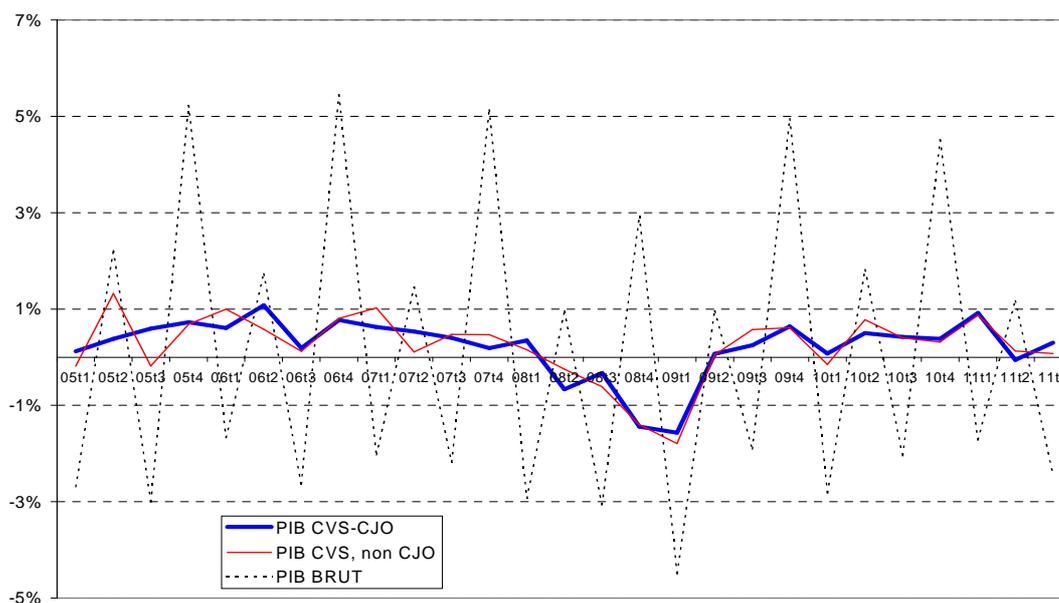
En pratique, le travail statistique pour estimer l'effet des jours ouvrables s'effectue en suivant différentes étapes, dont un certain nombre de tests. Ces étapes et la méthode d'ensemble sont formalisées en *annexe 5*.

Le choix des modèles et l'estimation des coefficients sont effectués une fois par an, au moment de l'élaboration du compte provisoire de l'année écoulée.

À titre d'illustration, l'effet absolu des jours ouvrables représente couramment de 0,1 à 0,5 point de croissance trimestrielle du PIB (*graphique 3*), et de 0,1 à 0,2 point de croissance annuelle. L'effet CJO sur le PIB est plus faible que celui de la CVS, mais l'impact peut être occasionnellement très important sur l'estimation de la croissance trimestrielle.

Plus précisément la croissance du PIB en données CVS (non CJO) au troisième trimestre de 2011 est de + 0,1 %, contre + 0,3 % pour le PIB en données CVS-CJO, lors de l'estimation des résultats détaillés du troisième trimestre 2011. L'impact de la correction des jours ouvrés sur le taux de croissance du PIB est de + 0,2 point au troisième trimestre, après - 0,2 point au deuxième. Cet effet au deuxième trimestre s'explique : le mois de mai 2011 comportait exceptionnellement peu de jours fériés, donc plus de jours ouvrés qu'habituellement, le 1^{er} mai comme le 8 mai tombant un dimanche.

Graphique 3 : Produit intérieur brut trimestriel en volumes chaînés, en données brutes, données CVS, et en données CVS-CVJO



Source : comptes nationaux, base 2005, publication des résultats détaillés du 3^e trimestre 2011.

2.2.2. La correction des variations saisonnières

La plupart des séries de comptes (production, consommation,...) présentent une allure saisonnière très marquée. La production est par exemple moins dynamique en juillet et août où beaucoup d'entreprises réduisent leur activité en raison des congés d'été. La consommation d'énergie est très élevée pendant les mois d'hiver du fait des dépenses de chauffage qui s'ajoutent à la consommation habituelle. Comme le montre aussi le *graphique 3*, les évolutions du PIB trimestriel brut sont très fluctuantes, avec des régularités saisonnières : le PIB est systématiquement en baisse à l'été. Il ne faut évidemment pas en conclure que la situation conjoncturelle y est dégradée.

Pour comprendre les évolutions conjoncturelles sous-jacentes, il est nécessaire de disposer d'une mesure indépendante des effets saisonniers.

Dans les comptes trimestriels, l'estimation des effets des variations saisonnières est effectuée sur les indicateurs, préalablement corrigés des effets de jours ouvrables. Deux méthodes de désaisonnalisation sont utilisées. La première, dite méthode de Buys-Ballot, est une méthode d'écart à la moyenne. Elle consiste à estimer l'effet saisonnier d'un trimestre donné comme la moyenne de la valeur de la série pour ce trimestre, sur l'ensemble des années observées, diminuée de la moyenne totale (*annexe 6*). C'est une méthode fruste, qui ne prend en compte aucune évolution de la saisonnalité. Elle n'est utilisée que pour les séries très courtes, par exemple lorsqu'un indicateur montre une rupture de saisonnalité une année donnée et doit être désaisonnalisé séparément sur deux périodes.

La seconde méthode de désaisonnalisation est la méthode implémentée dans le logiciel « X12-ARIMA » (*annexe 7*). Elle repose sur une série de moyennes mobiles appliquées à la série. Celles-ci permettent d'extraire les composantes saisonnières, mais nécessitent tout d'abord d'extrapoler la série sur plusieurs mois. L'extrapolation est effectuée à partir de modèles de type ARIMA.

La correction des variations saisonnières d'un indicateur donné est ré-estimée en continu, avec l'introduction de la ou des dernières périodes connues. Ceci est une cause fréquente de révisions. Le type de méthode (schéma additif ou multiplicatif) et les modèles ARIMA qui permettent d'extrapoler les séries avant d'évaluer la saisonnalité sont revus tous les ans, au moment de l'élaboration du compte provisoire de l'année écoulée.

2.3. Volumes aux prix de l'année précédente versus volumes à prix constants

Un des objectifs majeurs des comptes nationaux est de fournir les évolutions des **volumes** des grands agrégats économiques, pour éliminer les effets des prix dans l'analyse de la croissance du produit intérieur brut, de la consommation... Les volumes se rapprochent de la notion de quantité. Cependant, additionner tout simplement les quantités des produits élémentaires n'a pas de sens : la quantité consommée de voitures n'est pas comparable avec celle de bicyclettes. Il faut rendre ces quantités commensurables et ceci est effectué en attribuant aux quantités élémentaires le prix qu'elles avaient à une période donnée.

Les valeurs, elles, sont comparables : la valeur d'un agrégat est obtenue par somme des valeurs des produits élémentaires, définies, à chaque période, par le produit de leur quantité et de leur prix moyen au cours de cette période.. Le prix d'un agrégat est alors défini de façon symétrique à son volume : c'est le ratio de la valeur agrégée au volume.

Le calcul du volume d'un agrégat suppose de pondérer les volumes des éléments qui le composent par leur prix. Le choix de la période qui définit la structure des prix utilisée comme pondération pour évaluer les volumes est donc crucial. Deux choix sont envisageables :

- **le calcul de volumes aux prix constants de l'année de base** (par un raccourci parfois trompeur le terme de « volumes constants » est fréquemment utilisé) ; en 2011 par exemple, cela consiste à pondérer les niveaux élémentaires composant un agrégat (produits ou branches de l'économie) par leur prix relatif de l'année de base (en « base 2005 » l'année de référence ou de base est l'année 2005) ;
- **le calcul de volumes aux prix de l'année précédente** : en 2011, on pondère les niveaux élémentaires composant un agrégat par leur prix relatif le plus récent, à savoir celui obtenu dans les comptes de l'année 2010.

À titre d'exemple, pour calculer l'évolution de la consommation totale en volume entre 2010 et 2011, le calcul en volumes aux prix de l'année précédente amène à pondérer l'évolution en 2011 de la consommation de produits électroniques en volume par son prix relatif en 2010, qui est plus faible que ce qu'il était en 2005 compte tenu des fortes baisses de prix intervenues entre 2005 et 2010. Si la consommation de produits électroniques est dynamique en 2011, la consommation totale est donc moins dynamique en 2011 calculée en volumes aux prix de l'année précédente qu'en volumes aux prix constants de l'année de base. Le problème est d'autant plus aigu que l'on s'éloigne de l'année de base.

Les comptes nationaux annuels sont eux-mêmes calculés et publiés en volumes aux prix de l'année précédente, **chaînés** (par un raccourci parfois trompeur le terme de « volumes chaînés » est fréquemment utilisé), avec une année de référence fixe (l'année de base). L'idée des indices chaînés est de conserver les évolutions, et non les niveaux, des volumes aux prix de l'année précédente d'une année sur l'autre, et de chaîner ces évolutions à partir de la valeur d'une année de référence donnée.

En effet, les niveaux des volumes aux prix de l'année précédente non chaînés ne peuvent pas être utilisés en séries temporelles puisque l'évolution entre les volumes de deux années consécutives comprend à la fois une évolution des prix (entre les deux années de référence) et une évolution des volumes.

L'avantage des estimations de volumes aux prix de l'année précédente chaînés est donc à la fois d'être utilisées en séries temporelles et de capter les déformations de la structure des prix relatifs : autrement dit, ils fournissent une description plus satisfaisante de la réalité économique lorsque les prix de certains produits évoluent de façon très différente des autres prix - comme c'est le cas des produits des nouvelles technologies. Les volumes chaînés posent en revanche certains problèmes :

- ils peuvent entraîner une dérive lorsque les prix oscillent au lieu d'évoluer sur une tendance, comme ce peut être le cas par exemple des prix de l'agriculture ou des prix de l'énergie¹¹ ;

¹¹ À ce sujet, voir par exemple l'analyse effectuée par Berthier J.-P. [4].

- par ailleurs, ils perdent la propriété d'additivité par rapport aux volumes aux prix de l'année précédente non chaînés ou aux volumes à prix constant. Ainsi les équilibres comptables emplois-ressources ne sont plus respectés en niveau et les agrégats ne peuvent pas être obtenus directement comme somme des niveaux inférieurs. L'*annexe 8* permet de comprendre pourquoi les volumes chaînés n'ont pas cette propriété. L'absence d'additivité rend plus complexe l'élaboration des comptes, ainsi que leur publication. Ceci peut en outre poser des difficultés pour certains utilisateurs qui se servent du cadre comptable : dans les modèles macro-économétriques, par exemple, le cadre comptable est utilisé comme garant de la cohérence des prévisions.

D'un point de vue pratique, l'évaluation de volumes chaînés est plus complexe pour des comptes trimestriels que pour des comptes annuels. Ainsi, pour une fréquence inférieure à l'année, plusieurs techniques de chaînage peuvent être utilisées¹².

Pour prendre en compte rapidement les déformations de prix, il est possible de chaîner les volumes aux prix du trimestre précédent. Cette méthode est déconseillée, à la fois par le SCN 93 [20] et Eurostat [9]. En effet, elle peut entraîner une dérive parce que les données infra-annuelles fluctuent souvent plus que les données annuelles, en particulier lorsqu'elles sont saisonnières.

Il est alors conseillé d'effectuer le chaînage à partir d'une base de prix annuelle. Celui-ci peut être construit d'au moins trois façons différentes.

- Le chaînage le plus proche du chaînage annuel correspond à des comptes trimestriels évalués aux prix moyens de l'année précédente (« recouvrement annuel »). Ceci entraîne une discontinuité méthodologique (sans nécessairement une forte portée empirique) aux premiers trimestres : l'évolution de volume du premier trimestre comprend à la fois un changement de volume et un changement de prix, puisque le quatrième trimestre de l'année précédente n'est pas évalué sur la même base de prix. L'ampleur de l'« effet de chaînage » sur l'évolution des premiers trimestres dépend de la dispersion de l'inflation selon ses différentes composantes. L'avantage de cette méthode est qu'elle permet de préserver la propriété d'additivité infra-annuelle : la somme des quatre comptes trimestriels brut d'une année est égale au compte annuel de la même année.
- Pour résoudre la discontinuité méthodologique sur le premier trimestre, une autre méthode consiste à évaluer le quatrième trimestre de l'année N à la fois aux prix moyens de l'année précédente et aux prix de l'année en cours pour que le maillon entre le premier trimestre de l'année N+1 et le quatrième trimestre de l'année N corresponde à l'évolution du volume aux prix de l'année N (recouvrement sur un trimestre). Cette méthode, en revanche, ne préserve pas l'additivité infra-annuelle.
- Dans une dernière approche, les maillons sont des glissements annuels, entre par exemple le premier trimestre de l'année N+1 calculé aux prix moyen de l'année N et le premier trimestre de l'année N aux prix moyens de l'année N (recouvrement sur toute l'année). Le manuel du FMI [12] déconseille cette dernière méthode, qui n'évite pas les discontinuités et entraîne une forte volatilité.

¹². Toutes ces techniques sont décrites en détail dans le manuel sur les comptes trimestriels du FMI [12] ainsi que dans Arnaud F. [1].

2.4. Le calcul des volumes dans les comptes trimestriels

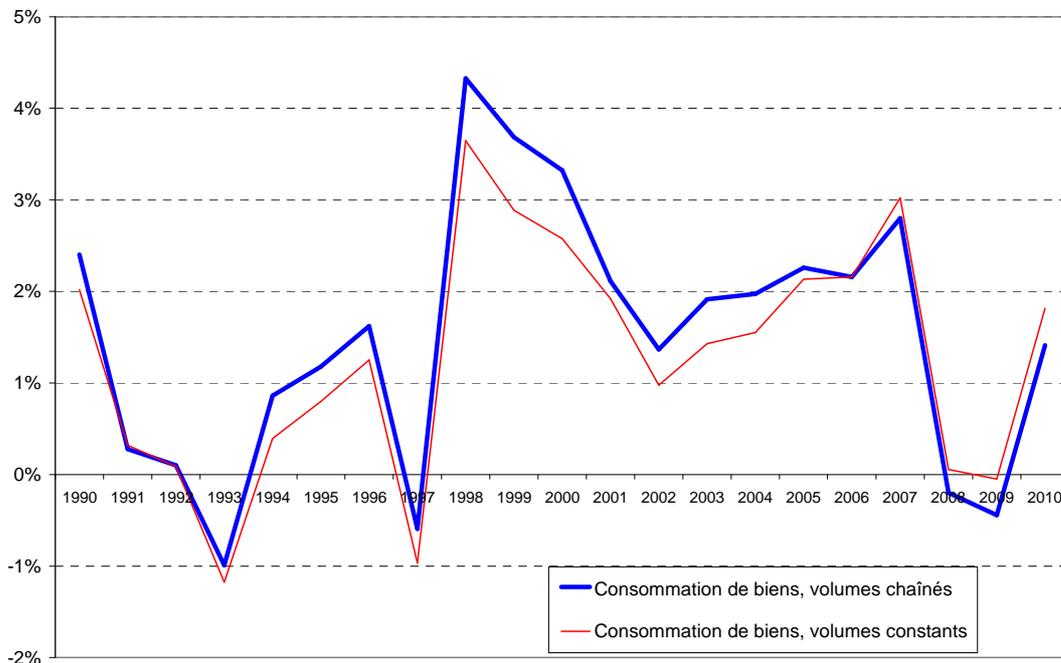
2.4.1. Des comptes publiés en volumes chaînés...

Face à ces difficultés, les comptes trimestriels français ont jusqu'en 2006 fait le choix de publier des volumes calculés aux prix constants de l'année de base. L'un des arguments reposait sur le fait qu'en France, les écarts entre volumes chaînés et volumes aux prix constants de l'année de base étaient relativement faibles. C'était en partie imputable au fait que les produits des nouvelles technologies, dont les prix peuvent baisser rapidement, n'avaient pas un poids très élevé dans l'économie française.

Ce dernier argument devenait toutefois de moins en moins pertinent avec la diffusion des nouvelles technologies dans l'économie française et l'accentuation de l'écart entre les deux concepts (*graphique 4*). En base 2005, sur l'année 2010, l'écart entre le taux de croissance annuel des dépenses de consommation des ménages en biens (agrégat publié mensuellement par l'Insee) en volumes chaînés, et celui en volumes aux prix constants de l'année de base, est de 0,4 point.

Le souci d'harmoniser les méthodes de construction des comptes trimestriels au niveau européen, conjugué au constat d'un creusement des écarts entre volumes chaînés et volumes aux prix constants de l'année de base sur certains couples produits \times opérations, a donc conduit à revenir sur ce choix : les comptes trimestriels français sont depuis 2007 publiés en volumes aux prix de l'année précédente chaînés, selon la méthode dite de recouvrement annuel (« *annual overlap* ») dont le principal avantage est de préserver la propriété d'additivité infra-annuelle.

Graphique 4 : Évolution annuelle des dépenses de consommation des ménages en biens - Comparaison des volumes chaînés et des volumes aux prix constants de l'année de base 2005



Source : comptes nationaux, base 2005, données CVS-CJO (résultats détaillés du 3^e trimestre 2011)

2.4.2. ...mais élaborés en volumes aux prix constants de l'année de base

Même s'ils sont publiés en « volumes chaînés », les comptes trimestriels continuent toutefois d'être élaborés en volumes aux prix constants de l'année de base (« volumes constants »).

En effet, dans la mécanique de construction des comptes trimestriels, les équilibrages des ressources et emplois (ERE) sont obtenus en soldant sur certaines opérations. Par exemple, s'agissant de la plupart des biens, les ERE sont soldés sur les variations de stocks (cf. chapitre 3), ce qui suppose que la propriété d'additivité soit vérifiée. Cela n'est évidemment pas propre aux comptes trimestriels : les comptables annuels vérifient également l'équilibrage des ERE qu'ils calculent en volume aux prix de l'année précédente (non chaînés). Mais à la différence des comptables annuels qui travaillent sur une année à la fois, les comptables trimestriels travaillent sur des séries temporelles. Il faut donc que la propriété d'additivité de ces séries soit vérifiée. Or :

- les séries en volumes aux prix de l'année précédente chaînés ne sont pas additives ;
- les séries en volumes aux prix de l'année précédente (non chaînés) sont certes additives, mais les sauts qui se produisent à chaque premier trimestre les rendent impropres à la démarche économétrique de l'étalonnage-calage.

Il aurait été théoriquement envisageable de procéder à l'estimation des étalonnages directement sur des séries en volumes chaînés, puis de passer aux volumes aux prix de l'année précédente (non chaînés) pour équilibrer le TES et chaîner ensuite les comptes obtenus. Mais une telle approche aurait été très lourde en termes de programmation. Il a paru plus simple de continuer d'élaborer le compte en volumes aux prix constants de l'année de base et, une fois obtenu le compte en volumes aux prix constants de l'année de base, à le passer en volumes aux prix de l'année précédente, puis à le chaîner. Au demeurant, le choix de continuer à travailler en volumes aux prix constants de l'année de base est sans incidence pour l'utilisateur puisque, *in fine*, les séries trimestrielles publiées en volumes chaînés sont complètement cohérentes avec les séries des comptes annuels.

Techniquement, trois étapes se succèdent.

- Dans une première étape, des volumes annuels à prix constants sont reconstruits à partir des comptes annuels chaînés (étape de « déchaînage »). Construire des comptes à prix constants par enchaînement d'indices de volumes ou de prix demande en principe d'additionner des volumes chaînés au niveau de détail le plus fin. Cependant, pour des raisons de simplicité de mise en œuvre, les comptes trimestriels utilisent les volumes chaînés à un niveau relativement agrégé, celui des étalonnages, c'est-à-dire en général le niveau de 48 produits, pour reconstruire à partir des comptes annuels des volumes à prix constants.
- Dans une deuxième étape, la technique d'étalonnage-calage (décrite dans ce chapitre précédemment) et d'évaluation du TES (décrite au chapitre 3) est appliquée pour construire des comptes trimestriels en volume aux prix constants de l'année de base.
- Dans une troisième étape, les comptes trimestriels en volume aux prix constants de l'année de base sont transformés en comptes en volume aux prix de l'année précédente, puis chaînés à tous les niveaux pour obtenir les comptes qui seront *in fine* publiés.

Cette description schématique rencontre toutefois une difficulté spécifique : lors de la première étape, dans la mesure où les volumes chaînés ne respectent pas les équilibres comptables et où les comptes trimestriels ne partent pas des niveaux les plus fins mais généralement du niveau à 48 produits, les équilibres comptables (équilibres ressources-emplois, comptes de branches, etc.) ne sont pas assurés au niveau de travail des comptes trimestriels. Ainsi, pour une branche donnée, le volume aux prix constants « spontané » de la valeur ajoutée n'est pas égal à la différence entre le volume aux prix constants de la production et celui de la consommation intermédiaire de la branche. Dans les comptes trimestriels aux prix constants, l'équilibrage est alors obtenu lors du « déchaînage » des comptes annuels grâce aux postes comptables qui sont construits par soldes. Il en résulte un écart de chaînage sur ces postes. En simplifiant, pour l'équilibrage des comptes de biens et services, l'écart de chaînage est reporté sur les variations de stocks pour les biens, et sur la production pour les services. Pour équilibrer le tableau des entrées intermédiaires, l'écart de chaînage est reporté sur les marges (consommations intermédiaires totales des branches et consommations intermédiaires totales en produits) tandis que pour le compte de production, c'est la valeur ajoutée qui est obtenue par solde et les transferts des services aux entreprises qui permettent de solder la somme des transferts à zéro.

Il importe, pour que les comptes trimestriels publiés soient bien calés sur les comptes annuels à la fin de la troisième étape (au moment du passage du compte aux prix de l'année précédente), d'effectuer l'opération inverse en retranchant les écarts de chaînage aux postes soldes obtenus après équilibrage du TES aux prix constants de l'année de base. De la sorte, ces postes soldes seront bien calés sur leurs équivalents annuels en volumes chaînés.

Cette séquence assure que les comptes trimestriels sont cohérents à tous les niveaux avec les comptes annuels publiés : ce qui veut dire que, quels que soient le produit et l'opération considérés, la somme des volumes (bruts ou CVS) des quatre trimestres d'une année est égale au compte annuel en volume publié pour la même année.

Certains agrégats sont généralement considérés comme n'étant pas susceptibles d'être chaînés, et de fait, ne le sont pas. C'est notamment le cas du solde extérieur (exportations moins importations), ainsi que des variations de stocks. Ces deux cas présentent la particularité de présenter des fluctuations très importantes impliquant des changements de signe fréquents.

Notamment, l'établissement de séries de variations de stocks en volume pose des problèmes spécifiques d'au moins deux ordres :

- des fluctuations très fortes, avec des changements de signe et des passages éventuels par des montants voisins de zéro ;
- la difficulté qu'il y a à maîtriser les indices de prix, produit par produit, du fait que les variations de stocks constituent le poste d'ajustement, en valeur comme en volume.

Combinés, ces deux problèmes peuvent entraîner des phénomènes explosifs en cas de chaînage.

Une voie alternative pourrait être d'appliquer une méthode originale que l'on peut qualifier de chaînage additif avec un contrôle des indices de prix. C'est ce qui est proposé par Berthier [4] et retenu pour chaîner les comptes annuels (Braibant M. et Pilarski C. [6]).

2.4.3. Le calcul des contributions dans les comptes en volumes chaînés

Le fait que les comptes en volumes chaînés ne soient pas additifs complique les calculs des contributions, familiers aux utilisateurs des comptes trimestriels. L'*annexe 9* explicite ces difficultés et les choix faits pour définir des contributions dont les propriétés soient satisfaisantes.

2.4.4. Construction des comptes en valeur, volume, prix, au niveau élémentaire

Lors de la construction des comptes trimestriels en volumes, valeurs et prix, ceux-ci sont le plus souvent élaborés à partir d'un indicateur de volume et de l'indicateur de prix correspondant. Parfois, l'information n'est pas disponible en volume ou s'avère être de meilleure qualité en valeur. Les comptes trimestriels associés sont alors construits à partir de l'indicateur en valeur et d'un indicateur de prix associé. Il est plus rare qu'un indicateur de volume et un indicateur de valeur soient utilisés conjointement pour un compte donné.

Plus précisément, les comptes sont établis suivant différents schéma-types.

- Si l'indicateur est en volume (ex : nombre d'immatriculations de véhicules neufs) :
 - élaboration du compte en volume : indicateur brut de volume -> correction des variations saisonnières (CVS) et des jours ouvrables (CJO)-> indicateur CVS-CJO de volume -> étalonnage-calage -> compte CVS-CJO en volume ;
 - indicateur de prix : indicateur brut de prix -> correction des variations saisonnières¹³-> indicateur CVS de prix ;
 - élaboration du compte en valeur : indicateur en valeur = compte en volume * indicateur de prix -> étalonnage- calage -> compte en valeur ;
 - compte de prix = compte en valeur / compte en volume
- Si l'indicateur est en valeur (ex : indice de chiffre d'affaires) :
 - on calcule d'abord le compte en valeur ;
 - l'indicateur du compte en volume est alors le compte en valeur déflaté par l'indicateur de prix.
- Si un indicateur existe en valeur et en volume :
 - les deux comptes sont établis séparément par le schéma d'étalonnage-calage ;
 - le compte en prix se déduit comme ratio des comptes en valeur et en volume.

2.5. La rétropolation dans les comptes trimestriels

2.5.1. Des principes généraux appliqués depuis 1980

Les principes généraux d'élaboration des comptes trimestriels nécessitent des séries temporelles qui aient un recul important. Ils sont mis en œuvre en base 2005 à compter du 1^{er} trimestre 1980. Tous les comptes estimés depuis cette date et jusqu'au dernier trimestre connu sont publiés et peuvent donner lieu à des révisions. En général, plus on remonte dans le temps et plus l'ampleur des révisions est faible. Les révisions sont de plus forte ampleur lors de la publication des premiers résultats du premier trimestre au mois de mai, avec l'intégration des comptes annuels définitif, semi-définitif et provisoire, et la ré-estimation des relations d'étalonnage ainsi que des coefficients CVS et CJO.

Pour bénéficier de ce recul des séries depuis 1980, un travail de rétropolation des indicateurs a souvent été nécessaire. Par exemple, s'agissant de l'IPI, dont les indicateurs sont disponibles et publiés depuis 1990 en base 2005, chaque série a été rétropolée avant 1990. Pour ce faire, les anciens indicateurs servant aux comptes trimestriels des précédentes bases ont été utilisés, moyennant des matrices de passage entre ancienne et nouvelle nomenclatures. Ces matrices de passage ont été estimées à un niveau relativement agrégé, sur chacune des opérations des comptes. Elles ne pouvaient être aussi précises que les matrices de passage utilisées par les fournisseurs de données, le plus souvent directement au niveau individuel. Lorsqu'un biais systématique apparaissait dans l'utilisation des matrices agrégées, la rétropolation des indicateurs en était corrigée.

L'existence d'indicateurs n'est pas toujours assurée sur le passé. C'est le cas notamment pour les familles d'indicateurs dont le champ s'est développé au cours du temps. Une série de compte peut alors résulter d'un lissage sur une première période, et d'un étalonnage-calage sur une seconde période. Une série de compte peut également résulter du traitement de deux indicateurs successifs, lorsque le deuxième s'avère meilleur que son prédécesseur mais n'a pas le même recul temporel (cas des salaires par branche, cf. chapitre 4).

¹³ Les indicateurs de prix sont en général insensibles aux jours ouvrables, ils ne sont donc corrigés que des variations saisonnières.

2.5.2. Une rétopolation spécifique des comptes entre 1949 et 1979

Par ailleurs, comme en base 2000, la base 2005 a donné lieu à un travail de « rétopolation » des comptes trimestriels avant 1980, c'est-à-dire d'adaptation des séries historiques aux nouveaux niveaux et à la nouvelle nomenclature.

La méthodologie liée à cet exercice est plus simple que celle d'estimation des comptes trimestriels en production courante, dans la mesure où :

- il n'y a pas nécessité d'extrapoler des données non encore connues ;
- plutôt que de repartir des indicateurs traditionnels, tels que l'IPI, on utilise comme indicateurs d'anciennes versions des comptes trimestriels, notamment car beaucoup d'indicateurs ayant servi à les construire ne sont plus disponibles ;
- le niveau de travail est plus agrégé qu'en production courante : en production courante, les comptes trimestriels travaillent à un niveau bien plus détaillé (une cinquantaine de postes, cf. chapitre 3) que celui de publication (17 postes). Pour la rétopolation, le niveau de travail retenu est directement le niveau de publication en 17 postes (« A17 »). Il n'était pas envisageable de retenir un niveau plus désagrégé car les comptes utilisés comme indicateurs ne sont eux-mêmes archivés qu'avec un niveau de désagrégation très limité (entre 15 et 20 postes), et dans des nomenclatures différentes de celle utilisée aujourd'hui par les comptes nationaux ; au total, le niveau de travail est directement celui de la publication.

Cependant, tous les principes généraux de construction des comptes trimestriels sont respectés (étalonnage, calage sur données annuelles, données CVS-CJO), si bien que les séries sont homogènes avant et après 1980.

Arnaud F. et Mahieu R. [3] explicitent la façon dont la base 2000 avait été rétopolée avant 1980 et depuis 1949, à partir des comptes annuels rétopolés d'une part, des comptes trimestriels en bases 1956 et 1970 d'autre part.

La base 2005 a donné lieu à l'application des mêmes principes, en partant de ces travaux publiés sur la base 2000, et en utilisant les différentes matrices permettant de passer de la nomenclature économique de synthèse (NES) à 16 positions de la base 2000 à la nomenclature agrégée à 17 positions de la base 2005.

Sauf changements ponctuels et rares des données annuelles, les séries avant 1980 ne donnent pas lieu à révision.

Annexe 1 : Exemple de choix de modèle d'étalonnage

Un modèle d'étalonnage se choisit parmi 6 types de modèles résumés dans le tableau suivant :

Numéro Modèle	Description synthétique	Description générale		
		Niveau/variation	Constante	Estimation (1)
1	$\Delta Y_t = \alpha + \beta \Delta X_t + \varepsilon_t$	En variation	Oui	MCO
2	$\Delta Y_t = \beta \Delta X_t + \varepsilon_t$	En variation	Non	MCO
3	$Y_t = \alpha + \beta X_t + \varepsilon_t$	En niveau	Oui	MCO
4	$Y_t = \alpha + \beta X_t + u_t$	En niveau	Oui	MCQG
5	$Y_t = \beta X_t + \varepsilon_t$	En niveau	Non	MCO
6	$Y_t = \beta X_t + u_t$	En niveau	Non	MCQG

(1) MCO : moindres carrés ordinaires; MCQG : moindres carrés quasi-généralisés

Avec : ε_t : résidu (bruit blanc)

u_t : qui suit un modèle autorégressif d'ordre 1

Source : comptes nationaux trimestriels

L'exemple détaillé ci-dessous illustre le choix du modèle d'étalonnage de la consommation des ménages en volume aux prix constants d'automobiles neuves. L'indicateur provient du comité des constructeurs français d'automobiles (CCFA) et résulte du produit des immatriculations des particuliers par un indicateur de puissance fiscale des véhicules nouvellement immatriculés. Le tableau suivant correspond au résultat des estimations d'étalonnage entre le compte et cet indicateur, sur la période [1980-1990] pour les 6 modèles définis précédemment.

Tableau : résultat des modèles, cas des dépenses en automobiles neuves par les ménages

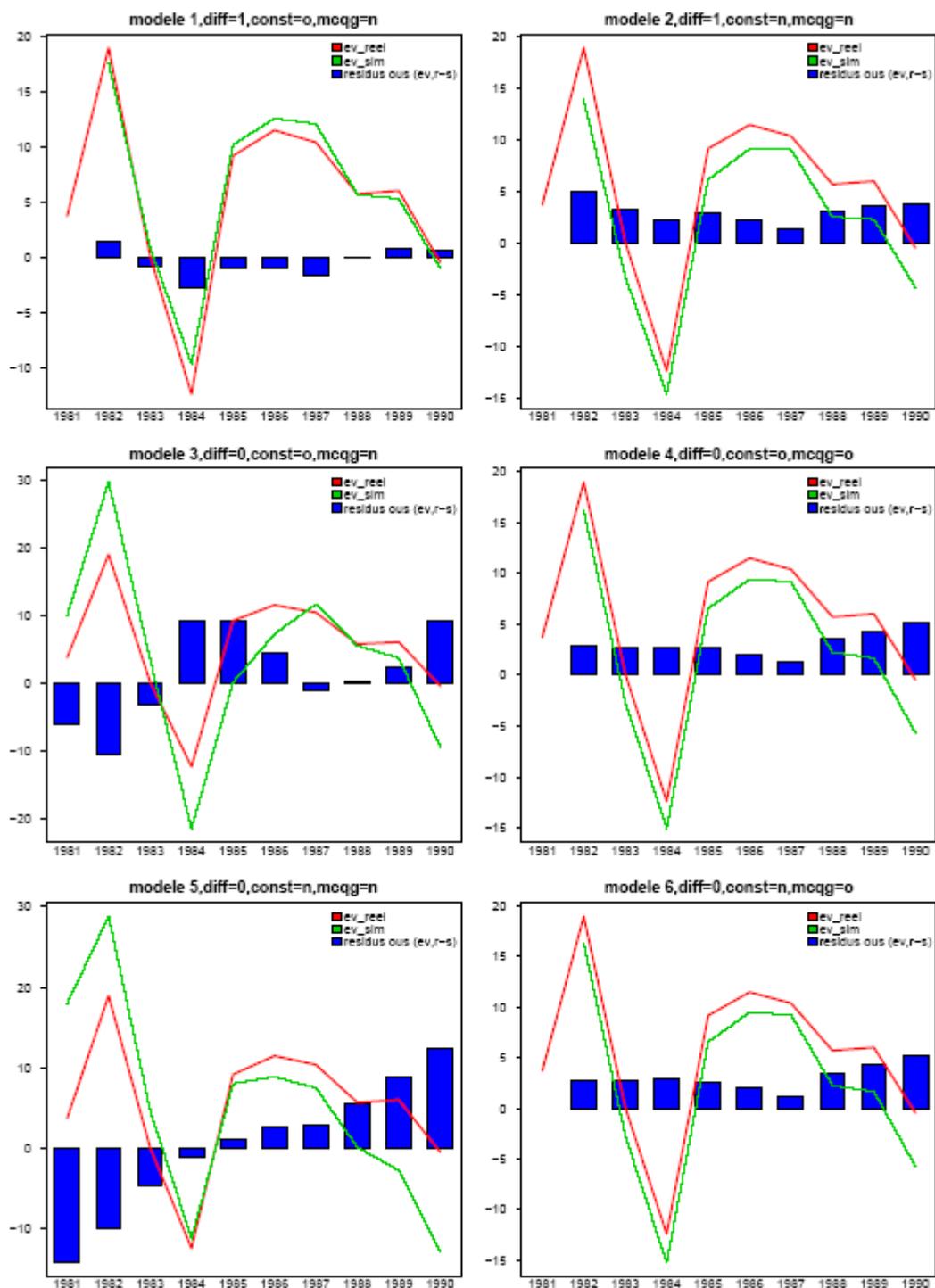
	Modèle 1	Modèle 2	Modèle 3	Modèle 4	Modèle 5	Modèle 6
Résidus récurrents (RMSE) evol	0.7	3.6	8.9	5.1	12.2	5
Test de Ljung Box L = 1 (p.value)	0.2	0.28	0.17	0.27	0.02	0.3
Coefficients estimés						
constante (Student)	65.92 (5.85)	0 (0)	-1029.61 (-2.07)	22.52 (0.07)	0 (0)	0 (0)
indicateur (Student)	1.87 (11.13)	2.02 (5.62)	3.44 (5.68)	2.14 (5.91)	2.19 (33.7)	2.16 (15.81)
rho	0	0	0	0.88	0	0.88

Source : comptes nationaux trimestriels

Les modèles 3 et 5 sont les modèles en niveau sans auto-corrélation de la cale, avec et sans constante. Le test de Ljung Box (une des formes de test de portmanteau qui permet de juger de l'auto-corrélation des résidus) rejette le modèle 5 (probabilité faible, le résidu n'est pas un bruit blanc). Les modèles 4 et 6 sont les modèles avec auto-corrélation de la cale, avec et sans constante ; dans le modèle 4 le test de Student indique que la constante n'est pas significative et qu'il faut privilégier, parmi ces deux modèles, le modèle 6. Les modèles 1 et 2 sont les modèles en différence, avec et sans constante ; d'après le test de Student, la constante est significative et il faut dans ce cas privilégier le modèle 1 avec constante.

Finalement, le choix doit donc être effectué entre les modèles 1, 3 et 6. Dans le tableau, l'écart-type des résidus récurrents montre que le pouvoir prédictif du modèle 1 est meilleur. Par ailleurs, les graphiques suivants montrent que l'erreur de prévision est plus faible pour ce même modèle 1, et n'est pas biaisée (les courbes comparent l'évolution réelle de la série compte et l'évolution simulée par le modèle, les bâtonnets sont la différence entre les deux, soit l'erreur de prévision chaque année). Sur la période considérée, c'est donc le modèle 1 qui est privilégié. Notons que des tests complémentaires dits de racine unitaire indiquent par ailleurs si les séries de comptes et d'indicateur sont intégrées ou stationnaires, et permettent de légitimer le recours aux modèles 1 et 2 par rapport aux autres modèles.

Graphique : Résultat des estimations « in sample » selon les modèles, cas des dépenses en automobiles neuves par les ménages, 1980-1990



Source : comptes nationaux trimestriels

Annexe 2 : Test de portmanteau

Le test est effectué sur les résidus de l'équation d'étalonnage. Il repose sur l'estimation du coefficient d'auto-corrélation des résidus par l'équation :

$$\varepsilon_t = \rho \times \varepsilon_{t-1} + \eta_t$$

On a alors :

$$\hat{\rho} = \frac{\text{cov}(\varepsilon_t, \varepsilon_{t-1})}{\text{var}(\varepsilon_t)}$$

et, sous l'hypothèse de normalité de η_t :

$$\hat{\rho} \approx N\left(\rho, \frac{\sigma_\eta^2}{\sum \varepsilon_t^2}\right)$$

Sous l'hypothèse nulle $\rho = 0$, et en notant T le nombre d'observations, on a :

$$\hat{\rho} \approx N\left(0, \frac{1}{T}\right)$$

et le test de portmanteau (de type Ljung-Box) repose alors sur :

$$\hat{\rho}^2 * T \approx \chi^2(1)$$

$$\text{où } T = \begin{pmatrix} 1 & 0 & & & \\ 1 & 1 & & & \\ 1 & 1 & 1 & 0 & 0 \\ & & & 1 & 0 \\ 1 & & & 1 & 1 \end{pmatrix} \text{ de dimension } [4A, 4A]$$

Le lagrangien associé au programme s'écrit :

$$\begin{aligned} L &= y'y + 2\lambda' \left[MT \begin{pmatrix} x_{1,1} \\ y \end{pmatrix} - X \right] \\ &= y'y + 2\lambda' [m_1 x_{1,1} + \tilde{m}y - X] \end{aligned}$$

avec

$$m_1 = MT \begin{pmatrix} 1 \\ 0 \\ \dots \\ 0 \end{pmatrix} \text{ et } \tilde{m} = MT \begin{pmatrix} 0 \\ I_{4A-1} \end{pmatrix}$$

En dérivant par rapport à $x_{1,1}$, y et λ , on obtient

- (1) $\lambda' m_1 = 0$
- (2) $y = \tilde{m}' \lambda$
- (3) $m_1 x_{1,1} + \tilde{m}y = X$

En introduisant (2) dans (3), on obtient λ en fonction de $x_{1,1}$:

$$(3') \quad \lambda = (\tilde{m}\tilde{m}')^{-1} (X - m_1 x_{1,1})$$

À partir de (1) et de (3'), on obtient $x_{1,1}$:

$$x_{1,1} = m_1' (\tilde{m}\tilde{m}')^{-1} X / m_1' (\tilde{m}\tilde{m}')^{-1} m_1$$

À partir de (3'), on en déduit λ et à partir de (2), on en déduit y . x en découle puisque

$$x = T \begin{pmatrix} x_{1,1} \\ y \end{pmatrix}$$

Annexe 5 : Méthode d'évaluation de l'effet des jours ouvrables

On postule la relation suivante entre une variable brute et son équivalent CJO, « a » désignant l'année et « m » le mois :

$$Y_{a,m} = \sum_{i=1}^7 \alpha_i N_{a,m}^i + \beta F_{a,m} + \gamma \delta_{m,2} \delta_{a,biss} + Y_{a,m}^{CJO} \quad (1)$$

où :

$Y_{a,m}$ désigne la variable brute

$Y_{a,m}^{CJO}$ désigne la variable CJO

$N_{a,m}^i$ désigne le nombre de jours ouvrables de type i (i=1,...,6) le mois considéré, et $N_{a,m}^7$ désigne le nombre de dimanches le mois considéré (férié ou non)

$F_{a,m}$ le nombre de jours fériés (ne tombant pas un dimanche) le mois considéré

$\delta_{m,2} \delta_{a,biss}$ une indicatrice valant 1 les mois de février des années bissextiles

La série CJO, $Y_{a,m}^{CJO}$, est ainsi obtenue comme le résidu de l'équation économétrique estimée.

Les variables calendaires sont au préalable calculées en écart à la moyenne pour ne pas incorporer de composante saisonnière (ce qui ne change pas fondamentalement la spécification du modèle).

De plus, pour raisonner sur des séries ne présentant pas de composante saisonnière et qui soient stationnaires (autant que possible), le modèle est estimé en glissement annuel :

$$\Delta_{12} Y_{a,m} = \sum_{i=1}^7 \alpha_i \Delta_{12} N_{a,m}^i + \beta \Delta_{12} F_{a,m} + \gamma \delta_{m,2} I_a + \Delta_{12} Y_{a,m}^{CJO} \quad (2)$$

où : $I_a = 1$ si a est un multiple de 4

$I_a = -1$ si a-1 est un multiple de 4

$I_a = 0$ sinon.

Cependant, le total des jours ouvrables, des dimanches et des jours fériés ne tombant pas un dimanche est égal au nombre de jours du mois. Par exemple, pour janvier (m=1) :

$$\forall a, \sum_{i=1}^7 N_{a,1}^i + F_{a,1} = 31$$

Février présente évidemment la particularité d'avoir un nombre de jours fluctuant, de 28 ou 29 jours. De sorte que l'on peut écrire :

$$\sum_{i=1}^7 \Delta_{12} N_{a,m}^i + \Delta_{12} F_{a,m} = \delta_{m,2} I_a \quad (J)$$

La variable $\delta_{m,2} I_a$ s'interprète donc comme un terme correctif lié à l'existence d'années bissextiles.

Surgit donc un problème de colinéarité, puisqu'il existe une relation linéaire entre les différentes variables explicatives. Il suffit donc d'intégrer cette relation de colinéarité directement dans la régression, et on obtient :

$$\Delta_{12} Y_{a,m} = \sum_{i=1}^7 \varphi_i \Delta_{12} N_{a,m}^i + \theta \delta_{m,2} I_a + \Delta_{12} Y_{a,m}^{CJO} \quad (3)$$

avec $\varphi_i = \alpha_i - \beta$ et $\theta = \gamma + \beta$

Ainsi, l'effet du nombre de jours fériés n'est pas identifiable.

Les coefficients φ_i doivent alors être interprétés en référence aux jours fériés. Ainsi, le coefficient des lundis ouvrables ne correspond pas à l'effet brut d'un lundi ouvrable, mais à l'effet de ce jour relativement à celui d'un jour férié. Par exemple, la consommation des ménages en hôtels, cafés, restaurants, liée au tourisme, est plus forte lors d'un jour férié que lors d'un jour ouvrable. Les coefficients devant le nombre de jours ouvrables sont alors négatifs.

➤ **Construction des variables N_t^i**

Afin de retenir dans l'estimation les effets uniquement dus aux jours ouvrables, les variables de nombres de jours ouvrables sont désaisonnalisées par écart à la moyenne, selon la méthode de Buys-Ballot (*annexe 6*).

➤ **Choix des paramètres du modèle**

Deux évaluations de l'effet des jours ouvrables peuvent être menées : une évaluation additive ou multiplicative. L'évaluation multiplicative consiste à estimer l'effet des jours ouvrables sur le logarithme de l'indicateur. Cela suppose que cet effet est proportionnel au niveau de l'indicateur ; plus celui-ci est élevé, plus l'effet du calendrier est important. Le choix du modèle retenu est fait conjointement pour les deux types de corrections (CJO et CVS).

Pour faire ce choix entre modèle additif ou multiplicatif une première analyse graphique de la série vise à déterminer si la saisonnalité croît en même temps que la tendance sous-jacente, autrement dit si la saisonnalité est plutôt multiplicative. Pour valider cette analyse, un test économétrique simple est effectué, qui consiste à régresser la série des amplitudes annuelles à celle de leur moyenne. Un coefficient significatif signale qu'une évaluation multiplicative est plus appropriée. En outre, le test de stabilité des coefficients des jours ouvrables permet de percevoir une forte instabilité qui pourrait résulter d'un mauvais choix de méthode. Des tests plus complexes existent (par exemple le test non linéaire de Box-Cox) mais ne sont pas utilisés, en raison notamment du temps de traitement et d'analyse nécessaire à leur mise en œuvre.

Dans la suite, l'explication de la méthode de correction des jours ouvrables sera poursuivie en supposant que la méthode est additive. Il suffit de remplacer l'indicateur par son logarithme pour mettre en place la méthode multiplicative.

• **Première étape : choix de la différenciation**

Le modèle est estimé en glissement annuel pour travailler sur des données stationnaires. Malgré tout, le glissement annuel de l'indicateur peut ne pas être stationnaire. Un test de Dickey-Fuller est effectué sur le résidu de l'équation (3). Si l'hypothèse de non stationnarité n'est pas rejetée, l'équation (3) est évaluée en différence première.

• **Deuxième étape : test de l'existence d'effets des jours ouvrables**

Il s'agit d'un test de Fisher sur les paramètres du modèle (3) (différencié le cas échéant). Si le test conclut à la non-significativité jointe des coefficients, aucune correction des jours ouvrables n'est effectuée sur l'indicateur.

• **Troisième étape : test de l'existence d'un effet bissextile**

Il s'agit d'un test de Student sur le paramètre θ de l'indicatrice bissextile. Si le test conclut à la non-significativité du coefficient, ou si celui-ci est d'un ordre de grandeur anormal comparé à celui des autres jours, aucune correction d'effet bissextile n'est effectuée sur l'indicateur.

- **Quatrième étape : test de différenciation des mois de juillet et août**

Les mois de juillet et août peuvent perturber l'estimation de l'effet des jours ouvrables ; ce sont des mois pendant lesquels de nombreux actifs prennent des vacances, ce qui peut conduire à atténuer l'effet des jours ouvrés. Pour tester l'effet du calendrier sur les mois de juillet et août, les effets sont tout d'abord évalués sur les seuls mois d'été. Un test de Fisher permet alors de conclure à la significativité ou non des coefficients. S'ils ne sont pas significatifs, les coefficients des mois de juillet et août sont contraints à être nuls, c'est-à-dire qu'aucun effet des jours ouvrables n'est retranché à l'indicateur pour les mois d'été ; les coefficients des jours ouvrables sont alors estimés sur tous les autres mois. Si le test conclut à la significativité des effets de calendrier pour les mois d'été, l'évaluation des coefficients est effectuée sur l'ensemble des mois de l'année, sans aucune différenciation.

- **Cinquième étape : test des regroupements**

Il s'agit de cinq tests pour retenir des regroupements courants. Il est en effet intéressant de retenir des spécifications parcimonieuses, qui aident à la compréhension des coefficients retenus et à la précision des effets estimés. Chacun des tests correspond à un test de Fisher sur des coefficients particuliers. Ces tests sont effectués avec un modèle dans lequel les résidus suivent un modèle auto-régressif d'ordre 1.

Les regroupements testés sont les suivants :

- le coefficient du dimanche est contraint à zéro (cas qui se retrouve souvent sur des indicateurs de consommation) ;
- les coefficients des jours de la semaine sont contraints à être égaux et le paramètre du dimanche est contraint à zéro (cas qui se retrouve également sur des indicateurs de consommation) ;
- les paramètres du samedi et du dimanche sont contraints à zéro (cas correspondant plutôt à des indicateurs de production) ;
- les coefficients des jours de la semaine sont contraints à être égaux et les paramètres du samedi et du dimanche sont contraints à zéro (cas qui se retrouve souvent sur des indicateurs de production) ;
- les coefficients des jours de la semaine sont contraints à être égaux.

- **Sixième étape : existence de rattrapages ou d'anticipation**

Une fois le regroupement choisi, il s'agit de tester si les variables explicatives retardées et avancées d'un mois sont significatives ou non. Il est en effet possible que certaines entreprises rattrapent l'effet d'un jour férié en augmentant ensuite la production sur plusieurs semaines, ou à l'inverse anticipent cet effet sur plusieurs semaines précédant le jour férié. De même, il est envisageable que la consommation des ménages en textile soit reportée sur plusieurs semaines lorsque par exemple un samedi férié a beaucoup réduit la consommation de ce jour-là. Dans ces deux cas, l'équation (3) est estimée en intégrant non seulement les variables de nombre de jours ouvrables du mois en cours mais également du mois précédent ou du mois suivant. Le test est alors un test de Fisher sur les variables avancées, retardées et les deux ensemble.

- **Septième étape : tests de stabilité**

Il est important de tester la stabilité des coefficients estimés. Plusieurs changements de comportement pourraient en effet avoir modifié l'impact des jours fériés sur l'économie française. Par exemple, l'ouverture de plus en plus fréquente des commerces les dimanches et les jours fériés change probablement l'impact de ces jours-là sur la consommation des ménages.

Des tests de Chow sont estimés pour chaque année. Si des ruptures apparaissent, les coefficients de jours ouvrables sont estimés sur des plages glissantes de plusieurs années. Le choix du nombre d'années de la plage est délicat : il faut arbitrer entre une plage suffisamment petite pour s'adapter aux ruptures de la série, et suffisamment grande pour estimer convenablement les coefficients.

Annexe 6 : Méthode de désaisonnalisation par écarts à la moyenne (Buys-Ballot)

La méthode de Buys-Ballot est une méthode de désaisonnalisation par écarts à la moyenne. Elle consiste à estimer l'effet saisonnier d'un trimestre donné comme la moyenne de la valeur de la série pour ce trimestre, sur l'ensemble des années observées, diminuée de la moyenne totale.

Le principe repose sur la régression de la série à désaisonnaliser $I_{a,t}$ sur des indicatrices de trimestres¹⁴ : $\delta_{a,t}^1$, $\delta_{a,t}^2$, $\delta_{a,t}^3$, $\delta_{a,t}^4$, avec par exemple $\delta_{a,t}^1 = 1$ si t est un premier trimestre ($t = 1$), 0 sinon :

$$I_{a,t} = \alpha_1 \times \delta_{a,t}^1 + \alpha_2 \times \delta_{a,t}^2 + \alpha_3 \times \delta_{a,t}^3 + \alpha_4 \times \delta_{a,t}^4 + \varepsilon_{a,t}$$

Le modèle est estimé sans constante car la somme des indicatrices est égale à 1 quelle que soit la date.

Les coefficients estimés sont en fait égaux aux moyennes de la série sur chacun des trimestres : $\hat{\alpha}_1 = \bar{I}_{\{T1\}}$, ..., $\hat{\alpha}_4 = \bar{I}_{\{T4\}}$

où $\bar{I}_{\{T1\}} = \frac{1}{A} \sum_a I_{a,1}$ et A est le nombre d'années.

La somme des coefficients est proportionnelle à \bar{I} , la moyenne globale de la série : $\sum_{i=1}^4 \hat{\alpha}_i = 4\bar{I}$.

Les coefficients saisonniers sont finalement définis comme les écarts entre les moyennes sur chacun des trimestres et la moyenne totale :

$$a_1 = \hat{\alpha}_1 - \bar{I} = \bar{I}_{\{T1\}} - \bar{I}, \dots, a_4 = \hat{\alpha}_4 - \bar{I} = \bar{I}_{\{T4\}} - \bar{I}.$$

La série désaisonnalisée est alors obtenue comme :

$$I_{a,t}^{cvs} = I_{a,t} - a_1 \times \delta_{a,t}^1 - a_2 \times \delta_{a,t}^2 - a_3 \times \delta_{a,t}^3 - a_4 \times \delta_{a,t}^4$$

La désaisonnalisation est neutre sur l'année puisque les coefficients vérifient: $\sum_{i=1}^4 a_i = 0$.

¹⁴ Dans cet exemple, on suppose que la série $I_{a,t}$ est stationnaire. Si $I_{a,t}$ suit une tendance linéaire, le modèle de régression utilisé devient : $I_{a,t} = \lambda + \beta \times (4 \times a + t) + \alpha_1 \times \delta_{a,t}^1 + \alpha_2 \times \delta_{a,t}^2 + \alpha_3 \times \delta_{a,t}^3 + \alpha_4 \times \delta_{a,t}^4 + \varepsilon_{a,t}$. Ce modèle n'est pas identifiable spontanément. C'est pourquoi, une contrainte sur les coefficients doit être rajoutée (par exemple $\sum_{i=1}^4 \alpha_i = 0$)

Annexe 7 : Méthode de désaisonnalisation du logiciel X12-ARIMA

La méthode du logiciel X12 consiste à extraire d'une série I_t (préalablement corrigée des effets des jours ouvrés) une composante correspondant à la tendance-cycle TC_t , une composante saisonnière S_t et une composante irrégulière R_t . La correction CVS consiste à estimer l'effet de la composante saisonnière pour corriger la série initiale de cette composante.

Il est important de noter que selon les concepts suivis par les comptes trimestriels, la composante irrégulière R_t doit figurer dans la série désaisonnalisée, et donc également dans la série CVS-CJO. Par exemple, les effets d'une grève dans une branche donnée doivent apparaître clairement dans la série de production de la branche en données CVS-CJO.

Le principe du logiciel X12 repose sur plusieurs moyennes mobiles appliquées aux séries. Ainsi, la désaisonnalisation d'une série trimestrielle I_t débute par 3 étapes :

Étape 1 : première estimation des composantes saisonnières et irrégulières

Avec l'application d'une moyenne mobile symétrique $M(\cdot)$, il est possible de séparer la composante tendance-cycle (TC_t) d'un côté de la composante saisonnalité-irrégulier (S_t+R_t) :

$$M(I_t) = \frac{1}{8}(I_{t-2} + 2I_{t-1} + 2I_t + 2I_{t+1} + I_{t+2})$$

$I_t - M(I_t) = \left(\hat{S}_t + \hat{R}_t \right)$ est alors une estimation de la composante saisonnière et de l'irrégulier (S_t+R_t).

Étape 2 : première estimation de la seule composante saisonnière

Afin de supprimer la composante irrégulière de $\left(\hat{S}_t + \hat{R}_t \right)$, on applique sur chaque trimestre la moyenne mobile de type M' définie par :

$$M'(I_t) = \frac{1}{9}(I_{t-8} + 2I_{t-4} + 3I_t + 2I_{t+4} + I_{t+8}) \quad (\text{moyenne sur 5 ans})$$

Ainsi $M'(I_t - M(I_t)) = \hat{S}_t$ est une estimation de la composante saisonnière.

Étape 3 : normalisation de la seule composante saisonnière

Cependant cette estimation de la composante saisonnière ne vérifie pas la contrainte de somme nulle chaque année. Cette contrainte est alors imposée : si t est un des quatre trimestres de l'année a et \overline{S}_a est la somme des \hat{S}_t sur l'année a , alors on définit S'_t la composante saisonnière normalisée par $S'_t = \hat{S}_t - \frac{1}{4}\overline{S}_a$.

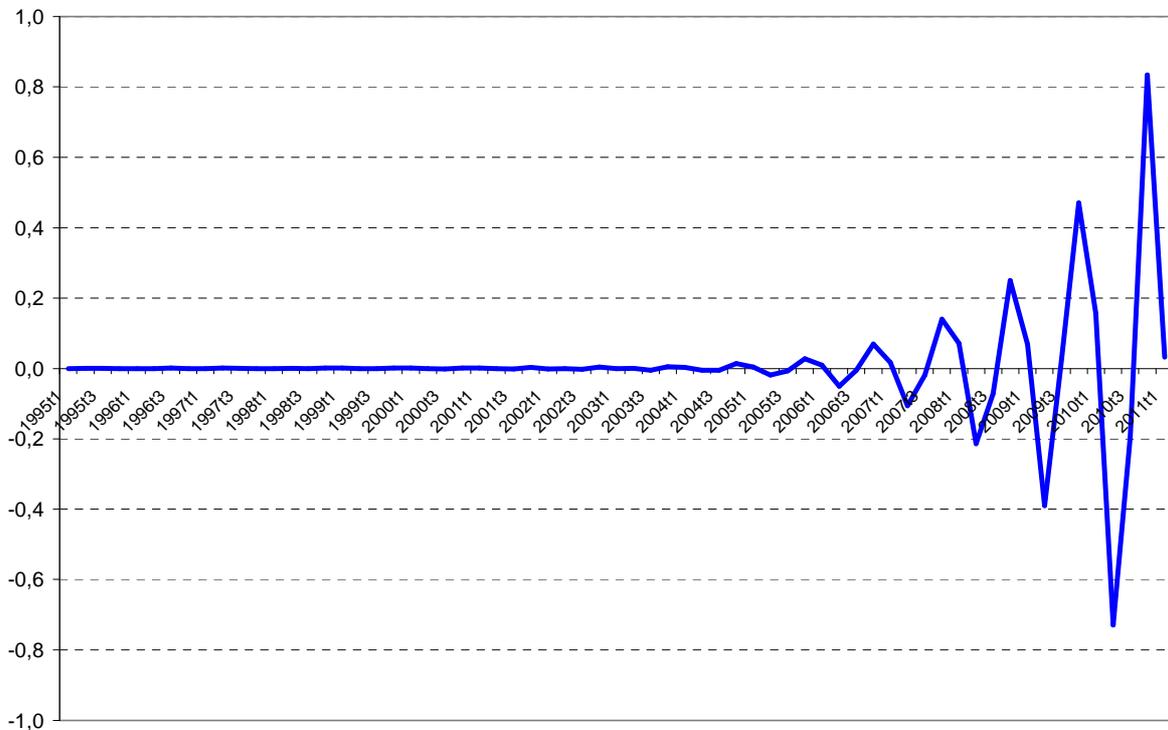
On obtient ainsi une première estimation de la série désaisonnalisée par $I_t - \hat{S}'_t = \hat{I}_t^{CVS}$.

Pour affiner cette première estimation, le logiciel réitère les trois étapes précédentes avec d'autres moyennes mobiles (en particulier des moyennes mobiles de Henderson), selon le même principe. On obtient alors une

seconde estimation des coefficients saisonniers que l'on transforme de nouveau afin qu'ils vérifient la propriété de somme nulle. En l'ôtant de la série initiale, on obtient la série désaisonnalisée finale.

Afin d'appliquer ce filtrage par moyennes mobiles en début et en fin de période, et de ce fait éviter les déformations liées au traitement aux « bords » de la série, il est nécessaire de la rétropoler et de l'extrapoler. La méthode du logiciel X12-ARIMA consiste à choisir un modèle ARIMA modélisant correctement la série, celui-ci servant de modèle prédictif pour extrapoler la série. C'est une cause de révision des séries CVS-CJO : lorsqu'un nouveau point est connu (ou une valeur récente révisée), l'extrapolation ARIMA est également modifiée ; en conséquence de ces deux facteurs de changement, l'effet saisonnier est estimé différemment. La révision a alors elle-même un profil saisonnier et s'amortit en remontant dans le passé (graphique suivant). Par ailleurs, par rapport à la version X-11, le logiciel X12-ARIMA dispose d'un module de détection et de correction des points atypiques (« outliers »). L'intérêt est de ne pas intégrer des valeurs aberrantes (qui peuvent être ponctuelles ou des changements durables de niveaux), pour ne pas perturber la détermination de la composante « tendance-cycle » et de la saisonnalité.

Graphique : Révision de l'évolution en % des importations en « aliments », entre les résultats détaillés du deuxième trimestre 2011 et ceux du premier trimestre 2011



Source : comptes nationaux trimestriels, base 2005

Annexe 8 : Des volumes à prix constants aux volumes aux prix de l'année précédente chaînés (et vice-versa)

La valeur d'un niveau élémentaire pour une année n donnée est égale au produit du prix et de la quantité de ce produit :

$$val_n = p_n \times q_n$$

où p_n est le prix et q_n la quantité.

Le volume à prix constants est défini par le produit du prix p_0 de l'année de base et de la quantité, de façon à ce que sur l'année de base, le volume soit égal à la valeur :

$$vol_n = p_0 \times q_n$$

L'indice de prix est alors défini par le ratio de la valeur et du volume :

$$IndP_n = \frac{val_n}{vol_n} = \frac{p_n}{p_0}$$

C'est donc l'évolution du prix entre l'année considérée et l'année de base.

Si l'on considère deux produits élémentaires i et j , l'agrégation de ces deux produits est simple : le volume (et la valeur) de l'agrégat sont les sommes des volumes (et des valeurs) de chaque produit.

$$vol_n(i+j) = p_0(i) \times q_n(i) + p_0(j) \times q_n(j)$$

et $val_n(i+j) = p_n(i) \times q_n(i) + p_n(j) \times q_n(j)$.

De cette façon, l'indice de volume, c'est-à-dire l'évolution du volume entre l'année considérée et l'année de base, peut s'écrire :

$$IndVol_n(i+j) = \frac{vol_n(i+j)}{vol_0(i+j)}$$

soit

$$IndVol_n(i+j) = \frac{q_n(i)}{q_0(i)} \times \frac{vol_0(i)}{vol_0(i+j)} + \frac{q_n(j)}{q_0(j)} \times \frac{vol_0(j)}{vol_0(i+j)}$$

Comme les volumes sont égaux aux valeurs pour l'année de base, l'indice de volume est la pondération des indices de volumes élémentaires par la structure des valeurs de l'année de base (c'est un indice de Laspeyres).

De façon générale, on peut écrire l'indice de volume d'un agrégat de niveaux élémentaires i comme :

$$IndVol_n(\sum i) = \frac{\sum_i vol_n(i)}{\sum_i vol_0(i)} = \sum_i \frac{q_n(i)}{q_0(i)} \times \left(\frac{p_0(i) \times q_0(i)}{\sum_j p_0(j) \times q_0(j)} \right)$$

En général, les volumes sont mesurés soit aux prix de l'année précédente, soit aux prix d'une année de base. Plus l'année courante est éloignée de l'année de base, plus les différences entre les structures de l'année précédente et de l'année de base s'accroissent.

Cependant, les volumes aux prix de l'année précédente ne peuvent pas être utilisés directement en séries temporelles. En effet, si l'on compare le niveau du volume de l'année 1 et le niveau du volume de l'année 2, exprimés aux prix de l'année précédente, on compare :

$$vol_1(\sum i) = \sum_i p_0(i) \times q_1(i) \text{ et } vol_2(\sum i) = \sum_i p_1(i) \times q_2(i)$$

c'est-à-dire que l'évolution entre ces deux volumes comprend à la fois une évolution des prix (entre l'année 0 et l'année 1) et une évolution des volumes (entre l'année 1 et l'année 2).

En revanche, l'indice de volume de l'année 2, aux prix de l'année précédente, correspond bien à une évolution de volume entre l'année 1 et l'année 2 :

$$IndVol_2(\sum i) = \frac{\sum_i p_1(i) \times q_2(i)}{\sum_i p_1(i) \times q_1(i)}$$

L'idée des indices chaînés est alors de conserver comme évolution de volume l'indice de volume évalué aux prix de l'année précédente et de chaîner ces indices à partir d'une année de référence donnée.

Ainsi, si l'année de référence est l'année 0, le volume chaîné de l'année 2 est égal à :

$$vch_2(\sum i) = vol_0 \times IndVol_1 \times IndVol_2 = \left(\sum_i p_0(i) \times q_0(i) \right) * \frac{\sum_i p_0(i) \times q_1(i)}{\sum_i p_0(i) \times q_0(i)} * \frac{\sum_i p_1(i) \times q_2(i)}{\sum_i p_1(i) \times q_1(i)}$$

ce qui peut se simplifier en :

$$vch_2(\sum i) = \frac{\sum_i p_0(i) \times q_1(i)}{\sum_i p_1(i) \times q_1(i)} * \sum_i p_1(i) \times q_2(i)$$

Lorsqu'on chaîne les volumes à un niveau élémentaire, on retrouve le volume à prix constant, base année 0 :

$$vch_2(i) = \frac{p_0(i) \times q_1(i)}{p_1(i) \times q_1(i)} \times p_1(i) \times q_2(i) = p_0(i) \times q_2(i)$$

On peut alors constater en développant, ou en appliquant à un exemple numérique simple, que :

$$vch_2(i + j) \neq vch_2(i) + vch_2(j)$$

Autrement dit, la propriété d'additivité n'est plus respectée. Pour obtenir des volumes chaînés il faut alors chaîner les indices à tous les niveaux d'agrégation et pour toutes les opérations.

Inversement, on peut, à partir de volumes aux prix de l'année précédente, chaînés, reconstituer des volumes à prix constants en partant des volumes chaînés au niveau le plus fin possible.

Annexe 9 : Le calcul des contributions dans les comptes en volumes chaînés

Dans le cadre de volumes calculés aux prix constants de l'année de base, une même formule, relativement intuitive, permet de calculer la contribution d'une variable Y aux évolutions d'une variable X, en annuel comme en trimestriel, formalisée par la fonction « *Contrib* » suivante :

$$Contrib(Y_a^{vol\ prix\ ct}, X_a^{vol\ prix\ ct}) = ev(Y_a^{vol\ prix\ ct}) \frac{Y_{a-1}^{vol\ prix\ ct}}{X_{a-1}^{vol\ prix\ ct}}$$

$$Contrib(Y_t^{vol\ prix\ ct}, X_t^{vol\ prix\ ct}) = ev(Y_t^{vol\ prix\ ct}) \frac{Y_{t-1}^{vol\ prix\ ct}}{X_{t-1}^{vol\ prix\ ct}}$$

où $ev(Y)$ désigne le taux de croissance de la variable Y par rapport à la période précédente.

La contribution de la variable Y à l'évolution de la variable X est donc donnée par le taux de croissance de la variable Y pondéré par le poids de la variable Y dans la variable X en volume à la période précédente.

En annuel, la formule précédente reste valable pour les volumes chaînés moyennant une légère modification. La contribution de la variable Y à l'évolution de la variable X est maintenant donnée par le taux de croissance de la variable Y pondéré par le poids de la variable Y dans la variable X **en valeur** à la période précédente.

$$\overline{Contrib}_a(Y_a^{vol\ ch}, X_a^{vol\ ch}) = ev(Y_a^{vol\ ch}) \frac{Y_{a-1}^{valeur}}{X_{a-1}^{valeur}}$$

Le calcul de contributions trimestrielles est en revanche plus complexe et suppose de faire certains choix. En particulier, la simple transposition de la formule annuelle précédente n'a pas été considérée comme satisfaisante (bien qu'elle garantisse l'additivité des contributions des différentes composantes d'un agrégat) car elle implique un saut au premier trimestre (puisque, au premier trimestre, « l'année précédente » n'est plus la même qu'au trimestre précédent).

Les comptes trimestriels français retiennent donc pour le calcul des contributions une formule dérivée de la formule annuelle, mais adaptée de manière à supprimer les sauts observés à chaque premier trimestre, tout en préservant l'additivité des contributions des différents éléments constitutifs d'un agrégat. La formule générique, qui comprend donc un terme correctif spécifique pour les premiers trimestres, s'écrit :

$$\overline{Contrib}_{a,t}(Y_a^{vol\ ch}, X_a^{vol\ ch}) = \left\{ \begin{array}{l} contrib(Y_{a,t}^{vol\ ch}, X_{a,t}^{vol\ ch}) \\ + contrib(Y_{a,t}^{vol\ ch}, X_{a,t}^{vol\ ch}) \left(\frac{Y_{a-1}^{prix}}{X_{a-1}^{prix}} - 1 \right) \\ + 1_{t=1} \times \left(\frac{Y_{a-1,4}^{vol\ ch}}{X_{a-1,4}^{vol\ ch}} - \frac{Y_{a-1}^{vol\ ch}}{X_{a-1}^{vol\ ch}} \right) \times \left(\frac{Y_{a-1}^{prix}}{X_{a-1}^{prix}} - \frac{Y_{a-2}^{prix}}{X_{a-2}^{prix}} \right) \end{array} \right.$$

Le premier terme correspond au calcul « traditionnel » des contributions (et fait donc appel à la fonction *Contrib*). Le second terme introduit un correctif lié à la dérive des prix entre la composante Y et l'agrégat X. Le troisième terme intervient uniquement pour le calcul des contributions au 1^{er} trimestre, et corrige le saut éventuel lié au passage de la référence des prix (de l'année « a-2 » à l'année « a-1 »).

Cette formule est au premier abord complexe. Cet inconvénient mérite toutefois d'être nuancé : d'une part, elle se programme assez aisément¹⁵. D'autre part, la préservation de l'additivité fait qu'elle permet de calculer par solde la contribution de variables qui, étant tantôt négatives tantôt positives (comme les variations de stocks ou le solde des échanges extérieurs), ne sont pas définies en volumes chaînés. La contribution du solde des échanges extérieurs à l'évolution du PIB en volumes chaînés est ainsi définie comme la différence entre les contributions respectives des exportations et des importations en volumes chaînés, alors même que la variable de solde des échanges n'est pas définie en volumes chaînés.

¹⁵ Sur la rubrique « comptes trimestriels/méthodologie » du site internet de l'Insee, figurent une note détaillée explicitant la méthode de calcul des contributions en volumes chaînés, et un fichier Excel dans lequel la formule de calcul de ces contributions est pré-programmée à partir d'un exemple.