NAPCS Discussion Paper*

Guidelines for Grouping and Aggregating Products in NAPCS

Discussion Paper
for

Trilateral Steering Group Meeting
Ottawa, Canada
November 30 - December 2, 1999

Prepared by
Michael F. Mohr
ECPC Coordinator for NAPCS

* NAPCS Discussion Papers are provided to foster and facilitate professional exchange on matters related to the development and implementation of NAPCS.

I. Introduction
There are three important parts to the process of creating NAPCS:

1. identifying and defining the products produced by each NAICS industry and developing a trilateral flat list;[1]
2. creating an aggregation structure for the products in the trilateral flat list and determining the level of aggregation necessary to achieve three-country comparability in the collection of product data for each NAICS industry; and
3. creating an overall aggregation structure for the detailed products of all NAICS industries that satisfies a market-orientated, demand-based, conceptual framework.

At this point in the process we have gone a long way towards both developing agreed-upon guidelines and perfecting a working procedure for implementing the first part of the process. In addition, at the August, 1999 meeting of the Trilateral steering Group in Ottawa, we also discussed at some length both principles and guidelines for implementing parts 2 and 3 of the process and the relationship between them. In that regard, I believe we achieved agreement on the following basic guidelines:

Guideline 1: the process of groupings and aggregating products is a pragmatic exercise that should be dictated by the unique characteristics of the flat list prepared for the given NAICS industry and driven both by collectability – the ability of businesses to report data – and by willingness to collect considerations.

Guideline 2: At the NAICS industry level, the working groups should seek agreement at the lowest level of aggregation possible – to minimize the potential implications for the treatment of secondary products and for the creation of overall NAPCS aggregation hierarchy.

Guideline 3: the task of creating logical groupings for three-country comparability for a given NAICS industry should be treated as distinct, but not independent, from the task of developing a demand-side hierarchy of aggregations for all NAICS industries.

For the concept of demand-side aggregation, it was recognized that there are many competing principles – substitutability, complementarity, and other market research criteria, etc. – for grouping and aggregating products. And, a proposal was made to adopt the principle of substitution as the basis for creating the demand-side hierarchy. It was also proposed that a separate three-country panel be established to review and comment on the implications for the

---

[1]In this paper, I use the term NAICS industry as shorthand for the level at which trilateral comparability exists in NAICS. However, it should be recognized that trilateral comparability does not extend below the NAICS subsector level for the industries in NAICS 522 (Credit Intermediation and Related Activities) and 562 (Waste Management and Remediation Services), and below the NAICS industry group level for the industries in NAICs 5231 (Security and Commodity Contracts Intermediation and Brokerage) and 5239 (Other Financial Investment Activities). In addition, the industries in NAICS subsector 525 (Funds, Trusts & Other Financial Vehicles) exist only in the U.S.

overall NAPCS aggregation structure that arise from the hierarchy of products created from the industry-level flat lists by the working groups. <u>No resolution on these proposals occurred.</u>

Since the August meeting, several of the subcommittees and trilateral working groups have begun to tackle part two of the NAPCS process for the first time, and several of the same questions and controversies discussed at the August meeting are again coming to the fore including:

1.   Should the subcommittees and working groups develop a hierarchy of product aggregations from the products on the trilateral flat list for a given NAICS industry, as a prerequisite for a systematic approach to determining "comparable products" for that industry?
2.   Should constraints be set on the levels of aggregation to be included in the hierarchies developed for NAICS industries?
3.   How do the industry-level hierarchy relate to the overall NAPCS hierarchy?
4.   Should industry-level product groups and hierarchies be constructed from market-oriented, demand-side aggregation principles; if so should one principle or multiple principles be used?
5.   Should one principle (substitutes only or complements only) be used to construct the overall NAPCS market-orientated hierarchy?

This recurring list of questions is in part a reflection of the fact that our discussions in August were exploratory, tentative, and evolving. On the other hand, the list no doubt also reflects the need to think through and articulate additional guidelines for the subcommittees and working groups. In doing so, however, I think it important that we don't needlessly impose on the deliberations of the subcommittees and working group with one-size fits all restrictions. In the main, we must trust those doing the work to find "the natural order of things" within the context of broad guidelines. Proceeding with some trepidation, therefore, I attempt in the remainder of this paper to explore and suggest additional guidelines for answering the aforementioned questions and, in particular, for aggregating products in NAPCS. I present my thoughts on these issues for consideration by the Trilateral Steering Group at its November 30, 1999 meeting in Ottawa.

II. Guidelines for Aggregating Products in NAPCS[2]
As indicated, the NAPCS process requires that we confront the issue of aggregating products on two levels – first, at the within-industry level and, second, at the all or across-industry level. Since our overall objective for NAPCS is a classification system in which products are aggregated together in a manner consistent with market-orientated, or demand-side, principles, it

---

[2]This section draws heavily on the discussion and references found in M. F. Mohr, "Capital Inputs and Capital Aggregation in Production," Discussion Paper 31, Bureau of Economic Analysis, August 1988, and in J. E Triplett, "The Theory of Industrial and Occupational Classification and Related Phenomena," Proceedings from the 1990 Annual Research Conference, March 18 -21, Arlington, VA, Bureau of Census, Washington, DC: U.S. Government Printing Office, pp. 9-25.

is axiomatic that the development of that system must begin at the industry level. Accordingly, I propose that we adopt an explicit guideline to that effect:

Guideline 4:   Products will be grouped and aggregated in NAPCS at both the within-industry and cross-industry levels according to market-orientated or demand-side aggregation principles.

On both levels, therefore, the fundamental questions that must be answered are the same:

- **What do we mean by the statement market-oriented or demand-side aggregation?**
- **How do we aggregate products in a manner that is consistent with a market-orientated or demand-based conceptual framework?**

There are two complementary economic perspectives from which to answer these question: (1) a common sense or heuristic perspective and (2) an economic theory perspective.

A. A Common Sense Guideline
In order to facilitate the efforts of the subcommittees and trilateral working groups, we must provide them with practical answers to these questions. To that end, I propose that we begin by adopting for NAPCS the following working definition and guideline for the subcommittees and working groups.

Definition: The statement market-oriented or demand-side aggregation means that the NAPCS classification scheme groups and aggregates products in a manner designed not only to serve the interests and needs of those who use these products – especially consumers and businesses – but also to conform to the understanding and record keeping practices of the industries which must report data for aggregate products.

Guideline 5: The products that are grouped and aggregated together in NAPCS will share a distinctive common property, from the perspective of both users and producers, that renders that family of products different from that of another family in the way, for example, that goods are distinct from services. Moreover, as one moves from higher to lower levels of aggregation in the NAPCS structure, the common property shared by the products in the aggregate will become increasingly refined in the way, say: (1) financial service products are distinct from information service products in the overall family of service products; (2) the common property shared by loan services products is distinct from the common property shared by deposit services products in the overall family of financial services products; and (3) the common property shared by real estate-secured loan products is distinct from that shared by unsecured loan products in the overall family of loan products.

I consider Guideline 5 to be the sufficient condition for defining what I have previously characterized as "logical groups." Moreover, when viewed from the lowest levels of aggregation, this guideline is akin to Joan Robinson's "gaps in nature" rule about which she says, "provides us with a rough-and-ready definition of a single commodity – sewing cotton or

Brussels sprouts – which is congenial to common sense and causes no trouble." (Robinson, 1933, p.5; see citation in Triplett).

B. Guidelines from Aggregation Theory
To establish market-orientated or demand-side rules that are more specific than those provided by guidelines (4) and (5) requires addressing the issue of aggregation in the more formal context of the economic theory of aggregation. To that end, suppose we have an array of detailed NAPCS products, $X = (x_1, \ldots, x_n)$, and suppose some of those products are used as consumption items by consumers and others as inputs to production by businesses. What we seek are rules for grouping and aggregating these many types of products in a manner that is consistent with the underlying theory of demand. For consumers the relevant body of theory is found in the literature on consumer behavior and for businesses the corresponding body of theory is found in the theory of production and costs. In either case, the economic agent (consumer or business) is assumed to behave in a manner that optimizes his objective in the context of a budget constraint and prices for the respective products. Thus, given the budget constraint, consumers seek to optimize on the feasible solution space of their utility function and businesses seek to optimize on the feasible solution space of their production function. Now, the utility function for the representative consumer is defined on the set of products used by consumers

$$U = F(x_{c1}, \ldots, x_{cs}) = F(X_c),$$

where U indicates utility and the subscripts c on the individual products x_i designates consumer product. Similarly, we can write the production function for the representative business firm as

$$Y = G(x_{b1}, \ldots, x_{bt}) = G(X_b),$$

where Y indicates real output and where the vector of all NAPCS products X is equal to the sum of the product subvectors for consumers and business; i.e., $X = X_c + X_b$. As we look at the two functions F and G, it is evident that the function underlying U is defined on just the products used by the representative consumer and the function underlying Y is defined on just the products used by the representative firm. These facts suggest a sixth NAPCS guideline for grouping and aggregating products:

Guideline 6: The NAPCS classification scheme will group and aggregate products in a manner that distinguishes between consumer and businesses uses of products, where this distinction is both relevant and consistent with industry record keeping practices. In other words, in NAPCS the array of products in $X = (x_1, \ldots, x_n)$ should contain a subset for consumer products $X_c = (x_{c1}, \ldots, x_{cs})$ and a subset of business products $X_b = (x_{b1}, \ldots, x_{bt})$.

The question now becomes: what are the sufficient conditions in economic theory for forming aggregates from the set of detailed products contained in F and G. Since the rules of aggregation are the same for both sets of users, I employ only the utility function to develop the balance of this section and, for convenience in notation, drop the subscript c in all subsequent equations. Let us

4

begin by supposing that the application of Guideline 5 to an n-dimensional array of detailed consumer products suggests that they can be logically grouped into m distinct product subgroups X_i (i = 1, ..., m) each of which contains an unique subset of the detailed types of products drawn from the array X. We can write this organized array as:

$X = (x_{11}, x_{12}, ..., x_{1p}, x_{21}, x_{22}, ..., x_{2q}, ..., x_{m1}, ..., x_{mr})$ or more compactly as $X = (X_1, ..., X_m)$. To make concepts more concrete, think of X_1 as a vector of food products, X_2 a vector of clothing products, X_3 a vector of consumer financial services, etc. In NAPCS we seek not only to group products but also to create aggregate products that are defined just on the components in each group. Below, I briefly describe theoretical requirements for creating such aggregates.

B.1 Functional Separability

The most general of the sufficient conditions for aggregation in economics is the condition known as weak functional separability. Applied to the problem at hand, that condition can be expressed as follows:

<u>Theorem</u>: If the representative consumer's utility function F (X) is weakly separable in the partition $(X_1, ..., X_m)$, then we can rewrite the representative consumer's utility function as

$$U = F(X) = F(f_1(X_1), f_2(X_2), f_3(X_3), ..., f_m(X_m))$$

where the $f_i(X_i)$, for i = 1, ..., m, are aggregate products.

<u>Definition</u>: The function F(X) is weakly separable in the groups X_i, for i = 1, ..., m, if and only if the marginal rates of substitution (MRS) between any two products x_k and x_l within a given group X_i is independent of any product x_s from outside of that X_i. More formally, the weak separability condition requires

$$\frac{\partial[\partial F(X)/\partial x_k \, / \partial F(X)/\partial x_l]}{\partial xs} = \partial MRS_{kl}/\partial xs = 0$$

Note: weak separability of F in the product groups X_i does not require the products in the respective groups to be substitutes for each other – they can be either complements or substitutes and they can be either weak or strong complements or substitutes. However, suppose either of the following polar (strong) MRS situations apply among the products in each group:

1. MRS_{kl} is constant for all k and l in X_i; i.e., the products in the group X_i are <u>perfect substitutes</u> for each other, or

2. $MRS_{kl} = 0$ for all k and l in X_i; i.e., the products in the group X_i are <u>perfect complements</u> to each other – always used in strict proportion to each other.

Then, not only is the weak separability condition satisfied but also the aggregate product $f_i(X_i)$ can be created as a linear combination of the products in the group X_i. Wassily Leontief proved this result for the case of perfect complements, and his aggregation theorem is frequently referred to as the "Leontief composite goods aggregation theorem".

B.2 Hick Composite Goods Aggregation
In contrast to Leontief, John Hicks examined the issue of aggregation from the perspective of the relationship between the prices of the products in a group. Specifically, rather than investigating aggregation when the products in X_i are always used in fixed proportion, Hicks demonstrated that, if the prices of the products in X_i always stay in fixed proportion to each other, then it is permissible to treat that product group as a single (i.e. aggregate) product.[3]

B.3 Sequential Optimization and Hierarchical Aggregation
Embodying a more heroic set of assumptions than either simple functional separability or Hicks aggregation theorems, sequential optimization theory hypothesizes that consumers maximize their utility in a sequential fashion. Moreover, if all the assumptions embodied in the theory hold, it has been shown that the level of utility achieved by the representative consumer is equivalent to what he would obtain by instantaneously allocating his total budget over the detailed products in X. Sequential optimization theory seeks to provide a behavioral paradigm that rationalizes the creation of sequenced aggregations, from small to large, typically required in empirical applications of optimization theory to consumers and business. Suppose, for example, we want to rationalize the following hierarchy of aggregations for food products:

Level 1 products: beef, pork, poultry, fish, frozen vegetables, fresh vegetables, wheat bread, rye
                              bread, white bread, milk, cheese, butter, etc.
Level 2 aggregates: meats, vegetables, bread, dairy, etc.
Level 3 aggregate: food.

Sequential optimization theory says the above aggregation structure is consistent with a behavioral construct wherein consumers make predetermined allocations of their total budget and maximize their utility in stages, starting from the highest level of aggregation to successively lower levels. It also requires that consumers have knowledge of aggregate price indexes that apply to the aggregates in play at each stage of optimization. Given his budget allocations and the price indexes, the consumer behaves as follows. First, he maximizes by allocating his overall budget

---

[3]Its important to note that the aggregation conditions developed in sections B.1 and B.2 extend to justifying more refined aggregations of products in F(x). Imagine, for example, that we want to partition the financial services array X_3 into a sub-arrays of loan products, X_{31}, deposit service products, X_{32}, etc. and that we want to write F(x) as F($f_1(X_1)$, $f_2(X_2)$, $f_31(X_{31})$, $f_{32}(X_{32})$, ..., $f_{3p}(X_{3p})$, ..., $f_m(X_m)$). In this case, we can form aggregate products over these more refined groupings of X_3 products provided either (1) F is weakly separable in the new groups or (2) the prices of the products in the new groups move in strict proportion to each other.

over the Level 3 aggregates – food, clothing, etc. Second, he maximizes his utility for food by allocating his food budget over the more refined Level 2 aggregate products: meat, vegetables, bread, and dairy. Third, he maximizes, say, his utility for meat by allocating his meat budget over the Level 1 products: beef, pork, fish, and poultry. When this behavior is extended to all major groups X_i, it is tempting to conclude that sequential optimization is a natural paradigm for rationalizing the entire NAPCS classification hierarchy. However, the AEA members of the Census Advisory Committee (April 22-23, 1999 Meeting) have advised against this proposition because the strong assumptions underlying the theory do not hold up in empirical studies of consumer and business behavior.   Ariel Pakes, for example, comments that this paradigm, "is a framework that has failed for economists and the Bureau should not be using it either."

C. Should Aggregation in NAPCS be Substitution Based
As discussed, aggregation theory does not require that a single demand-side principle be used to rationalize the creation of aggregate products in NAPCS, and I strongly recommend against adopting such a rule for aggregations formed at the NAICS industry level. Rather, I recommend adopting a guideline of the following sort:

Guideline 8: While giving primacy to the principle of substitution at the overall NAPCS level, aggregation at the industry level will employ the demand-side principle that best assures that Guideline 5 is satisfied.

In looking over the flat list of products and the product groups that have evolved to date, I believe the first requirement is to satisfy ourselves that the proposed groups and aggregates conform to some reasonable rule for demand-side aggregation in the sense Guideline 5 and one or more aggregation principles. However, no one demand-side aggregation principle can be used to rationalize every proposed aggregate, and any attempt to force every aggregate product in the NAPCS structure to be one created just from products that are substitutes for each other is likely to result in a structure that grossly violates Guideline 5. This because there is no one rule that can fit every situation and time period. In some situations, the products within a "logical group" may be rather unambiguously viewed as substitutes, while in others complements. Still, in other groupings the relationship among the products in question may be one of substitutes from the perspective of some users and one of complements from the perspective of others.[4] This fact can be demonstrated by a few examples.

---

[4]It is important to note that the conclusions I have drawn here are consistent with the advice given on this issue by the AEA panel at the April 22-23, 1999 Meeting of the Census Advisory Committee of Professional Associations. In particular, Professor Roger Bentancourt notes: "At the fundamental level, what constitutes a substitute or complement for consumption depends on the period of analysis. In the area of services, technological change is creating new substitutes that did not previously exist — e.g., teleconferencing now is becoming a substitute for actual physical meetings. This could lead to telecommunications and transportation as part of a related product group, but not doing so would violate the ... approach for classifying products." To better avoid such anomalies, Bentancourt suggests that we not try to build a system erected on a single classification principle. Rather, he suggests: "The classification *(framework)* adopted can be justified on the basis of convenience without any claims of ultimate truth, and it should emphasize flexibility to allow users to reconfigure the classification for their own uses."

Example 1: The flat list of products for NAICS 5221 contains a single product for the first mortgages sold to consumers rather than several products delineated, say, by duration and/or size of loan. In short, the working group has implicitly aggregated the various first mortgage products into a single aggregate product. As an empirical matter, it is probable that this implicit aggregate can be best justified by assuming that the Hicks-aggregation theorem applies rather than by assuming that the set of alternative first mortgage products are perfect substitutes for each other.

Example 2: The flat list for NAICS 5412 shows three types of financial audit products: traditional, tax, and specialized. Applying Guideline 5, I anticipate the working group will propose creating a product group and aggregate called, say, Financial Audits. In contrast to the above example, however, I find it is easiest to rationalized this aggregate by viewing it as a weakly separable aggregate of a group of products that are complements to each other in the production (and demand functions) of the representative business that uses these audit services.

Example 3: The flat list for NAICS 5412 contains a product group and product aggregate called Compilation and Review Services that consists of products – compilation, review, or both.[5] While the products included in this new group embody services that are similar to those embodied in audits, they are excluded from the audit products group, because they are less rigorous and lack either the same degree of professional competence or assurance. Nevertheless, applying Guideline 5, it seems natural that the working group will propose a higher level aggregate, say, Assurance Services, that is formed by aggregating over both Audit Services and Compilation and Review Services. In this case though, it is much less clear whether the products in the two groups should be viewed as substitutes for or complements to each other. For example, they are likely complements to each other for businesses that are required to do audits but substitutes otherwise.

III. The Process of Creating The NAPCS Classification Structure

To date we have charged the subcommittees and working groups with creating logical groups and aggregates from the products in industry flat lists and with determining the level of aggregation required to establish comparability across the three countries. However, we haven't been able to specify either the architecture (including levels of aggregation) for the overall NAPCS classification hierarchy or how it will evolve from the efforts of the working groups to organize and structure the products of individual NAICS industries. In my mind there is good reason for this state of affairs – we simply don't know enough at this time and may not for some time to come. Nevertheless, it is against this backdrop, that the Steering Group must address the following questions:

1.   Should the subcommittees and working groups develop a hierarchy of product aggregations from the products on the trilateral flat list for a given NAICS industry, as a

---

[5]I find the aggregate Compilation and Review Services consistent with Guideline 5 and I would again rationalize it as one created from products that are complements to each other.

prerequisite for a systematic approach to determining "comparable products" for that industry?

2.   Should constraints be set on the levels of aggregation to be included in the hierarchies developed for NAICS industries?

3.   How does the industry-level hierarchy relate to the overall NAPCS hierarchy?

I strongly recommend that the Steering Group adopt an affirmative guideline for the working groups with respect to question 1. And with respect to question 2, I recommend that the Steering Group either not establish a guideline or one with a very liberal constraint. These recommendations are based in large part on the view that the industry-level aggregation process serves several indispensable and interrelated objectives:

- Ordering the products in industry flat lists into logical groups and aggregates,
- Defining an ordered hierarch of logical groups and aggregates that can be used to systematically explore and determine both the comparable three-country product detail and the national product detail for that industry, and
- Providing suggestive building blocks for grouping and aggregating products in the overall NAPCS classification hierarchy.

Given these objectives, it is fundamental that every working group develop and seek agreement on a well ordered hierarchy of the products contained in the flat list of each industry. (To date, disagreements in this area have not been an issue but rather a positive factor in the deliberations of the working groups.) For some industries, the number of products in its flat list will be small and/or the number of legitimate groups and desired levels of aggregation will be simple and limited. In others, however, the list of products will be extensive and the number of groups and levels of aggregation complex and extensive. Further, it is not possible to determine a priori what level of aggregation will be necessary to establish comparability in each industry, due both to differences in the perceived availability and collectability of products across countries. (We are also a long way from knowing where common ground will be found between a proposed list of national products and collectability.)