PRICE DISPERSION IN U.S. MANUFACTURING*

by

Thomas A. Abbott, III

CES 89-7 October 1989

Abstract

This paper addresses the question of whether products in the U.S. Manufacturing sector sell at a single (common) price, or whether prices vary across producers. The question of price dispersion is important for two reasons. First, if prices vary across producers, the standard method of using industry price deflators leads to errors in measuring real output at the firm or establishment level. These errors in turn lead to biased estimates of the production function and productivity growth equation as shown in Abbott (1988). Second, if prices vary across producers, it suggests that producers do not take prices as given but use price as a competitive variable. This has several implications for how economists model competitive behavior.

*This paper presents results based on Chapter Two of my doctoral dissertation at Harvard University. The work was carried out at the Center for Economic Studies. The views expressed are attributed to the author and do not necessarily reflect those of the Census Bureau, Rutgers University, or any of the many individuals who commented on earlier drafts of the material. The author would like to thank Zvi Griliches, Richard Caves, Jerry Hausman, Robert McGuckin and Peter Zadrozny for their many useful comments and also Christina Palumbo for her skillful typing.

I. Introduction

Several papers have already been written on the subject of price dispersion. Theoretical explanations are provided by Burdett-Judd (1983), Carlton (1979, 1986), Perloff-Salop (1986), Salop-Stiglitz (1977), and Reiganum (1979). Much of this literature focuses on the incomplete information hypothesis proposed by Stigler (1961), although one can view this as a special case of the product differentiation and monopolistic competition models discussed by Salop-Stiglitz (1977), Stiglitz (1984) and others.

Empirical evidence supporting price dispersion is presented by Dahlby-West (1986), Isard (1977), Pratt et. al. (1979), and Stigler-Kindahl (1970).[1] These studies, however, have limited impact because they focus on relatively few products. Pratt et. al., for example, examin several products in the Boston area, while Stigler-Kindahl examin buyer-seller transactions for a few industrial goods. One cannot generalize these results to determine the extent of price dispersion across industries, or the level of price variation within industries. This paper examines prices for <u>all</u> 7-digit (SIC) products to determine the

---

[1]It is important to note that some of the studies cited are conducted with objectives other than showing the existence of price dispersion and questioning the assumption of a single market price. For example, Stigler-Kindahl's work focuses on obtaining accurate aggregate price deflators arguing that BLS should follow prices from many different producers. They do not focus on an explanation for why different producers received different prices.

extent and level of price dispersion in the U.S. Manufacturing sector.[2]

The outline of the paper is as follows. The next section uses the coefficient of variation and a normalized price range to measure price dispersion. Most products examined have significant levels of price dispersion: 95 percent have a coefficient of variation greater than 16 percent and a price range greater than 68 percent. Unfortunately, these statistics are sensitive to reporting errors and other data anomalies. Section III develops a measure of variation (RD) which is less sensitive to these problems. Under specific distributional assumptions, RD is comparable to the coefficient of variation. Approximately 75 percent of the products examined have RD greater than 10 percent. Although this is substantially lower than the estimated coefficient of variation, it still suggests that there

---

[2]The Standard Industrial Classification system (SIC) was established in the late 1930's to provide a method for the classification and aggregation of industrial statistics in the United States (see U.S. Office of Management and Budget (1972) for additional details.) The SIC system is composed of an ordered number scheme similar to the Library of Congress's classification of published material by subject matter. The first two digits of the SIC code are used to designate major industrial groups (e.g., Textile Mill Products (22) and Stone, Clay, and Glass Products (32)). The next two digits are used to break out specific industries within these major groups (e.g., Cotton Textile Weaving Industry (2211) and the Hydraulic Cement Industry (3241)). Finally, individual products from these industries are given seven digit codes (e.g., Finished cotton Broad Woven Fabrics - Bleached and White Finished (2211711) and Normal Portland Cement ASTM Type I (3241012)). The Bureau of the Census collects some data at the 7-digit product level.

is significant price variation at the 7-digit product level. The final section discusses additional areas of research needed to fully test the unique price theory, and to determine the impact of price dispersion for the analysis of production.

II. Empirical Price Dispersion

The Bureau of the Census collects data on value and quantity of shipments (FOB - plant gate) at the 7-digit product level as part of the Census of Manufactures. Implicit average prices (unit values) are constructed for each establishment-product in the 1982 Census of Manufactures. These prices provide the basis for this analysis.

Starting with 804,757 observations on annual establishment-product shipments, 144,377 observations are found to have usable value and quantity data.[3] For a general analysis of price dispersion across plants, the sample is further restricted to exclude certain types of broadly defined products: the "Not Specified by Kind" and "Not Elsewhere Classified" products.[4] And

[3]In many instances the Census does not collect quantity information due to the heterogeneity of the 7-digit product definition. In addition, to insure that the measured price dispersion is not the result of rounding errors, the sample is further restricted to only those plant-product observations with three or more significant digits.

[4]The SIC classification system is set up so that a seven digit product ending in '00' is generally an NSK (not specified by kind) classification - that is the manufacture did not report the specific product (seven digit code) that was being produced. Rather than contaminate the other data, these observations are

finally, the sample is restricted to only those products with 10 or more establishments.[5]

Imposing these additional restrictions limits the sample to a total of 112,630 establishment-product observations on 2,430 different products. For each product, two statistics are initially used to measure the level of price dispersion. The first is the coefficient of variation (CV), defined as the ratio of the standard deviation to the mean. The second is the normalized price range (RNG), defined as the price range (MAX - MIN) divided by the mean. Summary statistics on the coefficient of variation (CV) and the normalized price range (RNG) for these products are provided in Table 1.

In addition to these summary statistics, Figure 1 provides a histogram of the distribution of the coefficient of variation and normalized price range. From these statistics and figures, it is clear that price dispersion, as measured, is a widespread phenomenon. The average coefficient of variation is 69 percent,

---

pooled into one "general" category. Administrative Records are also frequently included in this NSK classification.

Products ending in a '98' or '99' are generally NEC (not elsewhere classified) product classifications. These products typically include a mixture of highly specialized products which get lumped together for purposes of data collection.

For an examination of price dispersion across "homogeneous" products, the author felt the NSK and NEC products clearly inappropriate for the analysis and could bias the results.

[5]In order to accurately measure the amount of price dispersion it is necessary to insure that there were more than just a few producers of the good.

the minimum 0, and over 95 percent of the products examined have more than 16 percent variation. The normalized price range provides a similar picture, the average is 398 percent, the minimum is 0 and over 95 percent have more than 68 percent price range.

It is also clear that there is a large tail to the distribution of price variation, as measured by the coefficient of variation and normalized price range. The maximum coefficient of variation is 1712 percent, and over 5 percent of the products have more than 226 percent variation. Similarly, the maximum price range is 42,429 percent and over 5 percent have more than 1300 percent price range. These latter results clearly indicate that something is wrong with these measures of price dispersion -- one would not expect any product to have such a large variation in prices across producers. Either there are many poorly defined products or there are a significant number of errors in this data.

III. A Robust Measure of Dispersion

Examination of this price data for a single industry -hydraulic cement, see Abbott (1988), reveals that the Census value and quantity data contain two types of errors which would affect our measure of dispersion: gross outliers and imputed data.

The first error is that of gross outliers; that is data, which could not be reported correctly. For example, in an industry with 60 producers, 59 producers sold the product at a price between $20 and $40, while the remaining producer sold at a price of $250. Clearly there is a units problem with the reported data. This type of error occurs in about 2 to 3 percent of the data and is most likely the result of reporting or keying errors.

The second type of data error is caused by imputations, observations with missing data for which the Census Bureau imputed a value based on the industry averages. Census data are collected, edited, and maintained for the purposes of constructing accurate aggregate statistics; not for the purpose of microeconomic analysis. As such, audit trails to specific microdata items are not well maintained and it is not possible for much of the historical data to determine if a particular observation is imputed, edited, or is the original reported data.[6]

One method of dealing with the problem of gross data errors is to "clean" the data and remove the individual observations which are deemed erroneous. With over 112,000 observations, this is not a simple task. Moreover, if one attempts to clean the

---

[6]In the vernacular of the Census Bureau, an impute is a computer generated value based on a key ratio and the current "hot-deck". An edit, on the other hand, is a replacement value provided by the industry analyst and may come from one of several sources including follow up calls and/or analysts estimates.

data by eliminating observations which are more than X standard deviations from the mean, it would necessarily bias the estimated measure of dispersion downwards. The estimated standard deviation of the truncated distribution is not an unbiased estimate of the true distribution.

An alternative method for handling the problem of gross errors is to use statistics which are more robust to the presence of gross errors, as discussed by Hempel et. al. (1986) and Abbott (1989).[7] Under additional assumptions about the distribution of the true prices, one can use order statistics to obtain robust estimates of the mean and standard deviation needed to construct the coefficient of variation. For example, under the assumption that the distribution is not skewed the median provides a robust measure of the mean.[8] Under the assumption that the true distribution is normal the inner quartile range is approximately

---

[7]In general, a robust statistic is a measure which is not greatly influenced by small deviations in the basic assumptions. In this context, we are looking for a measure of the dispersion which is not sensitive to the magnitude of the gross errors.

[8]The median is robust in that given a sample of observations, adding an erroneous observation to one of the tails results in only a small bias in the measure, for example, moving from the 50th to the 51st percentile. More importantly, the extent of the bias does not depend on the size of the error. The erroneous observation could be 1 standard deviation or 100 standard deviations and the impact on the median is the same. Clearly the impacts on the mean from such errors are very different.

1.348 standard deviations.[9]

The proposed robust statistic is the ratio of the inner quartile range to the median, properly scaled to be comparable to the coefficient of variation.

$$(1) \qquad RD = (Q_3 - Q_1) / (1.348 * Q_2)$$

Using this robust statistic (RD) we reexamine the Census data. Table 1 presents the comparable summary statistics for the RD measure. The average level of dispersion falls to 55 percent, the minimum is 0 percent, and over 75 percent of the products have more than 10 percent price variation with the robust statistic. Although these statistics are substantially lower than those found with the conventional coefficient of variation, they still suggest that the unique price theory does not apply to most of the 7-digit products examined. A similar picture is presented in Figure 1, where the entire distribution of the measure is shifted to the left in comparison with the coefficient of variation.

As shown in the figure, there is still a significant tail to the distribution of dispersion across producers, with 5 percent of the RD statistics being greater than 90 percent. As derived,

---

[9]The finite sample properties of the median and inner-quartile range as measures of central tendency and variation are examined in Abbott (1989).

the robust statistic is not impervious to gross errors in the data. That is, if there is a sufficient number of errors in a single product grouping the RD measure will not provide an accurate measure of the underlying parameters. Thus, the tail implies that in some cases the data are just too dirty, or that the products are poorly defined for even the robust measure.

Table 2 presents a list of the 20 products with the most dispersion, as measured by RD. From this list, it appears that these products are a mixture of "other" and "NEC" (Not Elsewhere Classified) products which did not conform to the usual principles used in defining the 7-digit SIC codes. Thus, it is not surprising that there should be such a wide range of reported prices for these product classifications. This evidence suggests that the tail to the distribution of observed dispersion is caused by a failure of the product definition rather than dirty data.

The final table examines price dispersion across two digit industrial groups using those products with less than 80 percent dispersion (RD). This truncation is used to remove the affects of the outlier products and leaves a 2,278 products for the analysis. From Table 3 it is clear that although the average level of price dispersion differs widely across the major groups, price dispersion is a general phenomenon regardless of which measure is used. Thus the results presented in Table 1 are not dominated by any particular industrial group.

IV. Conclusions

The analysis presented here establishes two important empirical facts concerning price dispersion. First, measured prices vary across producers, even at the 7-digit product level. Second, price dispersion is not isolated to a few manufacturing industries but exists, to some extent, in all industries. The analysis establishes these results using all of the product data available in the 1982 Census of Manufactures. These facts run counter to the assumptions of homogeneous goods and perfect competition usually made in analyzing economic behavior at either an industry or firm level.

The basic data used to arrive at these conclusions, however, suffer from two types of errors: gross outliers and imputed data. The gross errors bias the estimated dispersion upwards and are addressed through the use of robust statistics. The imputations, on the other hand, bias the estimated dispersion downwards and their affects cannot be easily eliminated from the data.[10] As a result, the current analysis only provides information on whether or not individual products exhibit price dispersion and does not provide reliable comparisons of the level of price dispersion across products or industries.[11]

---

[10]Some information on the dollar value of the imputations is o6btainable from the individual industry summaries. This information is not incorporated in the current study.

[11]A casual examination of the data reveals that the most important determinant of the level of price dispersion found in

Having established these facts, one must proceed to address two additional questions: What is the underlying source of the observed price dispersion? and does price dispersion imply market power? For the hydraulic cement industry (near the median level of price dispersion), this additional work has been addressed in Abbott (1988). The basic conclusions of that study suggest that price variation is real, i.e., it does not reflect differences in the product quality; the price variation is due to local conditions in both the output and input markets; and that manufactures do possess market power. Considering the industry, these findings are not surprising.

---

the data is the extent of imputations. For products with more than a 50 percent imputation rate, the robust measure of price dispersion would be zero as both the first and the third quartile would be imputed values.

Figure 1: Measures of Price Dispersion

```
PERCENTAGE

     |                                                        **
  14 +                                                        **
     |                                                        **
  10 +                ** ** **                               **
     |             ** ** ** ** ** **                         **
   6 +          ** ** ** ** ** ** ** **                      **
     |          ** ** ** ** ** ** ** ** ** ** **             **
   2 +    ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** **  **
     -----------------------------------------------------------
          5   15    25    35    45    55    65    75    85  95

                     COEFFICIENT OF VARIATION
```

```
PERCENTAGE

  20 +                                                         **
     |                                                         **
  15 +                                                         **
     |                                                         **
  10 +                      **                                 **
     |             ** ** ** ** ** **                           **
   5 +          ** ** ** ** ** ** ** ** **                     **
     |   ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** **
     -----------------------------------------------------------
       1 20    60   100  140   180   220   260   300   340   380

                         NORMALIZED RANGE
```

```
PERCENTAGE

  14 + **
     | **    **    **
  10 + ** ** ** ** **
     | ** ** ** ** ** **
   6 + ** ** ** ** ** ** **
     | ** ** ** ** ** ** ** ** ** **                           **
   2 + ** ** ** ** ** ** ** ** ** ** ** ** ** ** **            **
     -----------------------------------------------------------
         5    15    25    35    45    55    65    75    85    95
```

13

ROBUST DISPERSION

Table 1: Measures of Price Dispersion

	Coefficient of Variation	Normalized Range	Robust Dispersion
N	2430	2430	2430
Mean	69.4	398	55.4
Quantiles			
100% Max	1712	42430	39985
99%	512	3889	237
95%	226	1300	91
90%	135	736	65
75% Q3	65	310	38
50% Med	42	183	21
25% Q1	29	124	11
10%	20	85	1.5
5%	16	68	0.09
1%	8	35	0.002
0% Min	0	0	0

Table 2: Individual Product Price Dispersion

Product #	RD	Product Description
2392045	39985	Other Household furnishings - Napkins
2099967	8796	Perishable Food Products - Tortillas, Tamales, and other Mexican Food Specialties
3079030	2813	Misc. Plastic Products - Plastic Bottles
3131061	1834	Footwear Cut Stock - Other boot & shoe cut stock and findings
3079065	1657	Molded Plastic Products NEC - castings
2421896	1351	Softwood Flooring and Siding - other planing mill and sawmill products
2899597	1193	Essential Oil, Fireworks and Chemical NEC - other industrial chemical specialties including fluxes, plastic wood preparations and embalming chemicals
3079061	815	Molded Plastic Products NEC - injection molding
3691411	791	Storage Batteries, Lead Acid Type - Industrial Truck
3691419	706	Storage Batteries, Lead Acid Type - other motive power, including mining and industrial locomotive
2257820	667	All other Weft Knit Fabric - narrow fabrics under 12" wide
3873126	566	Clocks (not having balance wheel and hairspring) - all other including chime and strike
3551221	529	Commercial Food Products Machinery - Choppers, Grinders, Cutters, etc.
3494640	496	Hydraulic and Pneumatic Hose or Tube End Fittings and Assemblies except Aerospace
2851951	432	Miscellaneous Paint Products - Organosols and Plastisols, other than coatings
2599097	400	Furniture and Fixtures NEC - Other NEC except household
2299340	396	Scouring and Combing Mill Products - Tops and Noils
3634510	357	Electrical Housewares and Fans - Small household appliances, including razors
2599021	345	Furniture and Fixtures NEC - Hospital Beds
3079094	338	Miscellaneous Plastic Products

		- Building and Construction
3312192	299	Blast Furnace Products
		- Slag
3079066	298	Molded Plastic Products NEC
		- Other
3799988	286	Transportation Equipment NEC
		- Parts for Automobile and Light Truck Trailers
2843085	249	Surface Active and Finishing Agents
		- Bulk Surface Agents (detail reported ITC)
2299350	237	Scouring and Combing Mill Products
		- Scoured wool and other products

Table 3: Two Digit Industry Price Dispersion All Industries

Major Group		Number Products	Average CV	Average RNG	Average RD
20	Food and Kindred Products	590	46.532	247.511	25.731
21	Tobacco Manufactures	7	46.409	187.267	31.249
22	Textile Mill Products	101	70.191	332.998	37.619
23	Apparel and Other Textile Products	76	122.851	926.101	33.598
24	Lumber and Wood Products	143	70.675	666.787	14.330
25	Furniture and Fixtures	65	83.949	509.331	18.608
26	Paper and Allied Products	105	44.676	300.196	15.564
27	Printing and Publishing	94	105.513	876.306	30.596
28	Chemicals and Allied Products	179	55.886	275.694	24.099
29	Petroleum and Coal Products	47	89.435	521.475	20.024
30	Rubber and Misc. Plastic Products	34	62.122	283.529	16.549
31	Leather and Leather Products	25	80.961	452.479	30.230
32	Stone, Clay and Glass Products	86	74.368	490.113	18.481
33	Primary Metal Industries	107	66.430	330.023	27.633
34	Fabricated Metal Industries	269	64.602	378.982	17.645
35	Machinery, Except Electrical	154	58.636	279.295	19.821
36	Electric and Electronic Equipment	28	76.963	328.271	26.455
37	Transportation Equipment	76	81.192	440.681	33.087
38	Instruments and Related Products	29	73.577	265.709	24.625
39	Misc. Manufacturing Establishments	63	79.570	376.052	24.196

Bibliography

Abbott, T.A. III. "Producer Price Dispersion and the Analysis of Production." Unpublished doctoral dissertation. Harvard University. 1988.
____________. "Mean and Variance Estimators in Dirty Data: Small Sample Properties." Manuscript. March 1989.

Burdett, K. and K.L. Judd. "Equilibrium Price Dispersion." ECONOMETRICA. July 1983. p. 955-969.

Carlton, D.W. "Contracts, Price Rigidity, and Market Equilibrium." THE JOURNAL OF POLITICAL ECONOMY. 1979. p. 1034-1062.

___________. "The Rigidity of Prices." THE AMERICAN ECONOMIC REVIEW. September 1986. p. 637-658.

Dahlby, B. and D.S. West. "Price Dispersion in an Automobile Insurance Market." THE JOURNAL OF POLITICAL ECONOMY. April 1986. p. 418-438.

Hempel, F.R., E.M. Ronchetti, P.J. Rousseeuw, and W.A. Stahel. ROBUST STATISTICS: The Approach Based on Influence Functions. John Wiley & Sons: New York, NY. 1986.

Isard, P. "How Far Can We Push the 'Law of One Price'?" THE AMERICAN ECONOMIC REVIEW. December 1977. p. 942-948.

Perloff, J.M. and S.C. Salop. "Firm-Specific Information, Product Differentiation, and Industry Equilibrium." OXFORD ECONOMIC PAPERS. November 1986. p. 184-202.

Pratt, J.W., D.A. Wise, and R. Zuckhauser. "Price Differences in Almost Competitive Markets." THE QUARTERLY JOURNAL OF ECONOMICS. May 1979.

Reiganum, J.F. "A Simple Model of Equilibrium Price Dispersion." THE JOURNAL OF POLITICAL ECONOMY. 1979. p. 851-858.

Salop, S.C. and J.E. Stiglitz. "Bargains and Ripoffs: A Model of Monopolistically Competitive Price Dispersion." THE REVIEW OF ECONOMIC STUDIES. 1977. p. 493-510.

Stigler, G.J. "The Economics of Information." THE JOURNAL OF POLITICAL ECONOMY. June 1961. p. 213-225.

______________ and J.K. Kindahl. THE BEHAVIOR OF INDUSTRIAL PRICES. Columbia University Press. New York, NY. 1970.

Stiglitz, J.E. "Towards a more General Theory of Monopolistic Competition." ECONOMIC RESEARCH PROGRAM MEMORANDUM NO. 316. Princeton University. October 1984.

U.S. Office of Management and Budget. STANDARD INDUSTRIAL CLASSIFICATION MANUAL 1972. U.S. Printing Office: Washington D.C. Stock Number 4101-066.