Are producer prices good proxies for export prices?

Results of a recent study show that producer price indexes remain imperfect measures of export price trends; major conceptual and methodological differences underlie the differences between the two series

Bill Alterman

Accurate export (and import) price indexes have long been considered crucial for the construction and analysis of such critical economic data as measures of inflation-adjusted net exports, price elasticities of goods in world trade and estimates of the impact of exchange rate fluctuations on domestic inflation and U.S. "competitiveness." Historically, price indexes reflective of total output for U.S. companies, such as the Bureau's producer price indexes (PPI), were frequently used as substitutes for export price indexes. In order to assess the similarity between these two series, the Bureau of Labor Statistics (BLS) recently undertook a project to compare price trends of the Bureau's export price indexes (XPI) with producer price indexes.

The Producer Price Index is designed to measure price changes in a domestic industry's total output, regardless of the destination of that output. The Export Price Index is designed to measure price changes of only those goods that are physically shipped out of the country. Thus, the export price index would appear, at first glance, to be a subset of the producer price index and arguably derivable from it. In general, nearly all of the companies, products, and prices that make up the universe of U.S. exporters also are found in the universe of U.S. domestic producers.[1] However, there are significant differences between the two indexes, including differences in the underlying concepts and in the operational methodology used to construct the indexes.

Bill Alterman is an economist in the Office of Prices and Living Conditions, Bureau of Labor Statistics.

Concepts and methods

From a production theory point of view, the conceptual export price index underlying the XPI is similar to the conceptual output price index underlying the PPI. Both generally focus on a revenue-maximizing firm for given product prices, input quantities, and technology. There is, however, one significant difference. The PPI normally will not price goods that are being "sold" from one division or plant of a firm to another arm of the same firm. Under the definition of a *net* output index, these so-called intracompany transfers are considered out-of-scope of the PPI when the different branches of the same company are considered to be in the same industry. In contrast, the XPI represents everything that is exported.

In order for the PPI and the XPI to move similarly, two conditions must hold. The first is that identical goods in different countries sell for the same price when their prices are expressed in a common currency. This is usually referred to as the "law of one price." For example, if we were to assume that markets were competitive and that complete arbitrage were possible (that is, that there were zero price discrimination), there should not be any difference between the export price of a Chrysler minivan and its domestic price.[2] A number of empirical studies, however, have found that the law of one price does not hold. Possible explanations for these deviations

include transportation costs, tariffs, and nontariff barriers, as well as producers' attempts to maximize revenue by varying the price of a good based on destination; this is usually referred to as "pricing to market."[3] Short-term economic trends, such as exchange rate fluctuations or differences in aggregate growth rates, could also play a part. For example, if the dollar were to appreciate sharply against the British pound, the corresponding price of a minivan in terms of the pound would also rise sharply. As a consequence, the exporter might choose to lower his dollar price during the short term in order to maintain market share.[4]

Even if the law of one price holds between U.S. export prices and U.S. domestic prices, a second condition would have to be met in order for producer price indexes to be adequate proxies for changes in export prices: similar goods would have to have the same rate of price change. Because the XPI and the PPI are based on limited samples, it is likely that the actual products priced for the PPI and XPI indexes differ. For example, while the PPI may price a Chrysler minivan, the XPI may sample a Ford minivan. In order for the indexes to trend together, prices of the Chrysler minivan need to trend similarly to prices of the Ford minivan. However, because of differences in consumer tastes and other factors, the minivan that is exported may be very different from the one sold for domestic consumption. For example, a minivan sold to Great Britain would have the steering wheel on the right, while one sold in the United States would have the wheel on the left; or a minivan exported to Europe might be smaller and have different pollution control restrictions. To take the extreme example, suppose that the PPI index for a particular product group contained no exported products, but that such products had, in fact, been exported and are the basis for the XPI. Applying an export weight to the PPI index as part of an estimation of export price change would make little sense without the assumption that the price movements of similar products are the same. In short, to the extent that PPI indexes consist of products that are not in the XPI, and given that the two conditions necessary for index parity are unlikely to be satisfied, one might expect that many producer price indexes would be poor proxies for the export price indexes representing the same product groups.

Finally, it can be shown that in a general framework, one that includes all sectors of the economy, the appropriate share weights for index construction are product shares of gross domestic product. However, the conceptual indexes underlying the PPI and the XPI are conditioned on activity in differing sectors of the economy, and, thus, the relative importances for each grouping may be different at any and every level of aggregation. In the PPI, the weights are item shares in the total value of shipments; in the XPI, the weights are item shares of the value of only those U.S. goods that are traded externally. Furthermore, product categories that tend to be intracompany transfers will tend to have a greater weight in the XPI. This can have a significant impact in product areas in which substantial amounts of trade are intracompany transfers, such as semiconductors.

Exhibit 1 summarizes some of the major methodological differences between the two price programs. Some of these differences can be attributable to the different uses for which these indexes are intended. In 1978, the PPI was substantially revised in order to make it consistent with the input-output framework of the U.S. economy. The PPI program samples and estimates price changes across *industries* to produce net output indexes. Items produced and consumed within a given industry are *excluded* from the program's net output indexes. It is also important to note that, in the PPI, the same item from two different industries would show up in two *different* indexes. The most detailed classification structure used in constructing PPI indexes is based on items classified according to the Standard Industrial Classification (SIC) of the producing firm.[5]

The indexes produced by the BLS international price program are designed to deflate U.S. exports and imports in order to construct the real net export component of the U.S. National Income and Product Accounts, which are product-based as opposed to industry-based. Thus, the program samples and then estimates price change across *product* areas to produce import and export price indexes. The same items in two different industries would show up in the *same* index. Traded goods are classified according to the Harmonized System (HS) nomenclature used for documenting U.S. import and export trade flows.

Sampling. The impact of differing aggregation structures is first manifested in sample selection. Both the PPI and the XPI programs use a probability-proportional-to-size (PPS) multistage sampling technique to select sampling strata and establishments. Both programs also use PPS disaggregation techniques to select unique items for pricing. However, the programs draw their samples from different universes. The PPI sample is based on the number of employees, by company, recorded in the Bureau's Unemployment Insurance file, while the export sample uses the value of exports, by company, recorded by the Bureau of Census in its compilation of Shippers' Export Declarations. Furthermore, the programs have different sample rotation strategies. In the PPI, most industries are sampled only every 7 years.[6] In the XPI, each product category is resampled every other year. This faster sample rotation in the XPI could lead to differences in product mix between the two price indexes.

Even in similar product areas drawn in the same reference period, establishments and items selected for collection may vary dramatically between the two programs. For example, it is possible that a major U.S. producer also has production facilities abroad, and therefore may not export a significant

Exhibit 1. Comparing export price index and producer price index methodologies

Survey attribute	Export price index	Producer price index
Sampling		
Universe	Consists of Bureau of the Census trade tapes of all Shippers' Export Declarations filed with the U.S. Customs Service. Companies must be located within the 50 States and U.S. territories.	Consists of the Unemployment Insurance file of U.S. producers. Companies must be located within the 50 States; those located in the U.S. territories are excluded. This difference between the index programs can be critical for such index items as pharmaceuticals, for which much of U.S. production takes place in Puerto Rico.
Classification	Samples are drawn according to the Harmonized System commodity grouping.	Samples are drawn by industry (SIC classification).
Weights	Sampling probabilities are based on value of trade for the most recent 12 months' worth of census data.	Sampling probabilities are based on value of shipments and number of employees in an establishment.
Rotation	Products are sampled for 5 years. In addition, one-half of exports are resampled every other year.	Industries are sampled only once every 7 years.
Initiation		
Out-of-scope	Program has a fairly high out-of-scope rate, attributable in part to sampling by product. (The out-of-scope designation includes items no longer traded, those only sporadically traded, and so forth.)	Program has a fairly low out-of-scope rate, attributable in part to sampling by establishment.
Refusal	Refusal rates are about the same in the two programs.	Refusal rates are about the same in the two programs.
Repricing		
Reference period	Prices are requested for the first week of the reference month.	Prices are requested for the Tuesday of the week containing the 12th of the month.
Usable prices	Until recently, only prices of goods actually exported during the reference period were used in estimation. Intracompany transfer prices may be excluded if they are not influenced by market factors.	Estimated and average prices are used in estimation, in addition to actual shipment prices. Intracompany transfer prices are included.
Price basis	Inland transportation costs may be included.	Inland transportation costs and other intermediary costs usually are not included.
Number of prices	Data are collected for approximately 11,000 items.	Data are collected for approximately 100,000 items.
Estimation		
Stratum and group weights	Upper-level weights are updated every 5 years and are based on the value of U.S. exports for a given category. Currently, the program uses 1995 weights.	Upper-level weights are updated every 5 years and are based on gross census weights from the Census of Manufactures. Currently, the program uses 1992 weights.
Item weights	For data prior to January 1997, products were equally weighted at the item level.	At the item level, product weights are a function of the establishment's shipments in a given industry.

dollar value of goods, while a minor U.S. producer might export a larger dollar value in the same product category. Therefore, an establishment's probability of selection will differ between the two programs. In addition, for any given industry, the smaller the proportion of total U.S. production that is exported, the more likely it is that the PPI sample will not include exporters or exported products. The establishments eligible for selection also will differ because the PPI samples only producers in the sampled SIC, while the XPI includes in its sample brokers, freight forwarders, and any other establishments that complete export documentation. Furthermore, the PPI does not consider Puerto Rico, the Virgin Islands, and other U.S. territories as part of the United States, while the XPI does include them. For example, a good produced in Puerto Rico and shipped to Europe would be in scope for the XPI but not for the PPI.

Initiation. Although the two programs use a similar PPS disaggregation procedure for the selection of the actual items to include in the indexes, items selected for repricing during initiation may differ between programs because the weights and random numbers used in the PPS will differ.[7] In particular, for a given company, the PPI attempts to collect prices based only on the final stage of processing for each item. For example, if a semiconductor company internally produces silicon wafers (the basic building block of a semiconductor), which are then fabricated into microprocessors, the PPI does not attempt to price the wafer.[8]

The XPI also has a higher percentage of its sample that, for one reason or another, is not repriceable. For example, if a good is no longer exported, the item would be considered out-of-scope. This may be attributable to the fact that the PPI program tries to initiate pricing at a very broad item level while the XPI program, whose sample is based on much more detailed item information, attempts to initiate fairly narrow product categories.

Repricing. Both programs use very similar procedures for collecting prices from respondents. Similar procedures also are employed when prices must be adjusted to account for changes in quality or in the terms of transaction. However, there are also significant differences. For example, until recently the XPI program did not collect an estimated price if no transaction took place during the month being surveyed. In contrast, the PPI program collects reporters' estimated prices. Furthermore, if the same item had originally been included in both indexes but was no longer being exported, it would continue to be repriced for the PPI, but would be discontinued and no longer repriced for the XPI.

Perhaps more importantly, as mentioned earlier, over one-third of U.S. exports are intracompany transfers, shipments from one arm of a company to another. Economists have theorized, and have produced supporting evidence, that price setting by multinational companies for what are referred to as transfer prices is based on criteria very different from those applied to export transactions among nonrelated parties.[9] The PPI also includes some intracompany prices, but they appear less important in index construction, in part because the PPI program is primarily interested only in prices that represent the final selling price of the good.

In theory, the larger size of the PPI data base (there are approximately 100,000 items in the PPI, versus 11,000 in the XPI) will result in a less volatile index for any comparable index or product grouping. At the detailed level of some selected product areas (such as specific grains and similar commodities), however, there might be *more* volatility in PPI indexes due to the requirement that the price reported reflect the price on the specific reference day, Tuesday of the week containing the 12th of the reference month. This is in contrast to the international price program policy of using a weekly average for the reference price for certain product areas. The difference in reference points may also result in lags in trends between equivalent indexes of the two programs.

Different index levels may also be observed when comparing otherwise equivalent indexes because the value of shipping is included in XPI transactions but is excluded from the comparable PPI indexes.

Estimation. Both the PPI and XPI programs use a Laspeyres formula to calculate price indexes. There are, however, significant differences in the formula. Prior to January 1997, items in the XPI indexes were equally weighted at the cell level (the lowest level of aggregation).[10] In contrast, in the PPI, the item weights at the lowest level are a function of sample weights and the item selection process. The programs do use a similar cell-mean estimation technique to account for missing prices.[11] Recall, however, that the PPI attempts to collect an estimate from the respondent whenever no transaction has occurred during the reference period, while the XPI methodology, until recently, would not. Instead, it would impute the movement based on the cell-mean. Above the cell level, the XPI is weighted by export trade flows while the PPI is based on estimates of domestic shipments. Therefore, relative importances will vary because a product category's share of domestic production may be very different from its share of gross exports. Furthermore, while both programs update these aggregate weights every 5 years, the reference periods do not match; the export price indexes used in these comparisons are based on 1985 or 1990 trade weights, while the PPI uses weights based on either the 1982 or the 1987 Census of Manufactures.

Index comparison

Comparison using published PPI indexes. In the first phase of this analysis, a representative sample of 10 (out of a possible 130) of the most detailed levels of published export price indexes were chosen to represent each of the five major end-use groups (raw materials, consumer goods, and so forth), as well as a mixture of homogeneous and heterogeneous categories (for example, wheat versus fruits and vegetables.) These groups were then matched up with the published PPI series that most nearly corresponded to them.[12] For example, the published export price end-use index 12540, Industrial organic chemicals, was matched up with the PPI commodity group 0614, Basic organic chemicals. In two cases—Fruit and fruit preparations and Other industrial machinery—no individual PPI group would properly match up. For these cases, two PPI groups were equally weighted together in order to produce a proxy index. The 10 five-digit end-use groups accounted for roughly 18 percent of U.S. exports in 1990.[13]

Because the international price index program did not begin collecting *monthly* data until 1989, comparisons were undertaken using two different time frames, one using monthly data from September 1989 through March 1995, and the other using quarterly data from March 1985 through March 1995.[14,15] Comparisons were made of both the short-term and the long-term similarities between matches, because both are crucial in assessing the accuracy of a PPI proxy.

A graphical comparison of the export price index with its PPI counterpart can be found in charts 1 and 2. The PPI comparison index is labeled method 1. A review of these graphs highlights a number of instances in which the PPI series appears to differ significantly from the associated export price index over the long term. For example, the PPI index for Medicinal and pharmaceutical materials appears to rise much *more* rapidly over the 10-year period than the comparable export series. As a result, while real exports of medicinal and pharmaceutical materials (based on the export price index) grew 150 percent from 1985 to 1995, an estimate using the PPI index would show an increase of just 80 percent. In contrast, the PPI index for Measuring and testing instruments and related products rises much *less* rapidly during this time frame than does the export series, a pattern also repeated for Iron and steel products and for Household kitchen appliances. The estimated increase for real exports of Measuring and testing instruments based on the export price index is 54 percent, while the estimated increase using the PPI series would be 75 percent. In two additional areas, Fruit and fruit juices and Organic chemicals, the trends are similar but the short-term changes in the export price index are of a significantly greater magnitude than those in the comparable PPI proxy. A more rigorous statistical comparison of these series will be undertaken in a later section.

Comparison using PPI price data aggregated using the export price index market basket. In a limited study such as this one, it is difficult, if not impossible, to assess what role each of the specific factors discussed above may contribute to any differences between an export price index and its PPI counterpart. However, it is possible to delineate and even adjust for some of the variation. In particular, differences between these series can be decomposed into two components: a) divergence due to differences in the relative importance of a given commodity in the market basket, and b) divergence due to differences in domestic and export price trends for a given commodity. An analysis of this decomposition entails constructing a price index that combines PPI price data with relative importances based on export trade weights. The difference between this proxy PPI and the export price index would be attributable to differences in trends in the underlying price data.

The construction of a perfect and complete set of proxy PPI indexes for this study was not feasible for two reasons: first, to do so would entail manually classifying up to 80,000 PPI products and second, even at the lowest (or elementary) level of aggregation, there may be differences in the underlying product mix that are actually differences in relative importances of individual commodities. Instead, in order to attempt to control for differences in relative importances, this study tested three alternative methods for constructing export price indexes using PPI data and export weights. The methods used to calculate indexes differ primarily as to the level of PPI data (for example, an individual item, or so-called cell-level data) used to construct a proxy. Comparisons were undertaken for the same 10 product categories that were matched earlier.

Three proxy methods were developed, as listed below from the easiest to the most difficult to construct and maintain. In theory, this order of presenting the methods should also represent progressively better matches with the export price index by minimizing weighting differences. (Note that method 1 is the aforementioned direct comparison with the published PPI index.)

- *Method 2—PPI 'cell' to XPI 'classif group.'* Both the PPI and the XPI aggregate price relatives from individual items into detailed product groups. In the PPI program, these groups are called PPI 'cells,' while in the XPI program, they are referred to as 'classif groups.' Method 2 entailes matching up the PPI cells to the export classif groups.

  For this comparison, industry analysts made this match for the same 10 end-use groups identified in the preceding set of tests. The number of PPI cells that had to be matched ranged from five for wheat to several hundred for some of the more heterogeneous product areas. Many of these matches resulted in the mapping of more than one PPI cell group into a given export classif group, as well as the opposite case—one PPI cell mapping into a multitude of export classif groups. In these cases, the PPI cells mapped to a given export classif group were averaged together to construct a proxy export classif group index.[16] These proxy groups were then averaged together using the same weights employed in the export price index.

- *Method 3—PPI 'product' to XPI 'classif group.'* This method required the use of unpublished, item-specific data. Each of the individual PPI items was mapped directly to the export classif groups. Through the use of a simple average of the price relatives for all of the products mapped to each classif group, a long-term relative was constructed for each group, and these 'proxy' classif groups were aggregated together using export trade weights. If the PPI-cell-to-classif-group mappings in method 2 had all been exact one-to-one matches, the results obtained under method 3 would be nearly identical to the results under method 2. Given the numerous multiple mappings uncov-

Chart 1. **Comparison of export price indexes with PPI-based proxy price indexes using methods 1 and 2, quarterly data**

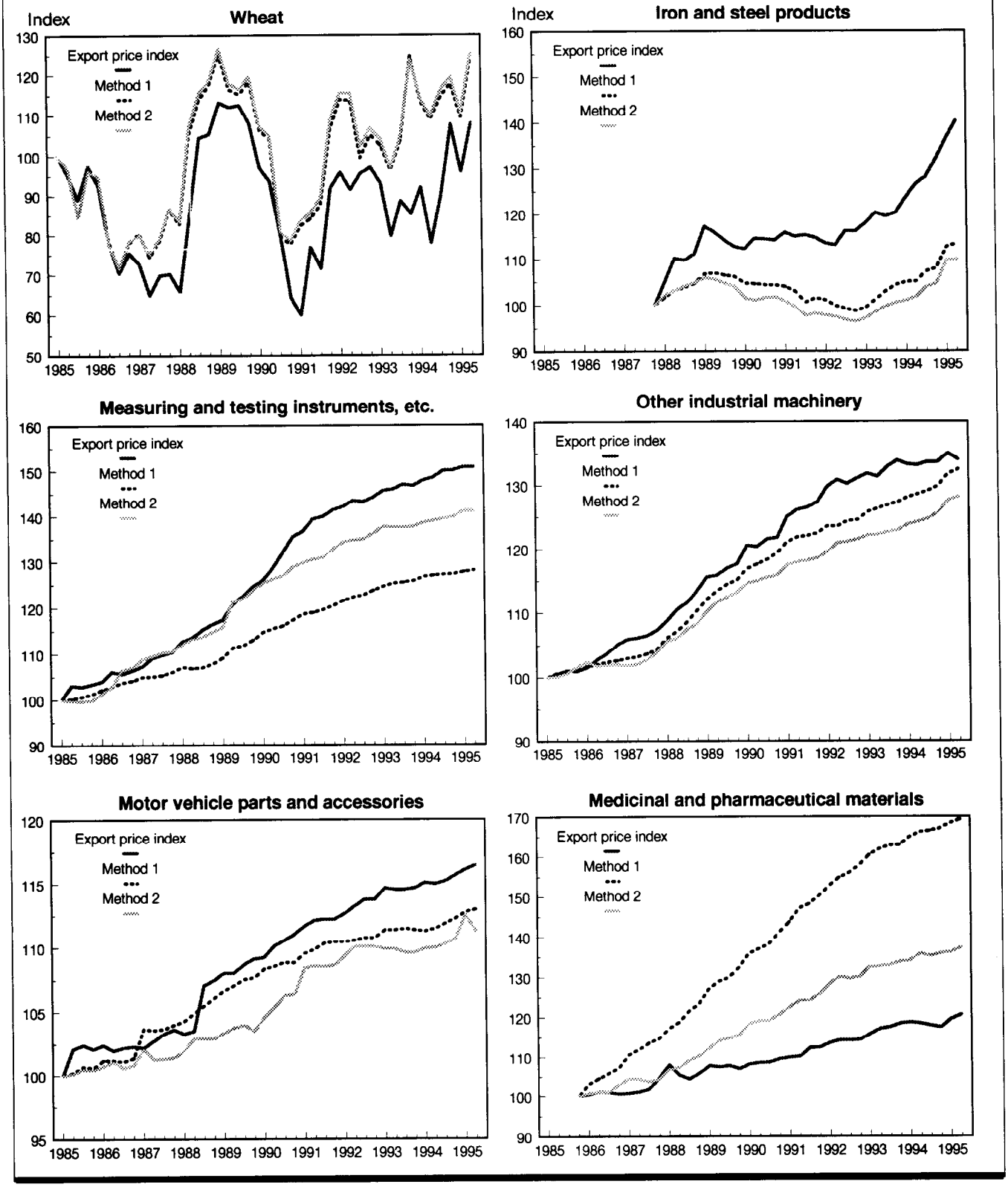
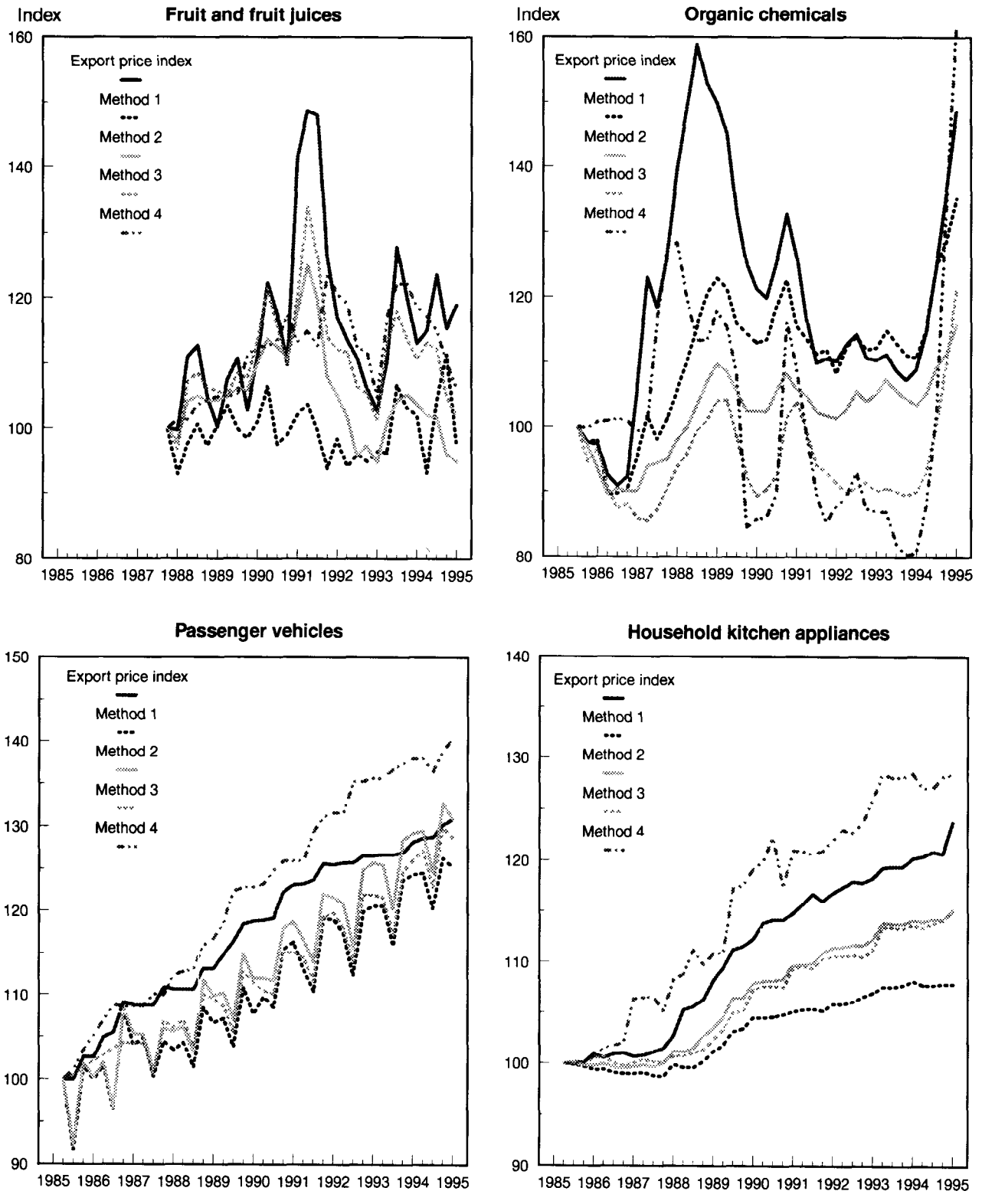


Chart 2. **Comparison of export price indexes with PPI-based proxy price indexes using methods 1, 2, 3, and 4, quarterly data**



ered in method 2, however, some of the indexes based on method 3 produced very different results.

Due to the extremely labor-intensive nature of this matching process (which requires that individual PPI item specifications be reviewed and classified), the method 3 procedure was conducted for only 4 of the 10 product categories covered by this study.

- *Method 4—PPI 'export product' to XPI 'classif group.'* At the time of initial contact with U.S. companies, representatives of the producer price program select individual products to be repriced on a regular basis. One of the items of information that they attempt to collect is whether the prices to be reported reflect domestic prices, foreign prices, or both. Method 4 thus is based on a subset of products from method 3, using only those products whose prices were coded 'foreign' or 'both' during the initiation process. Here, too, comparisons could be run only for the same four product categories tested in method 3. In theory, method 4 should provide the best match between the PPI and the XPI because it goes the furthest in eliminating any differences attributable to differing market baskets. Unfortunately, there are a number of questions regarding the coding and maintenance of these data. For example, the data are collected only during the initial visit to the company, and are not updated during monthly re-collection.

In addition to the comparisons with the published PPI index (method 1), chart 1 also includes quarterly comparisons for the six product groups that have PPI proxy indexes based on method 2, while chart 2 includes the comparisons for the four product groups that have proxy indexes based on methods 2, 3, and 4.[17] These graphs highlight a number of obvious disparities between the export price indexes and the proxy indexes. While in some cases, such as that of Medicinal and pharmaceutical materials, the proxy price index appeared to match up much more closely with the export price index than did the published PPI, in other cases—Iron and steel and Wheat, for example—using export weights did not produce a better match.

Analysis of comparisons

To supplement the graphic comparisons, five tests were run to help evaluate the effectiveness of using PPI data as proxies for export price indexes. Because some uses of these indexes (such as the analysis of U.S. competitiveness) might entail focusing on long-term comparisons of these data, while others (such as examining the impact of exchange rate fluctuations) may be more concerned with short-term trends, tests covered both long-term and short-term comparisons.

These tests were run separately for quarterly as well as monthly comparisons. Exhibits 2a and 2b present the results from the tests for the version of the PPI that appeared to best match up with the export price index, be it the published PPI series or one of the constructed proxies. The determination of the "best" match was based primarily on these statistical tests.[18]

Test 1—The correlation of the short-term relatives. In the measurement of the short-term correlation coefficients in test 1, a value of 1 would indicate a perfect correlation, while 0 indicates no correlation.[19] The median figure for the best proxies for each of the 10 groups was 0.522 on a quarterly basis, and just 0.244 on a monthly basis. Not surprisingly, the correlation values on a quarterly basis tended to be somewhat higher, presumably because there is less "noise" in the data over 3 months than over 1 month. In only one product area, Organic chemicals, were any of the correlations on the monthly percent changes over 0.50. Even the correlations for the groups that graphically appeared to have good proxies were relatively low, 0.29 for Other industrial machinery and 0.07 for Motor vehicle parts and accessories.[20]

Test 2—The frequency with which the sign of the short-term change for the proxy differed from the sign of the change in the export price index. In test 2, a count was kept of the number of times the sign (plus or minus) of the short-term percent change in the export price index was different from the corresponding sign for the PPI proxy index. The results of the sign test were surprisingly high, with signs of the comparison being different between 16.0 and 38.6 percent of the time on a monthly basis (a median of 27.2 percent), and between 13.3 and 26.8 percent of the time on a quarterly basis (a median of 19.5 percent). Method 1, which is essentially a simple comparison of the PPI and XPI indexes, appeared to generate the best fit, producing the lowest median percent of observations in which the signs were different.

Test 3—Five- and ten-year comparisons of the total percent change in the indexes. Test 3 compared the long-term changes in the PPI proxy indexes with those in the export price index.[21] One set of comparisons covered the full period for which data were available—usually 10 years—while another set compared the change for the last 5 years. The results indicate fairly widespread differences in long-term changes between the PPI proxy indexes and the export price indexes. These include several cases in which the PPI proxy and the XPI actually moved in opposite directions, such as the comparisons for Wheat and the 5-year comparison for Fruit and fruit preparations. Differences between the two measures also were significant for categories such as Iron and steel mill products and Medicinal and pharmaceutical products.

Exhibit 2a. Comparison of export price indexes (XPI) with best producer price index (PPI) proxies

Group	Test 1: correlation[1]		Test 2: Sign difference[2] (in percent)		Test 3: Long-term comparison[3] (in percent)				Test 4: regression[4]	
	Monthly	Quarterly	Monthly	Quarterly	10-year		5-year		Monthly	Quarterly
					XPI	PPI	XPI	PPI		
Wheat	0.407	0.657	32.1	19.5	−4.1	+9.4	−1.1	+2.8	0.014	0.404
Fruit and fruit preparations, including fruit juices	.343	.752	38.6	13.3	+19.0	+1.9	+7.0	−10.4	.793	.543
Iron and steel mill products	.212	525	28.4	26.7	+36.4	+12.7	+21.5	+7.6	.149	.066
Industrial organic chemicals	.553	812	32.1	19.5	+51.0	+38.7	+22.6	+19.7	.388	.085
Measuring, testing, and controlling instruments	.044	.269	19.8	14.6	+50.8	+41.2	+19.7	+12.7	.000	.000
Other industrial machinery	.292	.519	25.9	19.5	+35.0	+31.8	+11.9	+12.7	.971	.804
Passenger vehicles	.275	.739	23.5	26.8	+30.9	+32.2	+10.2	+16.9	.000	.000
Nonengine parts and accessories (motor vehicles)	.071	.069	16.0	17.1	+16.1	+12.8	+6.3	+4.0	.197	.004
Medicinal, dental, and pharmaceutical preparatory materials	−.490	.189	34.6	26.3	+19.6	+36.2	+10.5	+15.0	.000	.000
Household and kitchen appliances	.131	.321	21.0	26.8	+24.3	+15.6	+10.2	+6.6	.102	.011
Median	.244	522	27.2	19.5	+27.6	+23.7	+10.4	+10.2	.126	.039

[1] Test 1 was a test of the short-term correlation. A value of 1 indicates perfect correlation, while 0 indicates no correlation.

[2] For test 2, a count was kept of the number of times that the sign (plus or minus) of the short-term percent change in the export price index was different from the corresponding sign for the proxy index. The smaller the percent, the better the proxy.

[3] Test 3 generally compared the total change over a 10-year period (March 1985 through March 1995) and a 5-year period (March 1990 through March 1995) between the export price index and the PPI proxy. Due to data limitations, the comparisons for Fruit and fruit juices and for Iron and steel mill products begin in December 1987; the comparison for Medicinal, dental, and pharmaceutical preparatory materials starts in December 1985; the comparisons for Passenger vehicles and for Household and kitchen appliances begin in June 1985.

[4] For test 4, a regression equation was run, in which the short-term export relatives were the dependent variable and the short-term PPI proxies were the independent variable. An f-test was run with the restrictions that the coefficient on the PPI proxy was set equal to 1 and that the intercept was set equal to zero (the necessary condition for the proxy to be an exact match). Possible values are $0<P<1$. The higher the probability of these conditions not being rejected, the more likely it is that the PPI is a good proxy for the export price index.

Overall, the results of test 3 indicate that the best PPI comparison produced, on average, a slightly slower pattern of price rises over the long term than did the export price index.

Test 4—A regression equation with short-term export relatives as the dependent variable and short-term PPI proxy relatives as the independent variable. In test 4, regressions were run in which the short-term relative of the export price index was the dependent variable and the short-term relative of the producer price index was the independent variable.[22] This phase of the study was designed to give an indication of what would happen if BLS attempted to "model" export price indexes using PPI data. More importantly, each regression also included a test of the likelihood that the PPI series would be a good proxy for the XPI. This was accomplished by examining the possibility that the intercept in the regression was zero and the coefficient on the variable was equal to 1, the results required if the PPI were an exact proxy. The greater the probability that this possibility cannot be excluded, the more likely it is that the proxy index is a valid measure.

More formally:

$$\Delta EPg_i = B0g + B1g\Delta PPIg_i$$
$$H0:\ B0 = 0, B1 = 1$$

where,

ΔEP = monthly (or quarterly) percent change in export price;

ΔPPI = monthly (or quarterly) percent change in producer price;

g = product group; and

i = ith month (or quarter).

Exhibit 2b. Comparison of relative differences between export price indexes and the best producer price index (PPI) proxies—test 5

Group	Median[1] (in percent)		Standard deviation (in percent)		Test for unit root	
	Monthly	Quarterly	Monthly	Quarterly	Monthly	Quarterly
Wheat	4.7	10.4	8.4	9.8	Reject at 10-percent level of significance.	Cannot reject at 10-percent level of significance
Fruit and fruit preparations, including fruit juices	4.6	3.3	9.4	5.7	Reject at 5-percent level of significance.	Cannot reject at 10-percent level of significance
Iron and steel mill products	6.5	11.7	4.8	5.4	Cannot reject at 10-percent level of significance.	Cannot reject at 10-percent level of significance
Industrial organic chemicals	28.8	4.4	8.3	9.9	Reject at 5-percent level of significance.	Cannot reject at 10-percent level of significance
Measuring, testing, and controlling instruments	4.3	3.4	2.6	2.9	Cannot reject at 10-percent level of significance.	Cannot reject at 10-percent level of significance
Other industrial machinery	.7	2.8	1.1	1.5	Reject at 10-percent level of significance.	Cannot reject at 10-percent level of significance
Passenger vehicles	4.2	2.3	2.7	2.9	Cannot reject at 10-percent level of significance.	Reject at 5-percent level of significance
Nonengine parts and accessories (motor vehicles)	.8	1.6	1.0	1.3	Reject at 5-percent level of significance.	Cannot reject at 10-percent level of significance
Medicinal, dental, and pharmaceutical preparatory materials	7.4	9.3	3.0	5.1	Cannot reject at 10-percent level of significance.	Cannot reject at 10-percent level of significance
Household and kitchen appliances	1.2	4.5	.9	2.1	Reject at 1-percent level of significance.	Cannot reject at 10-percent level of significance
Median	4.4	3.9	2.9	4.0		

[1] Median of the absolute value of the relative differences between the export price index and the PPI proxy index for each period.

As reflected in the comparatively low median estimates, this test suggested that the PPI proxy was probably not a very good fit in the majority of instances. Some results, however, appeared to be more positive. For example, the regression of monthly data for end-use category 21180—Other industrial machinery—produced a coefficient of 1.020 and an intercept of 0.000. Even this match is suspect, however, in light of the extremely low R-square value, 0.085. Overall, no proxy appeared to produce a better fit than the others. The method 4 proxy, which presumably should have produced the best fit because it was limited to prices that the companies indicated during the initiation process were suitable as export prices, actually appeared to generate the weakest matches under this test

Test 5—A test of the relative differences between each pair of long-term index relatives. The set of statistics developed in test 5 is based on the construction of the differences between the export price indexes and the corresponding PPI index.

More formally:

$$Zg_i = {}_{\text{Log}}EPg_i - {}_{\text{Log}}PPIg_i$$

where,

EP = monthly (or quarterly) export price index;
PPI = monthly (or quarterly) producer price index;
g = product group; and
i = ith month (or quarter).

Each data point in the new series (represented by Zg_i) represents the percent difference between the export value and the PPI proxy.[23] For each set of comparisons, two statistics are presented in exhibit 2b. The first is an estimate of the absolute value of the median difference for each data point in the series.[24] The second is an estimate of the standard deviation of the series. For example, in the first category (Wheat), each of the 42 PPI quarterly index values is, on average, 10.4 percent different from (either higher or lower than) the comparable export price index. In addition, the standard deviation is fairly broad, 9.8 percent. This indicates that only approximately two-thirds of the 42 quarterly differences are within a 19.6-percent (9.8 percent times 2) range, centered on the average.

On a quarterly basis, the median estimated difference for all 10 groups is 3.9 percent, while the comparable monthly figure is 4.4 percent. This implies that, for any given period, the PPI proxy would differ from export price index by around 4 percent. Overall, however, there are significant differences in the distributions. For example, the median difference for the volatile Industrial organic chemicals index on a monthly basis is 28.8 percent, while the comparable figure for the comparatively flat Nonengine motor vehicles parts and accessories is just 0.8 percent.

One indication that a PPI proxy is an appropriate match for an XPI index would be a finding that each of these difference series (Zg) is stationary. A series is said to be stationary if its properties (mean and variance) are constant over time. If the condition for stationarity is satisfied in the case at hand, the series of differences between the PPI and the XPI would revert to the mean over time, with no tendency for the variance to increase or decrease. Ideally, if the proxy is good, the mean is zero. This would serve as an indication that any differences between the two series are likely to be random and not related to underlying differences in what is being measured.

The specific test used to check for stationarity in this study is the Augmented Dickey-Fuller (ADF).[25] This is one type of what is called a "unit root" test.

More formally:

$$Z_t = \beta_0 Z_{t-1} + \beta_1 \Delta Z_{t-1} + \beta_2$$

where Z is the relative difference between the export price index and the PPI for each product category for each period.

This relationship can be rewritten as:

$$\Delta Z_t = \gamma Z_{t-1} + \beta_1 \Delta Z_{t-1} + \beta_2$$

where $\gamma = \beta_0 - 1$.

The null hypothesis for the test for a unit root is H_0: $\gamma = 0$, which would indicate that the series is nonstationary. A rejection of the null hypothesis would indicate the series is stationary. An additional coefficient, β_1, is added in order to adjust for the fact that the two index programs collect data at different times during the month.[26] A constant, β_2, also is added.[27] The results are reported in exhibit 2b. Note that, on a quarterly basis, the possibility of a unit root (and, therefore, a systematic difference in trend between the XPI and the PPI) could not be rejected at even the 10-percent level of significance for 9 of the 10 comparisons. Interestingly, on a monthly basis, this held true in only four of the cases. This finding may be related to the strong possibility that systematic differences between the series are more likely to show up over a longer period.[28]

Summary and conclusion

This limited study was established to assess how well recent trends for U.S. export price indexes track with U.S. domestic price indexes for comparable product categories. Given some basic similarities in the supply and demand characteristics underlying these data, it was expected that some of these indexes would track fairly closely. By combining the PPI data with export trade weights, the study also attempted to evaluate any differences between the PPI and export price indexes and to estimate what proportion could be attributable to differing market baskets, and what was due to differences in the underlying price trends. The results highlight significant differences between the two indexes and, thus, the limitation of using PPI data to proxy export price indexes.

Exhibit 3 provides a cross-reference of study findings by end-use product group, description, PPI code, and proxy method tested.[29] It also highlights the PPI proxy that appears to best match up with the export price index, and discusses the accuracy of the match. In the majority of instances, the best PPI proxy for an export price index appears to be significantly different from the actual export price index, regardless of how the PPI data are manipulated. For example, in two cases, the best PPI proxy demonstrates a significant long-term deviation from the export price index, one upward (Medicinal and pharmaceutical materials), and one downward (Iron and steel mill products). In two groups (Fruit and fruit juices and Organic chemicals), the short-term swings in the export price index tend to be much sharper than those in the PPI proxy even when the proxy is based on export weights. In two other instances (Measuring and controlling instruments and Household

Exhibit 3. Summary of the results of comparisons of export price indexes with PPI proxies

Group	Export code (end-use)	PPI code (w= commodity; p= industry)	Method 1 proxy: stratum	Method 2 proxy: cell	Method 3 proxy: product	Method 4 proxy: foreign price	Analysis of proxy[1]
Wheat	00000	w0121	X	X			In this homogeneous and volatile category, the PPI proxy index tended to track the export price index fairly well, both over the long term and on a quarterly basis. However, on a monthly basis, there were significant differences, particularly in the magnitude of the short-term changes. This was primarily attributable to subsidies given to wheat exports, which vary month to month.
Fruit and fruit preparations, including fruit juices	00320	w0111, w0242	X	X	X	X	The best PPI proxy follows the general trend of the export price index and is a considerable improvement—particularly in the short term—over the published PPI index. Nonetheless, the magnitudes of the short-term changes, both positive and negative, are significantly greater in the export price index than in the proxy indexes. For example, in 1991, the export price index jumped 35 percent while the best PPI proxy rose only 22 percent. Furthermore, on a monthly basis, the short-term relative of the proxy is different from the export price index 40 percent of the time. (Interestingly enough, on a quarterly basis, the signs are different only 13 percent of the time, and the correlation value is extremely high. This implies that there may be some lag effects in the timing of the data collection that cause some of the differences in short-term trends of these indexes.) The regression test lends some support for the feasibility of using the PPI as a proxy for an export price index.
Iron and steel mill products	12100	p3312	X	X			In producing one of the worst long-term matches in this study, the PPI proxy index appears to be consistently and substantially biased downward relative to the export price index. Between 1987 and 1992, the actual export price index increased at a rate 3 times faster than that posted by the PPI proxy. Because the proxy index, which was designed to use export weights, did not match appreciably better than the straight comparison (method 1), which uses domestic weights, it is not likely that weight differences are the cause of the differences between the export price index and the PPI index.
Industrial organic chemicals	12540	w0614	X	X	X	X	Over the long term, the best PPI comparison roughly matched the export price index and the correlation values were very high. However, the magnitude of differences in short-term trends was often substantial enough to make the proxy comparatively useless. For example, between 1986 and 1988, the export price index jumped 72 percent while the largest increase in a proxy index was just 27 percent. In addition, on a monthly basis, the sign of the best proxy differed from the sign of the export price index one-third of the time. Here, too, the published PPI index matched up as well as or better than the proxy indexes, indicating that differences in the series were not attributable to weight differences.
Measuring, testing, and controlling instruments	21160	p382	X	X			This is an example of a comparison in which the accuracy of the proxy changed over time. From 1985 through 1990, the proxy did a fairly good job of mirroring the export price index. However, from December 1990 through September 1991—a time during which the dollar depreciated substantially—the export price index increased 11.1 percent while the proxy index rose only 4.5 percent, indicating a serious downward bias in the PPI proxy. In addition, the short-term correlation is worse than average. Differences in market baskets appeared to account for approximately one-half of the long-term differences between the export price index and the published PPI series.

Exhibit 3. Continued—Summary of the results of comparisons of export price indexes with PPI proxies

Group	Export code (end-use)	PPI code (w= commodity; p= industry)	Method 1 proxy: stratum	Method 2 proxy: cell	Method 3 proxy: product	Method 4 proxy: foreign price	Analysis of proxy[1]
Other industrial machinery ..	21180	w114, w1166	**X**	X			Surprisingly, despite the likely heterogeneity of this group, the published PPI series for Other industrial machinery provided the closest approximation to the comparable export price index of any comparison in this study. From 1985 through 1995, the export price index rose 34 percent while the published PPI index advanced 32.5 percent. This match was actually closer than the comparison of the export price index with the PPI proxy index. The regression test also signified a comparatively good fit. Despite this good long-term match, however, the percent changes on a monthly basis moved in different directions 26 percent of the time.
Passenger vehicles	300	p3711	X	**X**	X	X	This comparison produced an excellent long-term fit, but a very poor short-term match. This was attributable to the seasonal pattern in the PPI data, which manifests large price drops every September coinciding with price declines associated with model yearend close-outs. This pattern does not hold, however, for export prices. The method 4 proxy, which does not make use of prices coded "domestic only," and which excludes major U.S. automobile manufacturers, does not demonstrate this seasonality. It does not, however, produce a good long-term fit.
Nonengine parts and accessories (motor vehicles)	30230	p3714	**X**	X			Over the long term, prices of exported parts tended to rise about 1.6 percent per year, slightly more rapidly than the PPI index. (However, much of the difference occurred in just one period, third-quarter 1988.) In addition, the short-term correlation figures were extremely low. (This may be due in part to the relatively small changes that occur each month.)
Medicinal, dental, and pharmaceutical preparatory materials	40100	w063	X	**X**			The PPI proxy made some improvement in matching the PPI index to the export price index, indicating that differences in market basket may have accounted for one-third of the difference between the export price index and the published PPI series. Nonetheless, the proxy index still overstated the long-term rising trend in export prices by a factor of nearly 2, indicating that the PPI proxy is seriously biased upward. The short-term correlation values were also extremely low, consistent with the results of the regression test.
Household and kitchen appliances	41030	p363	X	**X**	X	X	The best PPI proxy index did produce a significantly closer fit with the export price index than that noted for the published PPI index, indicating that approximately one-half of the long-term difference between the export price index and the published PPI index was attributable to differences in the market basket. However, the proxy index still showed an overall increase from 1985 to 1995 that was approximately 30 percent less than the increase posted by the export price index. Here, too, the accuracy of the proxy appears to vary over time, with most of the discrepancy occurring between 1985 and 1988. (Note that the method 4 proxy, which used only products coded as valid for export, appeared to match fairly well in the long term, but, due to its short-term volatility, was not a good match overall.)

[1] The best XPI-PPI match is indicated by boldface X. The best match was primarily determined by the results of the five statistical tests.

kitchen appliances), the proxy appears to match up well for half of the period under review, but not very well for the other half. In yet another comparison (Passenger vehicles), the best proxy manifests sharp seasonal fluctuations that do not occur in the export price index. In some cases, up to one-half of the difference between the published export price index and the published producer price index is attributable to variations in market basket. In five matches, however, it does not appear that constructing a proxy index produces any better results than a direct comparison with the published PPI index (method 1). This latter point argues against variations in market baskets as the primary reason for the discrepancies between the indexes, and is evidence that much of the differences in export price indexes and producer price indexes may be related to underlying differences in the individual products or prices, or both, used in the two surveys. Even a homogeneous product area such as wheat demonstrated some difference between the export price and the domestic price. All of this leads to the conclusion that there is no practical method for constructing export price indexes using PPI data.[30]

These findings may not be too surprising, given the number of conceptual and methodological differences in the design and implementation of the two indexes. In fact, it was similar research showing the limitations of using wholesale prices to describe international prices that led BLS to establish the International Price Index Program a quarter of a century ago. In their seminal 1971 work on the price trends of exports, entitled *Price Competitiveness in World Trade*, Irving Kravis and Robert Lipsey found that "the major discrepancies between domestically weighted wholesale price indexes and the corresponding international price indexes with international weights are not mainly the result of weighting differences but rather differences in the price movements reported for individual commodities and in the samples of commodities."[31] It appears that the value of relying upon domestic U.S. prices to track export prices has not appreciably improved during the ensuing period. □

Footnotes

[1] The export indexes do, however, include nonmanufacturers such as brokers and freight forwarders, which are not included in the PPI.

[2] In the case of price indexes, this assumption can actually be modified to require simply that the price *trends* of identical goods in different markets be the same.

[3] For a summary of some of these studies, see Ken Rogoff, "The Purchasing Power Parity Puzzle," *Journal of Economic Literature*, June 1996, pp. 647–88. In addition, Irving Kravis and Robert Lipsey found evidence in their survey that "prices of a substantial fraction of international trade in manufactured metal products and machinery differ from those in domestic markets." See Irving B. Kravis and Robert Lipsey, *Price Competitiveness in World Trade* (New York, National Bureau of Economic Research, 1971.) There is, however, some counterevidence supporting the idea that U.S. firms do not vary price by market (at least not as much as foreign producers do). Firms currently in the PPI sample generally report that they charge the same to domestic and foreign buyers. Unfortunately, companies in the export sample are not asked the same question, which would have provided a crosscheck on the PPI experience.

[4] While this analysis covers the period of the dollar's long-term depreciation that began in 1985, it does not include the dollar's sharp appreciation from 1981 to 1985, making an assessment of the impact of the exchange rate on these comparisons difficult.

[5] The PPI data also are rearranged into a product-based classification system to construct (in conjunction with the Consumer Price Index) the implicit price deflator for the private consumption expenditures, gross private investment, and government purchases components of final demand.

[6] The PPI does attempt to re-sample high tech industries more frequently.

[7] In addition, because both programs are voluntary, during initiation (which is usually done via a personal visit from a Bureau employee), a given company may respond differently to each program, refusing to cooperate with one while supplying data to the other.

[8] In theory, because the PPI is defined as a series of industry net output indexes, the indexes should also exclude shipments between different companies that are classified within the same 4-digit SIC. In practice, however, the PPI does not find it feasible to exclude items that are shipped between separate companies. Instead, the weights of the net output indexes are adjusted to exclude the value associated with these intra-industry shipments.

[9] See for example, Subramanian Rangan and Robert Lawrence, "The Response of U.S. Firms to Exchange Rate Fluctuations: Piercing the Corporate Veil," *Brookings Papers on Economic Activity*, no. 2 (Washington, The Brookings Institution, 1993), pp. 341–79. In fact, the authors argue that export prices among multinationals are more likely to be based on sourcing considerations than on demand considerations. This is backed up, in part, by data indicating that most U.S. transfer pricing takes place at cost valuation instead of the final selling price. See Mohammad Al-Eryani, Pervaiz Alam, and Syed Akhter, "Transfer Pricing Determinants of US Multinationals," *Journal of International Business Studies*, March 3, 1990; or Anita Benvignati, "An Empirical Investigation of International Transfer Pricing by US Manufacturing Firms," in Alan M. Rugman and Lorraine Eden, eds., *Multinationals and Transfer Pricing* (New York, St. Martin's Press, 1985).

[10] Beginning with the release of the January 1997 indexes, the item weights used in the XPI are based on the sample weights.

[11] The movement of a missing price would be based on the price trends of the most nearly similar items.

[12] The small size of the sample used in this analysis was primarily attributable to the highly labor-intensive nature of the detailed cross-mapping tasks.

[13] There was also some variation in the proportion of a product category's production that is exported. In some of the test examples approximately 10 percent of a domestic category is produced for export, while in others—Wheat and Measuring, testing, and controlling instruments, for example—the proportion of exports is considerably higher.

[14] An analysis of the quarterly data is more appropriate for assessing how the proxies trend over the medium term, while the monthly comparisons give a better estimate of PPI data as short-term proxies. Quarterly data actually refer to price data collected for the third month of each quarter (March, June, September, and December).

[15] Unfortunately, not all series have complete historical information. In particular, the series for Fruit and fruit juices and for Iron and steel products only go back to 1987.

[16] This also meant that if a given PPI cell was mapped to more than one classif group, it would end up being used more than once in calculating an index. There would be no reapportionment of its weight.

[17] Monthly graphs are available upon request from the author.

[18] For the full table of test results, please contact the author.

[19] Note, however, that the correlation would be 1.0 if the change in the export price were always double or one-half of the change in the proxy variable.

[20] For comparison purposes, the correlation coefficient between the last 10 years of monthly changes for the Consumer Price Index for All Urban Consumers and corresponding data for the PPI for finished goods was 0.57.

[21] This test was similar to the analysis in Kravis and Lipsey, which compared data over the 11-year period from 1953 to 1964.

[22] The author also ran regressions using long term relatives. However, these results were fairly sensitive to the choice of base period.

[23] In each comparison, the XPI and the PPI have the same base period set equal to 100.

[24] The absolute value of the median rather than the actual median is used because the former measure is more appropriate for estimating the accuracy of short-term comparisons

[25] For further discussion of the ADF test and of the software used in the calculations, see *EVIEWS Users Guide* (Irvine, CA, Quantitative Micro Software, 1995), pp. 185–91.

[26] For example, trends in the PPI index may lag trends in the XPI index by a short period because the PPI program gathers prices around the 12th of the month, while the XPI program does so around the 1st of the month.

[27] In theory, the model should not include a constant because the series of differences should tend to fluctuate around zero. The use of a constant, however, serves to minimize the impact of more or less arbitrarily picking the starting reference month in these comparisons.

[28] See John Campbell and Pierre Perron, "Pitfalls and Opportunities: What Macroencomists Should Know about Unit Roots," in *National Bureau of Economic Research Annual 6*, 1991, pp. 141–201. Campbell and Perron state that "for tests of the unit root hypothesis versus stationary alternatives, the power depends very little on the number of observations per se, but is rather influenced in an important way by the span of the data." Recall that the quarterly comparisons tend to begin in 1985, while the monthly comparisons generally do not start until 1989.

[29] Note that, in six instances, the comparison is with two PPI proxies, while in four cases the comparison was conducted with four PPI proxies. For the PPI codes, a 'p' indicates an industry code, while a 'w' indicates a product code.

[30] While this study casts doubt on the feasibility of using PPI data in constructing an export price index, the proper test of the suitability of a PPI proxy index is determining the effect of using it in any economic analysis. For example, would it significantly impact estimates of real exports or real GDP? Would it produce different estimates of price and income elasticities? Would it produce a different estimate of the U.S. terms of trade (that is, the relative price of U.S. exports versus the relative price of U.S. imports)? These questions could only be properly answered in a study that constructed a PPI proxy for the entire range of U.S. exports of merchandise, an undertaking well beyond the scope of this study. It should also be noted that, given the dynamic nature of trade, what might have been an appropriate proxy in the past may not be suitable in the future.

[31] Kravis and Lipsey, *Price Competitiveness in World Trade*, p. 178.