quarterly U.S. real GDP demonstrates the importance of this phenomenon in practice.

The analysis presented in this note is not specific to the BK filter, but indeed applies to all bandpass filters. The simple act of differencing does not remove a stochastic trend; it merely renders it stationary. Therefore, although bandpass filtering can render an integrated series stationary, the properties of the filtered series will depend on the trend in the unfiltered series.

REFERENCES


THE GENERALIZED COMPOSITE COMMODITY THEOREM: STRONGER SUPPORT IN THE PRESENCE OF DATA LIMITATIONS

George C. Davis*

Abstract—Because of common data limitations, the existing testing framework for the generalized composite commodity theorem (Lewbel, 1996) is incomplete. This note clarifies and strengthens the testing procedure by implementing modified Bonferroni procedures. The conditions are established for consistency between the existing and modified Bonferroni tests. In an empirical application, the Bonferroni tests provide more powerful support for the generalized composite commodity theorem than is obtained from the existing test.

I. Introduction

One of the most problematic areas in all of economics is the consistent aggregation of commodities. For over fifty years there have existed two solutions: the composite commodity theorem (Hicks, 1936; Leontief, 1936) and the separability theorem (Leontief, 1947; Sono, 1961). Rejections of both are common. Recently, Lewbel (1996) proposed a generalization of the Hicks-Leontief composite commodity theorem that is appealing for several reasons. First, it imposes no restrictions on preferences. Second, it allows prices to be highly but imperfectly collinear, a fact frequently observed. Three, it applies the modifications of the composite commodity theorem. The note closes with conclusions.

II. The Generalized Composite Commodity Theorem

Drawing freely from Lewbel (1996, pp. 526–527), let $p_i$ and $w_i$ be individual prices and budget shares of individual goods $i = 1, 2, \ldots, n$, and $p$ and $w$ be the corresponding vectors. Similarly, let $P$ and $W$ represent vectors of group price indices $P_j$ and group budget shares $W_j = \sum_{i \in J} w_i$, where $J$ indexes groups of goods or commodities $J = \{1, 2, \ldots, N\}$ and $N < n$. Each good $i$ is an element of a group $J$. Furthermore, define $r_i = \log p_i$, $R_j = \log P_j$ and let $r$ and $R$...
represent the corresponding vectors. The price aggregation error is
defined to be the difference between the log of the individual good
price and the log of the group price index for which the individual
good is a member: \( p_i = r_i - R_i \) for \( i \in J \). The vector of aggregation
errors is denoted by \( \mathbf{p} \), and \( z \) denotes the log of a consumer’s total
consumption expenditure.

A good’s budget share \( w_i \) is defined to be composed of a systematic
Marshallian demand function \( g_i(r, z) \) plus an error term \( e_i \) with
a conditional mean zero such that \( w_i = g_i(r, z) + e_i \), so \( E(e_i|r, z) = 0 \)
and \( g_i(r, z) = E(w_i|r, z) \). Aggregate demands are similarly defined
as \( \sum_{i \in J} w_i = W = G(R, z) + \epsilon, \) so \( E(\sum_{i \in J} W_i|R, z) = E(W|R, z) = G(R, z) \),
given that the conditional expectation of \( \epsilon \) is zero.

Now assume:

1. The demand functions \( g_i(r, z) \) for \( i = 1, 2, \ldots, n \) are rational.
2. The distribution of the vector \( \mathbf{p} \) is independent of \( \mathbf{R} \) and \( z \).

The first assumption is equivalent to assuming utility maximization.
The second assumption relaxes the Hicks-Leontief condition that \( \mathbf{p} \)
is constant. Based on these and some milder assumptions, Lewbel
proves that the group demand function \( G(R, z) \) satisfies the properties
of a demand function. The independence assumption 2 is the focus
of the testing.

III. Data Limitations and Modified Bonferroni Procedures

The independence assumption 2 implies that if there are \( n \) disaggre-
commodity groups to be aggregated into \( N \) aggregate groups \( (N < n) \),
then the \( n \) aggregation errors \( p_i, i = 1, 2, \ldots, n \), should be
independent of each of the \( N \) aggregate nominal or deflated price
indices \( R_i, i = 1, 2, \ldots, N \). With sufficient observations and
stationary data, an \( F \)-test of significant coefficients in the auxiliary
regression of \( p_i \) on \( R_1, R_2, \ldots, R_N \) for each \( i \) is appropriate.
Theil (1971, chapter 11) discusses this test for aggregation bias. With
sufficient observations and nonstationary data, a multivariate cointe-
geration test (see for example Johansen, 1995) between each \( p_i \) and \( R_1, R_2, \ldots, R_N \) is appropriate. However, with insufficient observations
and nonstationary data the testing is more treacherous.

A. Implications of Limited Data

It is now well documented that with fewer than 100 observations on
nonstationary data there can be substantial size distortions (spurio-
sous cointegration) and power problems in testing for cointegration,
especially when there exist more than three variables (see, for exam-
ple, Haug, 1996; Ho and Sørensen, 1996). Most applied demand
studies involve more than three goods and use quarterly or annual
data, so the size and power problems will arise in exactly the situation
where the generalized composite commodity theorem would most
likely be applied. Indeed, Lewbel faced such data limitations in his
example.

In light of the data, size, and power problems associated with
multivariate procedures, some alternative must be pursued. Lewbel’s
simple and practical alternative is to apply a bivariate cointe-
geration test to the variables that would seem most likely to be cointegrated.

---

1 As in Lewbel (1996), because the demand function is homogeneous of
degree zero in prices and income, the price indices could be redefined
as deflated price indices and then the independence would be just between \( \mathbf{p} \)
and \( \mathbf{R} \). Therefore, it is sufficient to talk only about the independence
between \( \mathbf{p} \) and \( \mathbf{R} \).

Specifically, if the aggregation error \( p_i \) is going to be cointegrated with
any of the \( R_i, R_2, \ldots, R_N \) aggregate price indices, it would seem most
likely to be cointegrated with the price index from which it is
constructed, namely \( i \in J \). Lewbel’s simple test of the generalized
composite commodity theorem is then a cointegration test between \( p_i \)
and \( R_j \) when \( i \in J \). However, absence of own-group cointegration
does not guarantee absence of cross-group cointegration.

It is not difficult to imagine cross groups that also may be
cointegrated. For example, going to a restaurant and having wine
before going to the theater is common. If wine is an element of a food
aggregate and theater is an element of a recreation aggregate, then,
the aggregation error for theater may be cointegrated with the food price
index even though theater is not in the food group. More generally, the
bare assumption that \( p_i \) and \( R_j \) are not cointegrated when \( i \in J \) does
not imply that \( p_i \) and \( R_j \) are not cointegrated when \( i \notin J \). Rejecting
cointegration between \( p_i \) and \( R_j \) when \( i \notin J \) amounts to satisfying a
necessary condition but not the sufficient condition for the generalized
composite commodity theorem. The question is: how can a test of the
sufficient condition be constructed given the data limitations? The
testing can be improved by recognizing that if the bivariate cointe-
ration test when \( i \in J \) is acceptable, then a bivariate cointegration
test when \( i \notin J \) also must be acceptable. This simple fact can be
exploited to develop a more powerful test of the generalized compos-
tive commodity theorem using Bonferroni procedures.

B. Modified Bonferroni Procedures and Implications

Consider the general testing problem. Suppose there are \( N \) indi-
vidual hypotheses \( H_1, H_2, \ldots, H_N \) each being tested at the \( \alpha \)
level with corresponding \( p \)-values \( p_1, p_2, \ldots, p_N \). Let \( H = \{ H_1, H_2, \ldots, H_N \} \), and define the family hypothesis \( H_0 \) to be the intersection of all
hypotheses in \( H \), or \( H_0 = \bigcap_{j=1}^N H_j \) (Hochberg and Tamhane, 1987).
In the present setting, \( H_j \) can refer to the null hypothesis of no
cointegration between say, \( p_1 \) and \( R_j \). The test of independence
between \( p_i \) and all \( R_j \) is then just a test of the family hypothesis \( H_0 =
\bigcap_{j=1}^N H_j \). The\textit{familywise error rate} (FWE) is the probability of a
Type I error for a family hypothesis. To control the FWE, multiple
comparison procedures based on the Bonferroni inequality can be
implemented.

In econometrics, Savin (1984) discusses multiple comparison
procedures, and more recent applications are Maddala and Wu (1999)
and Dufour and Torres (1998). As Dufour and Torres point out, multiple
comparison procedures are especially useful when standard asympto-
tic methods are either not applicable or unreliable, which is certainly
the case here. Though Dufour and Torres recommend the regular
Bonferroni procedure, several more powerful procedures are used
here. These procedures can be classified into three categories: single-
step, step-down, and step-up. Step-down and step-up procedures are
in general more powerful than single-step procedures.

Table 1 summarizes Lewbel’s simple decision rule and four
Bonferroni decision rules.\textsuperscript{2} The simple procedure only calculates
one of the \( p \)-values associated with the elementary hypotheses \( H_1, H_2, \ldots, H_N \)
constituting the family hypothesis \( H_0 \). The calculated \( p \)-value is for the
in-group elementary hypothesis \( H_{i,j} \), and the other \( N - 1 \) cross-group elementary hypotheses \( H_{i,k} \) are ignored. This will

\textsuperscript{2} These are four of the most popular and simplest methods. Other
theoretically more powerful procedures exist but they are much more
complicated to implement. Moreover, several studies (such as Olejnik et
al., 1997; Sarkar and Chang, 1997) find that the gain in power is negligible
in many cases and that the procedures presented here are robust to
alternative assumptions about the distribution of the test statistics.
Table 1.—Alternative Test Procedures and Decision Criteria

<table>
<thead>
<tr>
<th>Test Procedure</th>
<th>Decision Criterion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single (1)</td>
<td>Reject $H_i$ if $p_i \leq \alpha$ for the calculated $p$-value $p_i$.</td>
</tr>
<tr>
<td>Bonferroni</td>
<td>Reject $H_0$ if $p_i \leq \alpha/N$ for any $i$.</td>
</tr>
<tr>
<td>Simes</td>
<td>Reject $H_0$ if $p_{(i)} \leq i \alpha/N$ for any $i$.</td>
</tr>
<tr>
<td>Step-down</td>
<td>Holm</td>
</tr>
<tr>
<td></td>
<td>Reject $H_{(j)}$ if $p_{(j)} \leq \alpha/(N-j+1)$ for all $j = 1, 2, \ldots, i$.</td>
</tr>
<tr>
<td>Step-up</td>
<td>Hochberg</td>
</tr>
<tr>
<td></td>
<td>Reject $H_{(i)}$ if $p_{(i)} \leq \alpha/(N-j+1)$ for any $j \geq i$.</td>
</tr>
</tbody>
</table>

be referred to as the single-test procedure. Alternatively, to implement the Bonferroni procedures, all of the $N$ $p$-values associated with the elementary hypotheses $H_1$, $H_2$, \ldots, $H_N$ constituting the family hypothesis $H_0$ are calculated. In the step-up and step-down procedures the individual $p$-values are first arranged in increasing order $p_{(1)} \leq p_{(2)} \cdots \leq p_{(N)}$ along with their corresponding hypotheses $H_{(1)}$, $H_{(2)}$, \ldots, $H_{(N)}$ before the testing commences. The decision criterion then relies on adjusting the individual significance levels by some function of the number of tests involved.

In the single-test procedure, if the single calculated $p$-value is less than the chosen significant level $\alpha$, the family hypothesis $H_0$ is rejected; otherwise, it is retained. The regular Bonferroni procedure rejects the family hypothesis $H_0$ if any individual $p$-value is less than $\alpha/N$. The Simes (1986) procedure rejects the family hypothesis $H_0$ if any $p_{(i)} \leq i\alpha/N$. The Simes procedure is more powerful than the regular Bonferroni procedure because the significance levels are adjusted with the ordering of the $p$-values. The Holm (1979) procedure rejects the hypothesis $H_{(i)}$ when $p_{(i)} \leq \alpha/(N-j+1)$ for all $j = 1, 2, \ldots, i$. If the hypothesis $H_{(i)}$ is rejected, then too are all of the subsets for $j < i$. Testing stops when the first nonsignificant result is encountered, and all remaining hypotheses are retained. Because the Holm procedure will reject any hypothesis rejected by the Bonferroni procedure, it is at least as powerful as the Bonferroni procedure. In the Hochberg (1988) procedure, the hypothesis $H_{(i)}$ is rejected when $p_{(i)} \leq \alpha/(N-j+1)$ for any $j \geq i$. Here a hypothesis is tested if and only if all its implied hypotheses are retained. If any hypothesis is rejected, then all implied hypotheses are rejected. The Hochberg procedure will reject all hypotheses rejected by the Holm procedure but may test and reject hypotheses not examined by the Holm procedure, and so it is more powerful than the Holm procedure. More detailed discussion of these tests can be found several places (for example, Olejnik et al., 1997).

As seen, a key aspect of the modified Bonferroni procedures is the ordering of all the $p$-values from the minimum $p_{(1)}$ to the maximum $p_{(N)}$ and calculating the corresponding adjusted significance levels from the minimum $\alpha_{\text{min}}$ to the maximum $\alpha_{\text{max}}$. Because of these characteristics, and the fact that the single-test procedure only calculates one $p$-value, it is easier to make comparisons across procedures by considering the ordering and the extreme cases.

Table 2 summarizes the decision criteria depending on the relationship between $\alpha_{\text{min}}$ and $\alpha_{\text{max}}$, $p_{(1)}$, and $p_{(N)}$. The Bonferroni procedures require all $N$ tests and control the overall Type I error or the FWE. The FWE is set equal to $\alpha_{\text{max}}$, note that $\alpha_{\text{max}} = \alpha_{\text{max}}/N$. To grant the single-test procedure maximum power, it will be assumed the chosen significance level in the single-test procedure is $\alpha_{\text{max}}$. Table 2 indicates that the single test procedure is consistent with all Bonferroni procedures in rows 1, 2, and 6. It is also consistent with the Simes and Hochberg procedures in row 4. The inconsistencies arise between the single-test procedure and the Bonferroni procedures in rows 3 and 5. In these cases, the problems arise because the single-test procedure only calculates one of the $N$ $p$-values and there is no way to make a decision without calculating the other $N-1$ $p$-values and adjusted $\alpha$-values. Without the other $N-1$ $p$-values and adjusted $\alpha$-values, there is little point in speculating about the validity of decisions based on a single $p$-value and $\alpha_{\text{max}}$ in cases 3 and 5.

IV. An Empirical Demonstration

To demonstrate the testing procedures, the same data analyzed by Lewbel are analyzed here and the results summarized. The data are constructed as described in Lewbel (1996, section IV, pp. 531–535). The data consist of 46 individual goods and 7 aggregate goods from the U.S. national income and product accounts (NIPA) for the period 1954 to 1993. Lewbel finds that in almost all cases the series are nonstationary. These results are taken as stated, and attention centers on the cointegration tests. To reduce distorting the comparison with Lewbel’s results, Lewbel’s implementation of the Engle-Granger (1987) test reported in his table 2 for the single test is followed verbatim for all tests. An Engle-Granger cointegration test with a trend is conducted for each of the seven nominal and deflated pairs ($p_1, R_1$), ($p_2, R_2$), \ldots, ($p_7, R_7$) for each $i = 1, 2, 3, \ldots, 46$. The FWE rate is chosen to be 0.15. As in Lewbel, the analysis is conducted on the nominal and deflated (homogeneity imposed) data.

For the nominal data, in all but nine cases, the minimum of the ordered $p$-values is greater than the maximum of the adjusted significance levels, which corresponds to row 6 in table 2. Consequently, all procedures are consistent, and the family hypothesis of independence is retained by all procedures in all but these nine cases. These nine cases correspond to row 5 of table 2, where $\alpha_{\text{max}} < p_{(1)} < \alpha_{\text{max}} < p_{(N)}$. In this case the less powerful Holm procedure would retain the family hypothesis, but the individual $p$-values and adjusted significance levels must be considered for the more powerful Hochberg and Simes procedures.

Table 3 gives an example of one of these nine cases and shows how the single test compares with the modified Bonferroni procedures when there are potential inconsistencies. The aggregation error is for spectator sports ($p_{(44)}$), which is an element of the recreation group. In this case the single test’s corresponding $p$-value is the largest (0.79) and is given in bold. Based on the single test procedure, the null of no cointegration is retained, as $0.79 > 0.15$. Similarly, based on the original Bonferroni procedure, the family hypothesis would be retained, as all $p$-values are greater than 0.02. Based on the Holm procedure, the family hypothesis is retained, because the smallest $p$-value, associated with the clothing group, is greater than the adjusted significance level (0.027 > 0.02). The Hochberg procedure rejects the family hypothesis, because the ordered $p$-value of 0.05 associated with the transportation group is less than 0.08 and therefore all remaining hypotheses are rejected. Similarly, the Simes procedure rejects the family hypothesis, because there is at least one $p$-value that is less than the adjusted significance level. Based on table 3, the family hypothesis and therefore the generalized composite commodity theorem is not satisfied for spectator sports.

The other eight possibly inconsistent cases in the nominal data were also investigated. Live theater was the only other case where the
modified Bonferroni procedures lead to a rejection of the family hypothesis but the single-test procedure leads to a nonrejection. In addition, the deflated data were analyzed, and the generalized composite commodity theorem was unanimously supported across all tests. So where Lewbel found no violations of the generalized composite theorem, spectator sports and live theater are found to violate the theory using the more powerful modified Bonferroni procedures, but only in the nominal data.

Some may argue that the modified Bonferroni procedures should be done over all tests—a grand test—and not just within the seven tests of the individual aggregation errors. That test was conducted as well, and after adjusting the significance levels appropriately, unanimous support was found for the generalized commodity theorem based on this grand test for the nominal or deflated data.

As discussed, the multivariate cointegration tests are known to have power and size problems for the type of data encountered here. However, for comparison purposes all of the corresponding 46 multivariate cointegration tests were also conducted. Specifically, using the Johansen (1995) error correction framework, cointegration was checked between the variables $p_i$, $R_1$, $R_2$, $R_3$, $R_4$, $R_5$, and $R_7$ for all $i = 1, 2, \ldots, 46$. The results turned out as expected and so are just summarized here. In general, there were either five, six, or seven cointegrating vectors in each of the 46 systems. Of course the cointegration may be just between the aggregate price indices, so a variable-exclusion test (Johansen and Juselius, 1990) was conducted as well, excluding $p_i$. If the exclusion hypothesis is rejected, then this would suggest that the generalized composite commodity theorem would be rejected as well. The tests were conducted for five, six, and seven cointegration vectors for each commodity. The exclusion restriction tests indicated that there were 39 (30) cases for the nominal (deflated) data where the generalized composite commodity would have been rejected at the 0.01 or lower level. But again, this would be expected in view of the documented literature, the small sample size, the large number of variables in the cointegration system, and the ensuing power and size problems associated with such tests.

### V. Conclusions

The generalized composite commodity theorem is an important contribution to economics because it offers hope that aggregate demand analysis may finally be based on an empirically supported aggregation theorem. Using modified Bonferroni procedures, this note provides a way for increasing the power of the tests of the generalized composite commodity theorem, especially in data sets where standard procedures are likely to be unreliable. The conditions under which the single test will be consistent with the modified Bonferroni procedures are presented and discussed. Applying the more powerful tests to the same data set Lewbel analyzed, the generalized composite commodity theorem is only rejected for two goods, live theater and spectator sports, for the nominal data. For the deflated data (homogeneity imposed), the generalized composite commodity theorem is unanimously supported by all procedures.

### REFERENCES


HOW EFFECTIVE ARE TRADE BARRIERS? AN EMPIRICAL ANALYSIS OF TRADE REDUCTION, DIVERSION, AND COMPRESSION

Jon D. Haveman, Usha Nair-Reichert, and Jerry G. Thursby*

Abstract—We analyze the effects of trade barriers using highly disaggregated data. The level of disaggregation allows us to separate the effects of tariffs and nontariff barriers (NTBs) into reduction, diversion, and compression effects. We find that multilateral tariffs significantly reduce trade flows and that trade preferences have a significant diverting effect. We also find that higher multilateral tariffs tend to shift trade towards larger exporters, suggesting that the desire to minimize fixed costs associated with trading dominates any preference for variety. In the case of NTBs, we find that, as often as not, the imposition of an NTB leads to an increase in the value of trade; in industries with low import demand elasticities, the influence of rising prices outweighs the decline in quantity.

I. Introduction

Some of the most important questions regarding barriers to international trade relate to the effectiveness of tariff and nontariff measures and their impact on multilateral and bilateral trade flows. The expected impact of measures of protection on trade flows is an integral component in investigating many issues related to protection—for example, the pattern and height of barriers to trade across industries, or the expenditure of lobbying effort by labor in search of protection. Similarly, the outcome of trade liberalization initiatives is seldom an across-the-board tariff cut; instead, it may be biased towards protective measures that provide relatively little in the way of protection for domestic industry.

In this study we analyze differences in the effectiveness of tariffs and four types of nontariff barriers (NTBs) across disaggregated industries. In the process, we identify specific distortions of bilateral trade flows that can arise from the imposition of barriers. We extend the literature on the effect of protection in several ways. First, unlike previous studies, we separate the effect of tariffs and NTBs into a reduction effect, which is a lowering of overall trade; a compression effect, which is a concentration of the source of imports into the largest exporters; and a diversion effect, which is a shift in trade patterns across exporters that is unrelated to size. This type of decomposition has policy relevance in that it permits the identification of specific channels through which protection affects trade, as well as the relative strengths (effectiveness) of these channels. Second, we use data that are more disaggregated than is typical of the literature on trade barriers. This allows for large samples, so that we can consider regressions for narrow commodity groups; this not only mitigates specification error due to structural differences in the determinants of trade flows across commodities, but it also allows for industry-by-industry comparisons of the effects of barriers.1

Our analysis indicates that, in addition to the expected reduction of trade caused by barriers, there is substantial diversion of trade flows resulting from the preferential application of tariffs, and that tariffs tend to shift trade towards larger exporters. We interpret this latter effect as due to a desire to minimize a fixed cost of trading that is related to the number of countries with which an importer trades. In other words, with fixed costs, even a constant multilateral tariff can redistribute trade towards large countries; that is, trade with small exporters is sacrificed first.

In the case of NTBs, we find that the price-raising effect of an NTB frequently dominates the quantity-reducing effect, resulting in an increase in the value of trade between two countries. In addition, we

Received for publication February 22, 1999. Revision accepted for publication February 26, 2002.

* Public Policy Institute of California, Georgia Institute of Technology, and Emory University, respectively.

The authors would like to thank several anonymous referees, Simon Evenett, Marie Thursby, various individuals at the Board of Governors, and participants at the Third Annual Empirical Investigations in International Trade Conference for many helpful comments. Financial support was provided, in part, by the Purdue University Center for International Business Education and Research.

1 Anderson (1985) states the case for the use of disaggregated data based on the finding that aggregation significantly reduces estimated elasticities.