



**12ème colloque
de l'Association de comptabilité nationale
Paris, 4-6 juin 2008**

**Session n° 5
Modélisation et utilisation des comptes nationaux**

**L'élaboration des données de comptabilité nationale et l'analyse
économique**

(version préliminaire)

**Stéphane Gregoir
EDHEC Business School**

Résumé

La comptabilité nationale en mesurant des concepts harmonisés d'agrégats macroéconomiques vise à décrire au mieux la situation économique d'un pays en autorisant des comparaisons dans le temps et entre pays. L'observation conjointe d'un ensemble de variables permet de proposer des interprétations sur les mécanismes économiques à l'œuvre et des schémas explicatifs sur ce qui peut survenir ou sur les actions à mettre en œuvre pour réduire un déséquilibre observé. L'usage de modèles macroéconométriques structurels estimés sur ces données assoie ces interprétations.

Pour l'analyse de court terme, il est fait naturellement recours aux données trimestrielles de comptabilité nationale. Cependant la lisibilité des données infra-annuelles et de leurs enchaînements temporels n'est pas aisée et ce pour au moins deux raisons. D'une part, les règles conventionnelles de construction des données (moments d'enregistrement) introduisent dans les données des éléments de rationalité économique supposés ce qui affecte l'interprétation des équations structurelles de comportement des agents. D'autre part, les méthodes statistiques usuelles de correction de saisonnalité ou des effets calendaires en traitant individuellement les séries chronologiques peuvent introduire des distorsions dans les relations entre les variables. Ces deux dimensions se mêlent parfois lorsque l'élément de rationalité correspond à la répartition en cours d'année d'un revenu prévisible comme un treizième mois.

Travailler avec des données non corrigées et sans imputation des flux à la date du fait générateur mais en respectant les mouvements de trésorerie permettrait en théorie de décrire les décisions des agents et d'en faire une modélisation des comportements plus proche de la réalité. Les données brutes ne sont cependant pas toujours représentatives des phénomènes sous-jacents. Elles peuvent présenter des régularités « saisonnières » infra-annuelles liées à la technique de mesure plutôt qu'au phénomène économique ou peuvent ne pas rendre compte de telles fluctuations alors qu'elles sont présentes. Il en est par exemple pour le premier cas des déformations saisonnières de la consommation des ménages induites par le déplacement des populations lors des périodes de vacances et dans le second cas de la fluctuation des consommations intermédiaires en produits énergétiques déduites des productions à l'aide de coefficients techniques. L'interprétation des équations économétriques estimées sur de telles données garde donc un caractère de fragilité.

Au total, l'analyse économique doit s'appuyer sur des données qui sont naturellement le produit de conventions et qui sont affectées par la méthode d'élaboration. L'usage de données facilement lisibles est donc privilégié. Pour qu'il soit informatif, il importe de maintenir des traitements homogènes et stables dans le temps pour pouvoir détecter les modifications sous-jacentes et tenter de produire une analyse économique. En ce sens, la distorsion éventuelle des relations entre variables induites par les traitements statistiques de correction des variations saisonnières ou autres appliqués séparément sur les séries temporelles peut plus drastiquement affecter la qualité et la pertinence des analyses. Car même si le traitement est stable, les distorsions peuvent mener à des pertes de simultanéité ou de chronologie des enchaînements qui peuvent induire en erreur. Après avoir évoqué ces différentes questions, nous illustrons sur un exemple l'ampleur des modifications des relations temporelles entre des données désaisonnalisées et des données non traitées.

Summary

National accounts, by measuring macro-economic aggregates through harmonized concepts, aim at describing the economic situation of a country in a way that permits comparisons both through time and across countries. The joint observation of a set of variables allows the economists to produce interpretations in terms of economic mechanisms at work, with consequences regarding future changes or actions to be taken in order to reduce some currently observed disequilibrium. The use of structural macro-econometric models, estimated on such data, strengthens such interpretations.

Quarterly national accounts data are generally being used for short term economic analysis. However, reading infra-annual data is not always easy particularly regarding the timing of changes in related variables, for at least two reasons. On the one hand, conventional rules applied to compile data (time of recording rules) introduce in the data some part of presupposed economic rationality, which is affecting the interpretation to be given to structural equations in terms of agents' behaviors. On the other hand, usual statistical methods for seasonal and calendar adjustments, by treating time series one by one, may introduce some distortion on the relations between variables. These two dimensions can sometimes interact, for example when the rationality dimension is related to the expectation of the yearly distribution of a foreseeable part of income (like a bonus distributed once a year).

Working directly on non adjusted data, and on data recorded on cash basis rather than on an accrual basis may in theory allow the economists to describe and to model agents' decisions in a more accurate manner closer to reality. However, non adjusted data may not always be representative of underlying economic phenomena : in some cases, the sampling technique itself may induce seasonal regularities disconnected from economic reality (e.g. household consumption variations due to population movements during holidays), or on the contrary may erase some real economic seasonal regularities (e.g. intermediary consumption of energy of a given industry, measured through the hypothesis that the intermediary consumption/output ratio remains stable on an infra-annual basis). Therefore, the interpretation of econometric equations estimated on such data can remain fragile.

In the end, economic analysis always relies on data that are subject to conventions and that are affected by the technique through which they are elaborated. This is why the use of adjusted - hence easily readable - data is generally preferred. In order to be informative, homogeneous treatments must be maintained through time, so that underlying modifications may be detected, hence allowing for economic analysis. However, potential distortion between variables induced by specific statistical treatments can significantly affect the quality and the relevance of some analysis. Even when the treatment is stable through time, distortions may appear, inducing some loss in the simultaneity or the chronology of economic events, which then can produce misleading interpretations. After a theoretical review of all these questions, the paper provides some illustration on a practical case, where temporal relations between economic variables are substantively different whether they are considered on a seasonally adjusted or non seasonally adjusted basis.

Introduction

L'analyse macroéconomique quantitative repose fortement sur les données de comptabilité nationale qui sont construites en s'approchant au mieux de la grille d'analyse qui est la sienne. Cependant, ceci ne garantit pas que l'usage de ces données soit aisée. Nous nous proposons ici d'illustrer quelques uns des problèmes que l'on peut rencontrer lorsque l'on veut utiliser des données de comptabilité nationale dans un exercice d'analyse macro-économique quantitative. Ces difficultés sont de plusieurs natures et doivent être analysées en fonction des objectifs assignés à l'analyse elle-même. Aussi, dans une première partie, nous rappelons les finalités usuelles des travaux macroéconométriques et soulignons les limites en termes d'interprétation qui sont présentes à l'origine dans ces travaux. Puis nous illustrons comment les hypothèses explicites ou implicites retenues dans les conventions d'élaboration des données comptables et dans les traitements statistiques appliqués aux données (en général de comptabilité trimestrielle) interagissent avec ces objectifs, principalement au niveau de l'interprétation des résultats et de l'usage en prévision. Les difficultés rencontrées sont essentiellement de deux ordres, elles peuvent être comptables en termes conceptuel ou méthodologique, ou statistiques en termes de nature des mesures à traiter ou des traitements qu'il faut leur appliquer pour autoriser un usage aisé aux dépens parfois de l'interprétation. Ces problématiques sont riches, nous proposons de n'en aborder que quelques aspects. Il ressort de ces éclairages que le caractère conventionnel des données macro-économiques est primordial, en partie certes parce que la construction de la description d'une économie repose sur un corpus de sciences sociales qui définissent leur propre grille de lecture, mais aussi par les hypothèses techniques qui servent à l'élaboration de données interprétables. Un arbitrage apparaît entre la nécessité de produire des descriptions comparables dans le temps, et donc d'user d'un ensemble de règles ou de procédures pérennes, et celle de décrire en respectant la chronologie des faits et leur ampleur, un monde en perpétuelle évolution. Les comparaisons des amplitudes des phénomènes considérés dans les différents exercices macroéconométriques envisagés ne sont pas informatives en termes absolus, mais relativement à ce jeu de règles. L'analyse quantitative des phénomènes économiques demeure un art, les données de la comptabilité nationale constituent un matériau de base dont la manipulation demeure délicate. Il est important que les utilisateurs connaissent les modes d'élaboration des données pour produire des interprétations correctes. Transmettre les enseignements apportés par ces exercices repose sur une rhétorique qui se doit d'être rigoureuse dans son argumentation, les approximations et extensions de sens sont aisées.

Pour montrer l'importance des traitements statistiques effectués sur les données et les conséquences que cela peut avoir sur leur interprétation possible, nous illustrons dans une seconde partie l'influence d'un traitement statistique sur l'analyse des liens au cours du temps entre les variables macroéconomiques.

Les objectifs principaux d'une analyse quantitative macro-économique

L'analyse macro-économique quantitative a trois objectifs usuels :

- estimer les relations entre plusieurs grandeurs macroéconomiques sur la base d'une théorie et ainsi évaluer la pertinence de mécanismes économiques,
- prévoir l'évolution d'un ensemble de variables économiques,
- simuler les conséquences d'une modification de l'environnement des décisions des acteurs suite à une réforme de la politique fiscale ou monétaire ou une modification de la situation internationale

Ces objectifs pour chaque ensemble de variables macroéconomiques considéré nécessitent une modélisation particulière dans laquelle intervient un ensemble parfois plus large de variables afin de décrire correctement l'environnement du phénomène. Seules les variables dont l'influence est d'un ordre de grandeur significatif (en pratique cela repose souvent sur des *a priori*) sont naturellement considérées. La façon dont ses variables interagissent entre elles contraint l'usage du modèle. Mener à bien ces travaux nécessite d'appliquer des techniques économétriques plus ou moins sophistiquées. L'emploi de ces techniques suppose qu'un certain nombre d'hypothèses soient satisfaites qui

garantissent la qualité des résultats et la validité de l'usage du modèle et de l'interprétation des résultats. Ainsi, à titre d'exemple, si l'on souhaite faire une prévision de l'évolution d'une variable et que celle-ci dépende d'une seconde variable, cela nécessite de proposer un chemin d'évolution de la seconde. Ceci n'est correct que si sur l'horizon de prévision considéré, la première variable n'affecte pas elle-même le chemin que pourrait suivre la seconde. Dans le cas contraire pour faire une prévision, il faut modéliser de façon jointe les deux variables et calculer leur prévision dans ce cadre. L'économètre a recours à plusieurs notions d'exogénéité des variables par rapport aux paramètres et aux équations du modèle pour définir ces usages possibles.

L'interprétabilité des résultats repose certes sur ces hypothèses techniques, mais aussi sur le contenu informationnel attaché aux mesures utilisées et aux analogies formelles que le modélisateur est amené à faire pour produire un discours susceptible d'entraîner l'adhésion des auditeurs. Le cadre d'analyse macroéconométrique classique implique par construction quelques problèmes en cette matière.

1. Modélisation structurelle

Une première difficulté repose sur l'interprétation des équations dites de comportement des agents économiques que ce soit des ménages, des entreprises, des autorités fiscales ou monétaires, ou un sous-ensemble de ces populations. Il est usuellement fait usage d'un argument de comportement microéconomique pour sélectionner les variables qui y interviennent et spécifier la formule mathématique qui les lie. L'interprétation repose alors sur l'analogie avec le comportement d'un individu représentatif. La notion de rationalité au sens d'un usage efficace de l'information disponible compte tenu des modes d'action des agents et des mécanismes en jeu peut aussi être explicitement utilisée pour décrire le comportement de cet agent représentatif. Une difficulté d'interprétation des résultats est directement liée à cette analogie. Nous proposons de la décrire brièvement avant de considérer les limites associées à la définition et aux méthodes d'élaboration des agrégats. Elle est d'au moins trois ordres : du point de vue de l'interprétation en termes de bien-être, de la conséquence de modifications de règles de décisions de politiques économiques ou enfin du mode de formation des anticipations.

La théorie de l'agrégation économique étudie les conditions dans lesquelles l'agrégation de décisions individuelles peut être décrite par une fonction « objectif » d'un agent représentatif. Elle illustre en particulier les limites de l'interprétation des décisions agrégées en termes de préférence des agents [16]. Jerison [14] construit une économie telle qu'il existe un individu représentatif dont les décisions correspondent à l'agrégation des décisions individuelles mais dont les préférences ne sont pas cohérentes avec celles des agents. Ceux-ci peuvent tous simultanément préférer une situation A à une situation B alors que la fonction d'utilité agrégée cohérente avec l'agrégation des décisions individuelles conduit au choix de B par rapport à A. Il s'ensuit que l'interprétation en termes de bien-être des résultats d'une modélisation macroéconométrique est fragile.

La critique de Lucas [18] est célèbre pour avoir mis en cause l'usage des modèles macroéconomiques pour évaluer les conséquences d'une modification de politique économique. Elle se plaçait dans le cadre des anticipations rationnelles selon lequel les agents prennent leurs décisions en utilisant toutes les informations disponibles. Une modification d'une politique pouvait entraîner une modification de leur règle d'action et rendait inopérantes les relations censées la décrire mais établies sur les observations passées. Afin de répondre à cette critique, les recommandations avancées portaient sur la nécessaire distinction entre les paramètres de préférences des agents et de la technologie de l'économie et ceux qui gouvernaient les choix des autorités. La modélisation conjointe des différentes variables agrégées devenait plus compliquée. Le raisonnement qui sous-tend la critique de Lucas est fait sur la base du comportement d'un individu représentatif, ce qui simplifie la présentation de l'analyse de ce phénomène sans en réduire la pertinence. En théorie, la situation est plus complexe. Geweke [9] a montré que la fonction d'agrégation des comportements individuels pouvait elle-même être affectée par la modification de la règle de politique (cf. aussi [1]) et plusieurs travaux ont montré que l'agrégation des anticipations rationnelles des agents pouvait être différente de l'anticipation rationnelle de l'agrégat (entre autres [11], [19]). Des exercices de simulation ou des études empiriques ont illustré les conséquences quantitatives de ces différents éléments. La critique de Lucas a ainsi généré un développement d'outils et de techniques qui ont permis de montrer qu'elle avait une portée

empirique limitée compte tenu des spécifications et des données utilisées ([6]). De même, elle a soulevé la question de la modélisation d'une modification de politique par des autorités et donc de leur processus de décision pour pouvoir l'intégrer dans la modélisation macroéconomique ([23]). La tâche est difficile, dépend des institutions et n'a pas connu à ce jour à ma connaissance de traitement satisfaisant.

Il faut donc garder en tête que dans sa forme usuelle, l'interprétation d'un travail structurel sur données agrégées repose sur des analogies qui théoriquement ne sont correctes que dans un nombre réduit de situations. Néanmoins, la stabilité dans le temps des mécanismes et des mesures utilisées dans la modélisation peuvent permettre de proposer un discours utile en permettant de comparer des situations dans le temps ou avec des prévisions. Les mécanismes évoqués pour expliciter ces comparaisons peuvent être approximatif mais d'un point de vue rhétorique, permettre de faire appel à l'intuition des utilisateurs et gagner en pouvoir de conviction.

2. La modélisation statistique

Une autre critique importante faite au début des années quatre-vingt a sensiblement modifié les pratiques de l'analyse macroéconométrique, il s'agit de la critique de Sims [22]. Elle portait sur la difficulté dans de nombreuses situations d'identifier les paramètres des équations structurelles. Pour qu'un seul jeu de paramètres soit compatible avec un modèle structurel, il faut en général imposer un certain nombre de contraintes dont l'interprétation ou le caractère légitime sont difficiles à établir en particulier dans un système d'équations simultanées dynamiques de grande taille. Ces contraintes pouvaient de plus être en contradiction avec les propriétés statistiques conjointes des variables. Une stratégie de modélisation alternative a été recommandée. Elle consiste à utiliser au mieux les propriétés statistiques des variables dans une première étape (modèles VAR), puis à rechercher un schéma d'identification des chocs en ligne avec les conséquences attendues par la théorie économique (modèles VAR structurels). Les propriétés de prévision de ce type de modèles sont alors plus proches de celles des données statistiques. La lecture des prévisions ainsi produite est aussi plus difficile car les régularités statistiques passées exploitées au mieux ne permettent pas de produire un discours explicatif causal des prévisions elles-mêmes, ce à quoi s'attend en général l'utilisateur. L'analyse « causale » se réduit à une analyse de propagation de chocs et la difficulté se porte alors sur la définition des chocs qui ont un contenu économique en termes d'action nouvelle des agents économiques. Un choc structurel de politique monétaire ou fiscale est ainsi identifié par ses conséquences à court et long terme sur quelques variables clés. A titre d'exemple, un choc d'expansion monétaire est supposé à court terme se traduire par une augmentation de la masse monétaire et du niveau des prix et une baisse des taux d'intérêt nominaux et à long terme de ne pas avoir d'effet significatif sur les grandeurs réelles du modèle. Le contenu informationnel de ces chocs dépend de l'unité de temps considérée. Plus le pas est long, plus les effets cumulés d'un choc qui a eu lieu en début de période peut être différent de ceux d'un choc en fin de période. L'identification sur des données historiques en est affectée. Les modélisateurs ont donc naturellement essayé d'utiliser des données à haute fréquence, mais la plupart des pays diffusent au mieux des comptes trimestriels. Cette approche a permis de produire des prévisions qui ont fréquemment dominé en qualité celles issues de modélisation plus structurelles qui nécessitent de définir des scénarios d'évolution pour un grand nombre de variables d'environnement.

Les difficultés induites par les traitements des données de comptabilité nationale

1. Les difficultés d'ordre comptable

Nous allons ici illustrer comment la mise en œuvre pratique de choix de mesures de la comptabilité nationale peut affecter l'interprétation des résultats d'une estimation économétrique. Ceci résulte souvent d'un souci de lisibilité des données mais correspond à l'imposition de comportements des agents qui interagissent avec le cadre d'analyse économique ou à des lissages des données qui peuvent affecter l'amplitude des révisions et rendre fragile des exercices de prévision. Nous aborderons la question du moment d'enregistrement et des différents systèmes de prix, mais il existe

d'autres situations pour lesquelles les mêmes considérations peuvent être faites. Il en est ainsi par exemple du traitement des investissements directs à l'étranger.

Ces difficultés diffèrent selon la fréquence d'observation. L'annualité des exercices comptables des entreprises et de l'Etat et de la plupart des sources statistiques qui leur sont associées rendent les choses plus aisées pour les travaux à pas annuel. La situation est plus compliquée pour les données infra-annuelles qui sont aujourd'hui plus souvent utilisées. Les règles et les concepts utilisés dans l'élaboration des comptes annuels sont adaptés *a minima* au cadre trimestriel si l'on désire assurer une cohérence comptable entre les mesures annuelles et les mesures infra-annuelles. Par exemple, la définition de l'investissement qui repose explicitement sur l'emploi d'un objet pendant une durée supérieure à un an ne peut être modifiée dans un exercice infra-annuel simplement parce que l'unité de temps est différente. Implicitement on vise à construire des évaluations de mesures qui sont pleinement interprétables dans le cadre des évaluations annuelles. L'analyse au cours de l'année des interactions entre différentes variables au sein d'un mécanisme économique peut ainsi être rendue plus délicate.

1.1. Le moment d'enregistrement et la description économique que l'on veut produire

1.1.1. Les problèmes liés au moment d'enregistrement

Lorsque l'on traite de phénomènes économiques (annuels ou infra-annuelles), la datation des opérations et le système de valorisation qui leur sont associés sont très importants pour décrire les ensembles d'information disponibles des agents et modéliser leurs règles de décision conditionnelles à ces ensembles. La comptabilité nationale repose sur une règle générale dont l'usage n'est pas immédiat. Selon les circonstances, une interprétation est nécessaire, mais il convient de tenter de rester le plus neutre possible et de ne pas introduire d'hypothèses liées aux comportements des agents issus d'*a priori* théoriques, tout en garantissant une description la plus exacte possible des flux et en assurant la cohérence entre les différentes approches possibles.

La règle retenue dans le cadre du SEC95 en matière de datation des opérations est la règle décrite par les termes anglais « *accrual basis* » traduits de façon restrictive dans le SEC95 par l'expression « droits constatés ». Son usage en comptabilité annuelle française n'est pas aisé, laisse parfois une large place à l'interprétation, et n'est pas toujours opératoire au vu des pratiques institutionnelles ou des informations dont on dispose. Les comptables français préfèrent en fait traduire l'expression « *accrual basis* » par l'expression « droits et obligations » qui se rapporte au moment où les dettes ou créances sont certaines. De fait, ceci recouvre, suivant les situations, les notions (i) de fait générateur, (ii) de droits constatés ou encore (iii) de moment de paiement. Suivant la nature des statistiques et des règles comptables observées par les différents agents, il sera plus ou moins aisé d'utiliser une des approches pour construire des mesures compréhensibles et interprétables par l'utilisateur. Quelques exemples vont illustrer ce propos.

Si nous regardons le cas de l'impôt sur le revenu, ce dernier devrait en toute logique suivant le principe du fait générateur être attribué à l'année de versement du revenu. Le décalage en France d'un an rend l'opération difficile. L'application stricte de la règle rendrait l'interprétation des comportements des agents assez abstraite, pour ne pas dire plus. De plus, le montant d'impôt à payer n'est en général connu qu'en début d'année où il est versé du fait de modifications régulières de l'assiette ou de règles de comptabilisation. Des révisions artificielles en résulteraient. Comme, il correspond à l'émission par l'Etat d'un rôle, les termes « droits constatés » permettent de fait de mieux décrire cette situation et garantir l'interprétabilité d'équation de comportement des ménages en ligne avec la réalité perçue.

De fait, s'éloigner des mesures directes implique l'introduction de corrections qui peuvent correspondre à des hypothèses implicites sur le comportement des agents et rendre l'analyse statistique ou économétrique plus délicate. Si les données sont rendues plus facilement lisibles du point de vue de l'utilisateur dans le sens où leur interprétation ne nécessite pas de connaissance détaillées des principes de la comptabilité nationale car un travail d'homogénéisation a été effectué, elles peuvent être plus difficilement utilisables dans des travaux de modélisation et d'analyse économétriques. Nous illustrons les conséquences de ces corrections des mesures de base dans deux cas : d'une part, le problème de la prise en compte de la production de produits qui nécessite un temps supérieur à l'unité de temps que

l'on considère implicitement et d'autre part, le problème des versements ou prélèvements irréguliers de revenu aux ménages.

1.1.2. Les productions à processus de production « long »

Lorsqu'un produit nécessite pour être réalisé un laps de temps correspondant à plusieurs périodes d'observation¹, la comptabilité nationale cherche à décrire la part de production effectuée chaque période à travers la mesure d'une production en cours qui se trouve en stock de produits en cours. Le problème auquel nous nous trouvons confrontés est de définir la valeur de ce bien et d'y associer les générations de flux de revenus pertinents. L'interprétation économique des grandeurs mesurées et leur chronologie peuvent être affectées. De même, l'usage de ces données dans un exercice de prévision peut se révéler délicat.

Deux cas peuvent être distingués. Cette production élaborée peut résulter d'une commande et donner lieu à paiement d'acomptes. Dans ce cas, la valeur de l'objet est connue et il est possible de valoriser l'immobilisation en stock en fonction de cette information et de construire un flux de revenus associés. L'équilibre comptable décrit par l'ensemble de ces données autorise une interprétation et il y a cohérence entre les différentes approches, que ce soit par la production, par les emplois ou par le revenu généré (du fait de l'existence de paiements partiels). En revanche, la situation est plus complexe lorsque l'on considère des processus de production dont on ne connaît pas forcément l'output, dans la mesure où il peut être affecté par un aléa ou pour lequel il n'y a pas prépaiement ou paiement partiel. Ceci décrit assez bien la production agricole² qui nous servira d'exemple pour illustrer que des difficultés d'interprétation et d'usage des mesures peuvent résulter des conventions utilisées.

L'approche préconisée par le SEC95 est de valoriser la production en cours d'une période en fonction de la somme des coûts induits corrigée d'un taux de marge (à définir) de façon à générer une valeur ajoutée associée au fait générateur de produire et un excédent brut d'exploitation dans la mesure où des salaires sont versés et des taxes sur les produits peuvent être perçues par les administrations. Ceci pose problème pour deux raisons. D'une part, nous ne connaissons pas le taux de marge qu'il faut utiliser et ce taux connaît des évolutions conjoncturelles qu'il importe de capturer si l'on veut décrire avec précision les évolutions et répartitions des rémunérations des facteurs. Le comptable en faisant choix d'une valeur particulière impose un comportement artificiel aux producteurs. D'autre part, le volume physique de ces productions demeure aléatoire. Un phénomène météorologique peut détruire tout ou partie d'une récolte (et jouer sur les prix et les marges). Le revenu des producteurs aurait donc été imputé de façon discutable lors des périodes qui ont précédé la destruction totale ou partielle de la production. Suivant le caractère exceptionnel ou non de cet aléa, le SEC95 propose des traitements différents. S'il s'agit d'un aléa régulier, le SEC propose de corriger ces flux à travers une variation de stock des produits en cours du montant concerné. Une valeur ajoutée a donc été créée puis annulée sans que ceci ait correspondu à une création transitoire effective de richesse. La cohérence des approches implique un mouvement marqué de l'épargne brute. Cependant, la nature de l'épargne dégagée par cette activité n'est pas de même nature que celle tirée d'activité « sûre ». Les confondre peut amener à des interprétations artificielles des comportements et des hypothèses a priori sur l'attitude des agents face au risque. S'il s'agit d'un aléa exceptionnel (catastrophe naturelle, parasite), le SEC propose de corriger ces flux à travers une variation du poste « autres changements de volume » du montant concerné, ce qui n'affecte pas la génération de revenu brut de la période. L'élaboration de ces données n'aide pas à l'interprétation des évolutions infra-annuelles et à la compréhension de prévisions qui ne sont pas révisées malgré la connaissance de la survenue de l'événement exceptionnel...Le modélisateur doit disposer de l'information détaillée pour corriger les flux et les aligner en cohérence avec une interprétation économique.

Les difficultés rencontrées ne sont pas de même nature selon l'unité de temps que l'on considère. A pas annuel, la production agricole ne pose pas de problème particulier dans l'hémisphère nord. Les semailles et la récolte ont lieu la même année. Cette dernière peut être excellente ou faible.

¹ C'est par exemple, le cas de la construction d'avions, de bateaux, de machines outils complexes,....mais l'on peut aussi considérer les réalisations de progiciels par des sociétés de services ou la production agricole.

² Il faut noter toutefois que l'année civile ne correspond pas toujours à l'année agricole.

L'évaluation ex-post qui sera réalisée en fin d'année par la comptabilité annuelle aura sa propre cohérence, l'aléa aura été observé. A pas infra-annuel, les conséquences des conventions sont plus lourdes et peuvent affecter la qualité d'exercice de prévision. Une alternative prudente consiste à valoriser sans marges les productions en cours suivant la valeur des coûts nécessaires à leur réalisation. De sorte que la valeur ajoutée soit générée lors de la réalisation de la récolte. Ceci implique néanmoins un profil contrasté au cours de l'année qu'il faut traiter à l'aide de procédures statistiques pour en autoriser une interprétation conjoncturelle. Ces mesures s'accompagnent de révisions au fur et à mesure que l'information devient disponible sur la réalisation des aléas.

1.1.3. Les versements ou prélèvements irréguliers de revenus

Les enjeux sont d'une nature différente lorsque l'on veut traiter des versements à des dates espacées de revenus de montants irréguliers auxquels sont confrontés les ménages. Là encore l'interprétabilité et la lisibilité des comptes peuvent être affectées par les conventions en usage.

Nous considérons dans un premier temps, le cas d'un revenu connu mais versé de façon irrégulière par rapport à l'unité de temps considérée. Sur la base du principe du fait générateur, il est relativement légitime d'attribuer aux ménages qui reçoivent un complément de salaire du type « treizième mois », un « douzième »³ du versement mensuel sur toute la période pendant laquelle ils y ont droit. Ceci est possible car la somme qui leur est versée est contractuelle. Il faut néanmoins noter que cela implique un traitement des données de base et altère la relation entre l'épargne des ménages et celle des entreprises telles qu'elles peuvent être décrites dans les comptes des agents institutionnels de la comptabilité nationale et ce qui peut par ailleurs être observé dans des comptes financiers. On attribue aux ménages qui en bénéficient un revenu qu'ils n'ont pas. Alors qu'ils peuvent rationnellement réduire leur effort d'épargne en début d'année du fait de ce sursalaire anticipé en fin d'année, on maintient leur niveau d'épargne constant. Par souci de simplicité d'usage et de lisibilité, on élimine ainsi des données une composante du comportement des ménages qui pourrait se révéler informative pour une estimation précise de l'équation de consommation.

Le problème couvre plus des dimensions voisines lorsque l'on considère le cas des paiements de primes de fin d'année fonction de l'activité de l'entreprise pendant l'année ou des décisions politiques relatives à une mesure exceptionnelle touchant les prestations ou les impôts⁴.

Dans le premier cas, le problème est lié à la nature de la description des flux de revenus que l'on veut construire et à l'ampleur des révisions que le comptable national accepte. Le principe d'une prime est acquis et correspond à la réalisation d'un travail pendant l'année écoulée, mais son montant reste à la discrétion de l'employeur et de son appréciation de la performance de l'employé et de la situation de l'entreprise. Si aucun traitement n'est fait a priori sur les données de base, le profil infra-annuel des revenus présente un accroissement marqué mais irrégulier en fin d'année. L'utilisation d'une procédure statistique (de type désaisonnalisation) peut être utilisée pour permettre une interprétation « correcte » de l'évolution de la variable, mais ne permet pas de traiter de manière satisfaisante cette déformation régulière dans la mesure où sa taille dépend de l'activité et peut présenter un caractère cyclique. Il peut néanmoins sembler prudent d'effectuer ce traitement à chaque publication afin de traiter progressivement les informations disponibles. Des révisions des évaluations s'en suivent. Une autre approche possible consiste à construire une estimation des montants attendus au cours de l'année et à attribuer en fin d'année l'écart entre le montant anticipé et le montant observé. En termes d'interprétation, la lisibilité des données est plus grande mais ceci correspond encore à une hypothèse sur le comportement des agents face à l'incertain et leur comportement d'épargne. Dans cette situation, l'analyse statistique ou économétrique des données doit alors plutôt s'effectuer sur les données non traitées (une base caisse) si elles sont disponibles, ce qui est rarement possible.

Dans le second cas, il faut définir l'interprétation du principe du fait générateur. Le fait générateur est la décision exceptionnelle du gouvernement d'introduire une mesure nouvelle sur une période donnée, éventuellement rétrospective. Sa prise en compte peut prendre plusieurs formes selon le calendrier de l'annonce et de la mesure. Considérons le cas d'un versement annuel unique dont le

³ ou au prorata temporis de leur présence dans l'entreprise.

⁴ Nous pouvons penser par exemple aux primes annuelles des personnes employées dans les banques (et plus particulièrement dans les salles de marché).

calendrier est connu (dans la seconde partie de l'année) mais dont l'annonce de l'augmentation exceptionnelle est faite quelques mois avant le versement. Si l'on considère que l'annonce n'est pas liée au fait générateur et que par souci de lisibilité des comptes on répartit usuellement le montant sur toute l'année, il est possible de réviser après l'annonce gouvernementale le montant versé et de le répartir sur l'année en cours. Il est aussi possible si l'on considère que l'annonce affecte le fait générateur, d'introduire une révision de la somme infra-annuelle versée seulement à partir de la date de l'annonce publique. On introduit ainsi une forme de rationalité dans le comportement des agents qui respecte la chronologie de l'ensemble d'information disponible, ce qui n'est pas pris en compte dans le premier traitement, mais correspond aussi à une hypothèse sur leur décision en situation incertaine. Enfin, il est possible de ne faire intervenir le montant qu'au moment de la date officielle de versement. Les agents ont pu modifier leur décision au moment de l'annonce en fonction du montant, l'épargne des ménages révèle leur arbitrage intertemporel en fonction de leur aversion pour le risque. L'analyse conjoncturelle et les exercices de prévision que l'on sera amené à faire devront prendre en compte explicitement ce traitement. De même, si l'on cherche à estimer des comportements économiques sur des données à haute fréquence. Il importe d'une part que cet aspect soit connu des modélisateurs, d'autre part que la règle retenue soit pérenne.

D'une façon générale, on voit que lorsque l'on procède à des « lissages » ex-ante des revenus irréguliers, on impose aux données un profil qui peut se traduire en termes de comportement des agents. Ceci affecte naturellement les estimations d'équation de comportement et rend leur usage délicat pour des analyses de modification de politique ou des exercices de prévision. Ces règles d'élaboration des données doivent être connues des utilisateurs.

1.2. Les mesures dans différents systèmes de prix

Afin de rendre les évolutions de grandeurs macro-économiques comparables entre deux périodes différentes, la comptabilité nationale propose d'effectuer des mesures dans un système de prix de référence. Lorsque l'on se trouve à une date trop éloignée de cette période de référence, la pertinence de la mesure peut être mise en question. Afin de capturer à la fois les déformations de la structure des prix relatifs et l'évolution des volumes, il est proposé de recourir à l'emploi d'indices chaînés de chaque poste comptable. Ce chaînage se fait en comptabilité annuelle en utilisant le système des prix de l'année précédente. Une de ses conséquences est de produire des grandeurs qui ne satisfont plus les équilibres comptables.

Un problème surgit dès que l'on veut envisager le même type de mesure pour des données infra-annuelles qui soient cohérentes avec les données annuelles par ailleurs disponibles. Ceci signifie qu'il faut pratiquer des évaluations des agrégats des périodes d'observation (mois, trimestre,...) d'une année donnée aux prix moyens de l'année précédente. L'interprétation des évolutions des agrégats en volume chaque début d'année peut être affectée dans la mesure où elles mêlent à la fois une modification du rythme d'évolution de l'activité et un changement de système de prix. Dans la pratique, les conséquences de ces modifications régulières de système de prix semblent négligeables.

L'usage des données chaînées dans des modèles macroéconométriques classiques qui reposent sur l'emploi d'un grand nombre d'équations comptables nécessite un traitement adéquat. Les équations comptables doivent être complétées d'un terme d'ajustement dû au changement de système de prix. La modélisation de ces termes est a priori peu simple. Ils reposent sur la structure des prix des agrégats impliqués dans l'équation comptable et sont donc liés entre différentes équations. A court terme, compte tenu de l'amplitude des effets observés, leurs conséquences sont vraisemblablement faibles. En revanche, lorsque l'horizon de prévision est long, exercice de moyen ou long terme (10 à 20 ans), des erreurs sur leur spécification peuvent avoir des conséquences lourdes. Il est certes légitime de rappeler que l'élaboration d'un scénario de moyen ou long terme dans un système de prix fixe ancien perdait aussi en pertinence. Il importe dans les deux situations de connaître les limites de l'exercice. Il faut garder en tête qu'*a priori* le compte central de ces projections est principalement affecté. Les évaluations des effets de différentes politiques économiques en écart à ce compte restent exploitables si leur ampleur reste modérée.

2. Les difficultés liées au traitement statistique

2.1. Des statistiques de base parfois lacunaires

La principale difficulté statistique rencontrée dans l'élaboration d'un système de comptabilité nationale à fréquence trimestrielle consiste en le volume plus faible de données disponibles. Les données ne couvrent parfois qu'une partie du champ désiré, dans une nomenclature éventuellement différente, avec des règles de moment d'enregistrement différentes,...Ceci entraîne que le niveau de détail d'élaboration des données et de leur diffusion est plus réduit. Des approximations de concept et de champ sont nécessaires. Il faut en contrôler les conséquences en termes d'interprétation. Là encore, la chronologie des phénomènes économiques peut être modifiée, ce qui peut se traduire en modification des comportements des agents. L'analyse conjoncturelle et l'interprétation des équations et des prévisions sont a priori plus affectées que l'estimation d'équations de comportement si ces approximations de concept ou de champ sont pérennes.

Par exemple, il est possible de mesurer la consommation des ménages en électricité en fonction de leurs dépenses effectives lorsque l'on ne possède pas d'information directe sur leur consommation. Les grandes entreprises qui produisent l'électricité offrent en général aux ménages la possibilité de payer leur consommation par acompte régulier au cours de l'année, un ajustement étant opéré à la fin de la période en fonction de l'observation de la consommation annuelle effective observée. Ceci n'est pas neutre sur l'interprétation que l'on peut faire de l'évolution de court terme du taux d'épargne au cours de l'année en particulier lors de périodes de conditions météorologiques exceptionnelles. Il en est de même lorsque l'on construit une mesure de la consommation des ménages en automobiles ou en produits pharmaceutiques en s'appuyant sur des statistiques basées sur des déclarations d'ordre administratif. La datation des phénomènes peut être affectée des pratiques de gestion des organismes en charge des enregistrements. Lorsque l'information est disponible, des corrections assises sur des observations doivent être apportées aux mesures collectées. Enfin, si l'on applique des mesures annuelles sur des phénomènes infra-annuels, il est possible d'induire des modifications dans les équilibres comptables qui peuvent affecter l'évolution infra-annuelle du PIB. Ainsi, lorsque l'on utilise les coefficients techniques annuels d'un tableau d'emplois intermédiaires en consommation d'énergie, on ne respecte pas le profil infra-annuel de cette consommation qui est affectée par les conditions climatiques. Il y a une sous-évaluation de la valeur ajoutée les trimestres où l'on ne chauffe pas les ateliers et les bureaux et une surévaluation dans le cas contraire.

2.2. Nature et traitement de la variabilité des données

2.2.1. Nature de la variabilité

Lorsque l'on considère des périodes courtes, les mesures des phénomènes économiques peuvent être affectées par des aléas de natures diverses qui peuvent se compenser lorsque l'on établit des observations à des fréquences plus faibles. Ces aléas sont par exemple liés aux saisons et aux conditions météorologiques qui leur sont en moyenne associées, à la structure des jours ouvrés dans la période considérée, à la présence de périodes de vacances scolaires et de jours fériés,... Cette variabilité représente en général la part principale de la variabilité de court terme des grandeurs macro-économiques usuellement regardées à fin d'analyse conjoncturelle. Il est alors nécessaire d'aider l'utilisateur dans sa lecture en lui proposant une mesure corrigée de ces sources de variabilité.

Il faut tout d'abord rappeler que la variabilité infra-annuelle que l'on mesure n'est pas forcément celle de la variable réelle que l'on cherche à mesurer mais peut être le résultat de la méthode d'élaboration des données. Ainsi, si l'on étudie la consommation des ménages en produits alimentaires, on observe une chute lors des mois de vacances estivales. Ceci peut résulter non pas d'une baisse de la consommation alimentaire des ménages mais de leur départ en vacances vers des lieux où les magasins qui font partie de l'échantillon utilisé pour la collecte des données tout au long de l'année sont pondérés par leur poids moyen annuel et non leurs poids représentatif estival qui lui sont supérieurs. Le plan de sondage n'essaie pas de capturer ces déformations dans la mesure où il est difficile d'évaluer l'évolution infra-annuelle de la densité moyenne des personnes présentes dans ces

sites et qu'il faudrait utiliser un système de pondération variant avec la date d'observation. En redressant les données collectées chaque mois par les chiffres d'affaires annuels, dans la mesure où les magasins sondés qui connaissent des fluctuations marquées de leur activité liées à des déplacements saisonniers de la clientèle représentent une part minoritaire du poids total, on surévalue la consommation lors des mois hors période de migration touristique et on sous-évalue celle réalisée en été⁵. Dans cette circonstance, il est important d'effectuer un traitement des données afin d'éviter des interprétations erronées des évolutions décrites sur données brutes. Ce traitement des données sera d'autant moins difficile à réaliser que la manière d'élaborer les données sera toujours la même au cours du temps et les déformations infra-annuelles régulières.

De manière générale, on est amené à traiter deux types de variabilité « régulière » qui nuisent à l'interprétabilité des données. Ce sont les variations saisonnières et les variations des nombres et types de jours présents dans chaque période de temps considérée. En pratique, les deux corrections sont liées dans la mesure où la structure des jours et leurs nombres présentent des régularités saisonnières.

2.2.2. Les traitements des corrections des variations saisonnières

Les méthodes de corrections des variations saisonnières reposent usuellement sur la décomposition des variables à traiter en la somme de quatre composantes : la tendance qui capture les évolutions sur le long terme, le cycle qui reproduit les évolutions de moyen terme, la composante saisonnière et une composante irrégulière composée de chocs transitoires sans persistance. La série C.V.S. est égale à la série précédente corrigée de la composante saisonnière. Il faut donc procéder à l'estimation de cette dernière. Cette approche descriptive peut donner des résultats en contradiction avec ceux qui seraient déduits du comportement d'un agent face à un phénomène saisonnier dans un modèle structurel ([9]). Deux approches descriptives sont principalement employées. Nous les décrivons brièvement et soulignons leurs limites.

D'une part, on trouve les méthodes non paramétriques qui proposent l'usage d'une famille restreinte de moyennes mobiles centrées. Le fait que ces moyennes soient centrées pose la question du traitement des débuts et fins de série dans la mesure où on ne dispose pas d'assez d'observations pour les appliquer. Ces méthodes ont alors recours à des modélisations parcimonieuses pour construire des prévisions et des rétrovisions de la série et appliquer à ces ensembles de données les filtres précédents. En fait le traitement est légèrement plus complexe, car il importe d'isoler et de traiter proprement des points aberrants qui pourraient polluer localement les calculs. La définition de ces points est fonction de la variabilité hors composante saisonnière de la série. En pratique, on observe des problèmes de stabilité des estimations des modèles utilisés dans le calcul des provisions et rétrovisions.

D'autre part, une autre approche plus paramétrique est possible. Elle repose sur l'estimation d'un modèle de type SARIMA sur la série, puis la paramétrisation de la densité spectrale de chacune des composantes évoquées précédemment (à savoir, tendance, cycle, saisonnalité et irrégulier) à partir de cette estimation et enfin l'usage d'un algorithme d'extraction de signal appliqué à la décomposition produite. Elle induit le calcul d'un filtre infini optimal en un certain sens, calculé à partir de l'estimation de la modélisation retenue de la série. Le filtre est adapté à chaque série et diffère donc pour chaque série. En pratique, des instabilités dans la paramétrisation ou dans la forme du modèle SARIMA retenu peut entraîner une instabilité des estimations des différentes composantes.

Deux questions sont liées à l'usage et aux conséquences de ce traitement. La première porte sur le niveau auquel et la fréquence à laquelle le traitement doit être appliqué pour produire des données faiblement révisées du fait de l'unique traitement de la saisonnalité et sans saisonnalité résiduelle qui induirait en erreur l'utilisateur. Le statisticien a parfois le choix de traiter les données disponibles à une fréquence plus haute puis de les agréger dans le temps ou de traiter séparément les composantes d'un agrégat plutôt que l'agrégat lui-même. Il peut aussi effectuer un traitement à toutes les publications ou effectuer une campagne de C.V.S. une fois l'an. Il n'existe pas choix optimal, il dépend des propriétés des données. Il s'ensuit que le producteur d'un grand nombre de données ne pouvant adapter sa pratique à chaque situation sera amené à publier des données de qualité variable susceptibles de plus ou moins grandes révisions qui compliqueront le travail du conjoncturiste. La

⁵ On peut aussi introduire une erreur de mesure sur le taux de croissance annuel moyen.

deuxième question porte sur la déformation des liens temporels qui existent entre plusieurs variables macroéconomiques. En effet, les méthodes rapidement décrites ci-dessus sont des méthodes univariées. Elles induisent en général l'emploi de filtres différents pour chaque série. Des travaux ([21], [24] et [25]) ont montré que l'emploi de filtres différents pour deux variables liées par une relation dynamique linéaire induisait un biais dans son estimation. Chaque donnée corrigée des variations saisonnières étant une combinaison linéaire de données passées et futures, on comprend intuitivement que lorsque les données ne sont pas modifiées de manière similaire, la relation qui les lie change de forme. Or ce sont ces liens aussi bien dans l'analyse conjoncturelle que dans l'estimation de modèles qui sont exploités pour construire des interprétations et faire des comparaisons dans le temps ou entre pays. Ces liens artificiels sont par ailleurs susceptibles d'être modifiés à chaque changement de filtre de désaisonnalisation. Le traitement de conjoint de la saisonnalité de plusieurs variables a été envisagé par Geweke [8] dans un cadre contraint qui n'a pas connu d'application. Gregoir [13] a proposé un cadre plus souple qui repose sur des filtres asymétriques multidimensionnels. Mais dans les deux cas, le nombre de séries qui peuvent être traitées simultanément est limité. Du point de vue de l'utilisateur, ces deux points soulèvent une difficulté quant à l'usage des données pour une analyse économique. Il faut bien entendu pouvoir évaluer les conséquences empiriques et pratiques de ces défauts. La diffusion de données brutes, lorsque cela est possible, permettrait néanmoins à chaque utilisateur expérimenté de faire choix des traitements qu'il souhaite appliquer. Pour le consommateur de base, la nécessité de produire des données aisées à lire ne permet pas en l'état actuel des connaissances de se passer des données désaisonnalisées de manière usuelle.

2.2.3. Les traitements des corrections des effets de calendrier.

Les effets liés aux nombres et à la nature des jours présents pendant la période de mesure peuvent être sensibles et entraîner des erreurs d'interprétation sur la tendance sous-jacente de l'activité s'ils ne sont pas pris en compte explicitement. Le traitement de la variabilité due au nombre et à la structure des différents jours présents dans la période de mesure se fait principalement suivant deux méthodes : une première itérative qui mêle désaisonnalisation et correction des effets de calendrier et une seconde paramétrique, préliminaire au traitement de désaisonnalisation, qui permet d'étudier la significativité et la stabilité des effets. Cette dernière est à privilégier. De même que pour le traitement des variations saisonnières, la question se pose de savoir sur quelle fréquence et quel niveau d'agrégation, il est le plus approprié d'effectuer le traitement. Les données corrigées des effets de calendrier servent à l'analyse conjoncturelle. Elles ne peuvent entrer dans une estimation d'une équation structurelle de comportement, ne serait-ce que parce que les postes des comptes de secteur n'ont pas vocation à être corrigés des effets de calendrier et ne sont donc pas homogène aux volumes corrigés.

2.2.4. Le traitement des points aberrants

L'influence des points aberrants sur les traitements évoqués ci-dessus peut être très grande, aussi les procédures statistiques considérées comprennent une étape de détection des points aberrants et des modifications persistantes. Il existe un degré de liberté assez grand dans le choix des critères statistiques utilisés pour identifier ces points exceptionnels, ce qui peut produire des résultats assez éloignés lorsque le traitement est fait en mode automatique sans information exogène.

Il arrive que des événements exceptionnels se produisent au même moment dans l'année deux ou trois années pas nécessairement consécutives. Ceci s'observe en pratique pour certains mouvements sociaux. Ne pas traiter a priori ces événements comme des points aberrants peut faire passer ce qui est dû à une cause exogène comme un nouveau profil saisonnier. De manière générale, il importe de traiter a priori les points aberrants en fonction de l'information exogène disponible de façon à réduire et contrôler leur impact sur les estimations. En pratique, cependant, lorsque le nombre de séries à traiter est élevé, une telle tâche peut apparaître irréalisable.

De façon plus critique, le problème soulevé par le caractère univarié du traitement de la saisonnalité se pose de façon plus aiguë pour le traitement des points aberrants. Il existe de nombreuses approches de l'analyse de la causalité entre variables économiques. La plus connue

introduite par Granger [12] assimile la causalité à une régularité en prédiction (voir [16]). Une variable cause au sens de Granger une autre variable si elle permet d'améliorer le pouvoir prédictif d'un modèle autoprojectif de cette dernière. Cette propriété peut résulter de l'existence d'une troisième variable qui cause les deux premières, si bien que cette notion de causalité est une forme de corrélation dynamique satisfaite en moyenne. Une approche alternative considère que ce sont les événements exceptionnels qui permettent d'identifier les liens causaux entre variables et les modes de propagation, dans la mesure où les causes de ces valeurs exceptionnelles sont parfois connues. Lorsque les données sont traitées de manière séparée de leurs points aberrants, cette information peut être déformée en termes d'ampleur et de chronologie. L'estimation d'une équation les reliant peut être biaisée.

Que fait-on lorsque l'on réalise une étude quantitative sur la base de données de comptabilité nationale ?

A la lecture de ces différents exemples, une question peut naturellement venir à l'esprit. Compte tenu de toutes ses dimensions, quelle est la nature d'un exercice d'analyse quantitative macroéconomique ? Cette question couvre deux domaines, d'une part le domaine technique et d'autre part le domaine de l'interprétation et du discours qu'un tel travail autorise.

D'un point de vue technique, la démarche statistique mise en œuvre dans un travail macroéconométrique correspond à un cadre de travail précis. Des hypothèses lui sont associées qui, lorsqu'elles sont satisfaites, définissent l'usage des résultats. La démarche aujourd'hui privilégiée est d'utiliser au mieux les propriétés statistiques jointes des variables modélisées. L'objectif est d'aboutir à une modélisation parcimonieuse qui puisse être disqualifiée aisément. Ceci n'est possible que pour des systèmes d'équations de petite taille. Il peut être coûteux de développer un grand nombre de modèles de taille restreinte, chacun adapté à l'étude d'une question particulière. L'estimation d'un grand système d'équations dynamiques peut alors apparaître raisonnable pour traiter plusieurs questions. Il est cependant dans ce cadre plus difficile de respecter scrupuleusement les bonnes pratiques. Des hypothèses a priori et des pertes d'efficacité réduisent la qualité du modèle. Qu'il s'agisse de modèles structurels ou de modèles moins contraints *a priori*, on comprend à l'issue de la section 2 de la partie précédente, qu'il peut être souhaitable pour respecter et exploiter au mieux les liens dynamiques entre les variables de travailler sur des données brutes sans aucune correction, même si les données obtenues peuvent présenter des propriétés dynamiques artificielles liées au mode d'élaboration statistique plutôt qu'aux propriétés de la variable mesurée. La contrepartie de ce choix est d'une part que le modèle est plus compliqué et en général mobilise plus de paramètres, et d'autre part, qu'il faut produire avec ce modèle un discours compréhensible par tous, *i.e.* corrigé des régularités saisonnières. Un traitement doit donc être réalisé *ex post* pour présenter sous la forme usuelle les résultats. Aussi peu de personnes travaillent sur données non traitées. Lorsque les données traitées de la saisonnalité sont utilisées, il importe que ces traitements soient stables dans le temps au risque d'aboutir à des estimations et des résultats instables. Néanmoins, il semble souhaitable dans tous les cas d'utiliser l'information la plus fraîche et les techniques les plus performantes.

L'interprétation des résultats et la possibilité de communiquer aisément sur leur contenu sont des points plus délicats. Ils couvrent plusieurs dimensions liées à la crédibilité des données produites et à la possibilité d'un discours rigoureux et pédagogique. Nous considérons dans un premier temps ce second point. Les exemples évoqués dans la section 1 de la partie précédente semblent indiquer que le contenu informationnel des données de comptabilité nationale n'est pas neutre en termes de comportements économiques. Le système de mesure part d'une réalité qu'il segmente, code, échantillonne et estime. Il propose une mesure d'un ensemble de concepts économiques mais est amené afin de rendre les données simples d'usage à effectuer des traitements qui ne permettent pas nécessairement de rendre compte des comportements des agents. Ceci rend fragile l'usage d'un modèle estimé sur ces données pour évaluer les conséquences d'un changement de politique économique car il semble alors raisonnable pour prendre en compte la Critique de Lucas de modéliser la façon dont les agents prennent leur décision en fonction de leur environnement. L'estimation ne permet pas de produire une mesure exacte de ce lien dans la mesure où les données sont construites en imposant un certain comportement. Les travaux qui ont conclu à la faible pertinence empirique de la

Critique de Lucas ont peut-être été victimes de cette propriété des données. Il serait nécessaire pour aller plus loin d'apprécier les conséquences empiriques de ces hypothèses de construction des données. Cela semble en pratique difficile. De même, déjà critiquée pour certains abus par rapport à la théorie économique, la pratique de recourir à une analogie microéconomique pour interpréter les équations dites de comportement perd encore de sa pertinence. Si elle permet simplement de proposer une lecture causale entre des équations (moyennant des hypothèses identifiantes), les données ne semblent pas fournir l'information pour décrire l'équation de décision optimale d'un agent représentatif. Cette approximation microéconomique peut être raisonnablement abandonnée au profit d'une lecture en termes de grandeurs liées dans un système dynamique. La spécification d'un modèle passe alors non plus par la forme des équations qui lient les variables mais par la nature des variables qui interviennent dans le modèle et donc supposées avoir un effet significatif sur le mécanisme modélisé (de par leur contenu informationnel). Il n'y a pas de lecture causale des équations, mais une description des liens entre les variables. Le discours produit est naturellement contraint et plus abstrait. L'analyse des résultats se fait en comparant différentes situations. La crédibilité du discours s'assoit sur la capacité à décrire avec pertinence dans le passé des situations qui ont été analysées par ailleurs. Ceci suppose qu'il y a une grande stabilité des méthodes d'élaboration et des traitements des données.

Par ailleurs, la crédibilité des données conditionne leur usage et l'utilité des résultats obtenus dans des exercices de modélisation macroéconomique. Elle repose sur plusieurs objectifs contradictoires, d'une part la pertinence des descriptions proposées, d'autre part, une politique de révisions qui ne nuit pas à la crédibilité, enfin, une permanence des méthodes et des concepts.

La pertinence des descriptions proposées par un système comptable est liée à son aptitude à capter les modifications structurelles, l'introduction de nouveaux produits, la mise en place de nouveaux modes de commercialisation, de nouvelles relations économiques entre les agents dès que celles-ci ont une influence significative sur les variables publiées, i.e. au niveau de détail retenu. En pratique, c'est une tâche ardue, en particulier pour des comptes trimestriels pour lesquels on dispose d'un faible recul pour juger de l'importance et de la significativité du phénomène et en général de statistiques lacunaires. Il faut aussi pouvoir intégrer des activités nouvelles dans des nomenclatures et des systèmes de concepts qui n'ont pas nécessairement été pensés pour les accueillir. La politique de révision des évaluations annuelles et infra-annuelles est elle-aussi très importante dans la construction de la crédibilité des données. Des révisions importantes et fréquentes induiront une méfiance chez le consommateur mais réduiront sensiblement la durée de vie des études conjoncturelles ou des produits des modèles économétriques à vocation prospective.

Au total, la publication régulière de données de la comptabilité nationale autorise un ensemble d'analyses descriptives à caractère économique, politique, social ou historique. En s'appuyant sur un cadre cohérent et pérenne qui prétend à la description exhaustive des flux significatifs de création de richesse, elle permet d'en faire le bilan pendant une période donnée et d'étudier sa répartition au sein de la population décrite à l'aide d'une nomenclature. Les économistes s'appuient sur ces données pour juger de la pertinence de mécanismes que la théorie propose. Afin d'analyser les propagations de modifications et de suivre plus précisément la chronologie des enchaînements, ils ont très souvent recours à des données infra-annuelles. Ces données connaissent des fluctuations liées aux saisons ou autres phénomènes institutionnels qui les rendent difficiles à lire, des traitements statistiques leur sont appliqués. Elles doivent aussi être traitées en cohérence avec des phénomènes annuels, ce qui amène les comptables à recourir à des hypothèses sur le comportement des agents pour construire les évaluations. Si les traitements comptables et statistiques sont stables dans le temps tout en s'adaptant aux modifications fréquente du monde économiques, les données obtenues permettent de procéder à des modélisations descriptives qui autorisent des comparaisons dans le temps de différentes situations pour illustrer des mécanismes, dater des phénomènes, ... Le passage à des modélisations plus causales reposant sur des *a priori* théoriques est délicat, car l'interprétation qu'il est possible de faire des équations de comportement peut être moins directement liée à la réalité économique sous-jacente des comportements des agents. Les évaluations qui résultent de simulation de ce type de modèles peuvent éventuellement être trompeuses. Il est néanmoins difficile d'évaluer la pertinence empirique de ces limites, mais le recours croissant à des données individuelles microéconomiques est une réponse simple pour analyser les comportements des agents face à des modifications de leur environnement.

Bibliographie

- [1] Altissimo F., S., Siviero et D. Terlizzese (2002), « How deep are the deep parameters ? », *Annales d'Economie et de Statistique*, 67-68, 207-226.
- [2] Bell et Tiao (1983) : « Modelling Considerations in the Seasonal Adjustment of Economic Time Series » in *Applied Time series Analysis of Economic Data*. Ed. A. Zellner.
- [3] Bell, W.R. et S.C. Hillmer (1984), `Issues Involved with the Seasonal Adjustment of Economic Time Series', *Journal of Business and Economic Statistics*, 2, 291-320.
- [4] Box, G.E.P., S.C. Hillmer et G.C. Tiao (1978), `Analysis and Modeling of seasonal time series', in *Seasonal Analysis of Economic Time Series*, Ed. A. Zellner, Washington D.C., U.S. Bureau of Census, 309-334
- [5] Burman J.P. (1980), `Seasonal Adjustment by Signal Extraction', *Journal of the Royal Statistical Society*, Serie A, 143, 321-337.
- [6] Favero C. et D.F. Hendry (1992), « Testing the Lucas Critique : A review », *Econometrics Review*, 11, 265-306.
- [7] Geweke, J.(1979), `The Temporal and Sectoral Aggregation of Seasonally Adjusted Time Series', in *Seasonal Analysis of Economic Time Series*, Ed. A. Zellner, Washington D.C., U.S. Bureau of Census, 411-427
- [8] Geweke, J.(1985), `Macroeconomic modeling and the theory of the representative agent', *American Economic Review*, 75, 206-211
- [9] Ghysels E. (1988) `A Study Toward a Dynamic Theory of Seasonality for Economic Time Series', *Journal of the American Statistical Association*, 83, (401),168-172.
- [10] Gomez V. et A.Maravall (1996), `Programs SEATS and TRAMO: Instructions for the User' working paper n°9628, Bank of Spain.
- [11] Goodfriend, M. (1992), "Information aggregation bias", *American Economic Review*, 82, 508-519
- [12] Granger C.W.J. (1969), « Investigating Causal Relations by Econometric Models and Cross-Spectral Methods », *Econometrica*, 37, 424-438.
- [13] Gregoir S. (2005), « Un cadre de travail alternatif pour la correction des variations saisonnières de séries temporelles univariées et multivariées » *JMS 2005*, Paris
- [14] Jerison M. (1984), "Social welfare and the unrepresentative representative consumer", mimeo
- [15] Hillmer (1982) : « *Forecasting Time Series with Trading Time Variation* » *Journal of Forecasting* Vol 1 385-395.
- [16] Hoover K. (2001), *Causality in macroeconomics*, Cambridge, Cambridge University Press.

- [17] Kirman A. (1992), "Whom and what does the representative individual represent ?", *Journal of Economic Perspectives*, 6, 117-136.
- [18] Lucas R.E. Jr (1976), "Econometric Policy Evaluation : A Critique", in K.Brunner and A.Meltzer, eds *The Phillips Curve and Labor Markets*, vol 1, Carnegie-Rochester Series on Public Policy, *Journal of Monetary Economics*, suppl. 19-46
- [19] Pischke J.S (1995), "Individual income, incomplete information and aggregate consumption", *Econometrica*, 3, 805-840
- [20] SEC 95 : Système Européen de Comptes (European system of Accounts 1995) EUROSTAT, version de juin 1996
- [21] Sims, C.A. (1974), 'Seasonality in Regression', *Journal of the American Statistical Association*, 69, 618-626.
- [22] Sims C.A. (1980), "Macroeconomics and reality", *Econometrica*, 48, 1-48.
- [23] Sims C.A. (1982), "Policy Analysis with Econometric Models," *Brookings Papers on Economic Activity*, 107-152.
- [24] Wallis, K.F. (1974), 'Seasonal Adjustment and the Relation between Variables', *Journal of the American Statistical Association*, 69, 18-32.
- [25] Wallis, K.F. (1976), 'Seasonal Adjustment and Multiple Time Series Analysis', in *Seasonal Analysis of Economic Time Series*, Ed. A.Zellner, Washington D.C., U.S. Department of Commerce, Bureau of Census, 366-397