Technical Report No. 4
Issued Under Office of Naval Research
Contract No. Nonr - 1858(02)

ESTIMATES OF NATIONAL AND REGIONAL STEEL PRODUCTION IN THE UNITED STATES FOR 1970

Robert E. Kuenne

Econometric Research Program
Research Memorandum No. 6
2 February 1959

This study was prepared under contract with the Office
of Naval Research and reproduction in whole
or in part will be permitted for any
purpose of the United States
Government

Princeton University
Econometric Research Program
92-A Nassau Street
Princeton, N. J.

Estimates of National and Regional Steel Production in the United States for 1970

Robert E. Kuenne

1. The Approach. A fundamental requisite of the manganese model which the Econometric Research Project published recently was an estimated bill of manganese requirements for the given years whose variables constituted the unknowns of the model. Moreover, since the model was originally envisioned as including the spatial relationships between points of ore concentration or import on the one hand, and of steel production on the other, these estimates were to be broken down by areas of steel production.

Since manganese is added in various forms to steel at stages from the blast furnace up through the ladle, steel for ingots and castings may be treated as containing all of the manganese in the final product. Also, should a differentiation of steels by manganese content prove useful in some future study, this stage of manufacture provides a workable basis of classification. Lastly, the manganese content of steel ingots can be computed with reasonable accuracy from known chemical compositions within the limits of tolerance established by such groups as the Society of Automotive Engineers and the American Iron and Steel Institute.

These facts and others bearing upon the availability and reliability of data argued for the treatment of manganese as a derived demand rather than for an approach which would treat directly with usage data for the metal itself. Moreover, since about 97 per cent is used in the metal industries, by treating it exclusively as a material whose demand is derived from the manufacture of steel for ingots and castings, we may concentrate the analysis upon a series of rather homogeneous problems without a great deal of omission of distortion.[1]

The method of analysis, therefore, will be to project the demand for steel ingots and non-foundry steel for castings in a set of regions for a given future period, with the ultimate goal of multiplying these projections by coefficients of manganese requirements per ton of steel. The present study addresses itself entirely to the estimates of steel production.

2. Defining the Time Period. The terminal year for which the final projections were sought was dictated by considerations of the broader study. It is the nature of this problem of dating that it is one of reasonableness, not one of exactitude in some undefined sense. Since the manganese study was one which envisaged the possibility of radical changes from present sources of the metal, time had to be allowed for further experimentation and development of processes on the technological horizon, and a sufficient period had to be postulated by the study for their acceptance and plant construction.

However, the general conditioning of the study by the international political situation tended to foreshorten the horizon which the purely technological might have dictated. For example, a twenty-year period to allow further research into the technological and economic nature of the processes, the building of pilot plants, the overcoming of inertia and starting frictions, the more exact determination of amounts and qualities of domestic ores, and the full-blown development of production facilities to the extent that the study determined to be optimal,

---

[1] Thus, the broader problem concerned with manganese estimates has been narrowed in two directions: (1) the neglect of chemical, battery, and other industry demand for manganese which totalled 3 per cent of consumption in 1952, as quoted in G. L. DeHuff, "Manganese," reprinted from the U.S. Bureau of Mines, Minerals Yearbook, 1953, p. 5; and (2) the exclusion of steel used for castings and produced in foundries for that purpose. This latter exclusion is a minor one, since castings made from steel in the pre-ignot stage in non-foundries are included in the American Iron and Steel Institute steel tonnage data. The second point is more important for the narrower analysis of steel output which is the major concern of this paper.

would not be overly conservative. On the other hand, to have projected the basic, underlying political matrix within which international power politics and decision-making operate, much beyond ten years would have been to compound an inordinate courage with outright fool-hardiness.

On the basis of considerations such as these, a period of roughly fifteen years was chosen over which the needed developments and adjustments could take place. The period ending in 1970 was selected. This medium-term period imposed the necessity of making formal assumptions upon which to found our analysis. Those having the most direct bearing upon the steel estimates are these:

a. full employment of resources. It will be assumed that the period between the present and 1970 is essentially one of full employment. The primary need for such an assumption springs from the dynamics of the steel industry's demand and supply implied by our methods, as well as from the fact that the data used for projection come from such a period. In past periods, existing steel capacity has become a rather rigorous limitational factor in the production of steel. The industry has revealed a good deal of caution in the expansion of its plant during periods of underemployment, and, moreover, has been slow to build far in advance of current demand. Consequently, our methods assume that the statistically-estimated production for the latter years of the period will not be unobtainable because of troublesome slumps in demand which interfered with the building of new capacity earlier.

b. continuance of the cold war. The international political climate now existent between the two major power blocs will be assumed unchanged during the period. It is difficult to specify the exact nature of this assumption, but for purposes of this study it will be acceptable to assume that conditions will continue to enforce upon the United States a military budget between 9 and 12 per cent of the gross national product, with no intervening "hot" wars. Thus, in the year 1970, the nation will be assumed to be facing a threat whose relative economic proportions are somewhat comparable to those of the present time.

c. continued relevance of present "economic potential for war". Of necessity, we must adopt the assumption that the possibility of limited and total war continues to impose upon the nation's economy the burden of preparing for it steel-wise in terms roughly similar to those the technology of war imposes today. For example, during the cold war which we assume will continue through the period of our projections, the supposition must be that military preparedness will continue to imply the existence of "conventional weapons" and the means to produce them, or that new weapons continue to demand existing and potential steel in amounts comparable to those now experienced and anticipated in preparing for limited and total war. Under such conditions, steel, and its demand derivative, manganese, will continue their roles as important war materials in a manner roughly proportionate to their present roles.

Our concern with that segment of steel demand which springs from military requirements reflects the present state of flux in military technology and strategy, and the potential changes in the parameters of our study which different courses in either area could cause. Although our study must also assume that civilian demands for steel will not be altered radically, at present no reason exists for supposing that large changes will occur in this sector of steel demand.

3. The Estimates. The method used to obtain the projections of steel output was to fit exponential trends to annual time series data adjusted for certain isolable short-run movements. Three sets of estimates for the six regions adopted as spatial units and two sets of estimates for the nation

investigated first were the ones requiring a minimum additional sacrifice of degrees of freedom.[4]

Ideally, in the absence of knowledge concerning the direction and extent of the expected observation errors in the data, the classification adopted for the study of steel production should be devised to minimize the error of estimation of manganese requirements. A stratification of steels by chemical composition has been developed by the Society of Automotive Engineers and the American Iron and Steel Institute. Since manganese is one of the chemicals listed, it was believed desirable in earlier stages of the study to obtain an analytical breakdown of steel by combinations of these listings reasonably homogeneous in manganese content.

For each of these types of steel the specified ranges for manganese content are available, as well as the allowable variations from either end of the ranges.[5] Since the goal of the metallurgist is to obtain a metal content which is the mean of the range,[6] it is conceptually possible to derive some distribution of actual manganese content for each category of steel used, and thereby obtain estimates of total manganese content with specified upper and lower limits.

---

[4] Some idea of the need to economize on degrees of freedom in time series analysis can be gained from Orcutt's and James's findings that the distribution of sample correlation coefficients between time series generated by a model very close to that assumed by our estimates of Set 3 follows the distribution of those from random series with 5 degrees of freedom when the time series are of 60 observations and 6 degrees of freedom when they contain 90 observations. The effective degrees of freedom in time series are much lower than the number of observations, so that nominal degrees of freedom must be jealously conserved. See G.H. Orcutt and S.F. James, "Testing the Significance of Correlation Between Time Series," Biometrika, XXXV (1948), 412.

[5] Society of Automotive Engineers, 1956 SAE Handbook, 51-6, 105-07.

[6] Statement in a letter from Mr. Charles M. Parker, American Iron and Steel Institute, December 12, 1956.

Three considerations led us to drop this more ambitious stratification. First, carbon steel so dominates the total production data that the time spent estimating the other categories would not bear returns sufficiently great to warrant its expenditure. In 1955, for example, carbon steel, plain and free-cutting, amounted to about 91 per cent of the total steel produced, and no other single category included as much as 2 per cent of the total.[7] Slight relative observation errors in the carbon steel category would swamp the added amount of precision gained from such disaggregation.

Second, the categories for which alloy steel data have been collected differ over the time period involved, so that a substantial aggregation would be necessary to obtain homogeneous categories between years, enforcing a sacrifice of the manganese-content criteria to an unacceptable degree. On the basis of the information available, some of the data could not be distributed between categories in any practicable scheme.

Third, it is impossible to obtain even such highly aggregated data for the regions adopted in the study. Since an important purpose of the study was to make regional projections, this was considered to be a serious fault with the method.

For these reasons no distinction is made between types of steel whose production is included in the data: all distinctions between types of steel and differential manganese requirements, as well as regional differences in the types of steel produced, must be introduced in the second stage of the study, when the manganese coefficients per ton of steel are computed.

---

[7] Cf. the Annual Statistical Reports of the American Iron and Steel Institute for classifications of alloy steels between 1943 and 1955.

b. The Regions. The regional classification adopted in this study is that used by the American Iron and Steel Institute.[8] It includes six districts whose boundaries are described below:

1. Eastern -- This will be symbolized by "e" in this study. It includes New England, Delaware, most of Maryland, and the Middle Atlantic States with the exception of the Pittsburgh-Monongahela area. The important steel production points within it are Sparrows Point, Philadelphia-Morrisville, Allentown-Bethlehem, and Buffalo.

2. Pittsburgh-Youngstown -- symbolized by "py" in the study. It includes the Pittsburgh-Monongahela area in Pennsylvania, West Virginia, Kentucky, and all of Ohio with the exception of a 50 to 60- mile band across the northern part of the state.

3. Cleveland-Detroit -- symbolized by "cd" in the study. It comprises all of Michigan and the band of 16 Ohio counties across the north of the state excluded from the Pittsburgh-Youngstown district.

4. Chicago -- symbolized by "c". It includes Indiana, Illinois, Iowa, Minnesota, Missouri, and Wisconsin.

5. Southern -- denoted by the symbol "s". It includes the states south of but not including Maryland, West Virginia, Kentucky, Missouri, and Kansas, and east of and including Texas and Oklahoma.

6. Western -- symbolized by "w" in the analysis. It includes all states west of, but not including, Minnesota, Iowa, Missouri, Oklahoma, and Texas.

c. The Models. The models used to obtain steel estimates are quite simple ones. The production of steel in the nation in period t, Y_t, and in region r in period t, Y_t^r, are considered to be functions of n systematic variables plus a random error term in the case of the nation:

$$[1] \quad Y_t = F(X_1, X_2, \ldots X_n, \mu_t)$$

where the X_i include different variables at the same period of time and the same variable at different periods of time, and μ_t is a randomly drawn error term from normally-distributed universes. These "national" factors are conceived of

---

[8] For a map of the districts and data on capacity in the past, see American Iron and Steel Institute, Steel Facts, February, 1954, 2.

as operating uniformly over regional production to impart uniform influences to all regions. As far as the regional model is concerned, these national factors are supplemented by regional forces which are peculiar to the regions. They are denoted X_{n+1}^r through X_s^r in the models:

$$[2] \quad Y_t^r = F^r(X_1, X_2, \ldots X_n, \mu_t; X_{n+1}^r, \ldots X_s^r, \mu_t^r)$$

For both the national and regional models a first subset of these variables, $(X_1, X_2, \ldots X_i)$, is looked upon as contributing a *trend* component to Y_t and an identical component uniform for all regions to Y_t^r. A second subset, $(X_{i+1}, \ldots X_n)$, of systematic variables acts in both national and regional models to impart a series of *short-run* impacts which are identical components of Y_t and Y_t^r. A third subset, $(X_{n+1}^r, \ldots X_p^r)$, affects each region differentially in contributing a regionally specific trend component to Y_t^r, and a fourth and final subset of systematic variables, $(X_{p+1}^r, \ldots X_s^r)$, imparts peculiar short-run movements to each region.

Therefore, for the national model we may distinguish three distinct movements in national production:

$$[3] \quad \begin{aligned} & G_t = F_g(X_1, X_2, \ldots X_i) \\ & D_t = F_d(X_{i+1}, X_{i+2}, \ldots X_n) \\ & \mu_t \end{aligned}$$

where G_t denotes the growth component, D_t the short-run component, and μ_t the stochastic element. Similarly, the regional production data are considered the resultants of six variables:

$$G_t$$

$$D_t$$

$$\mu_t$$

[4] $$G_t^r = F_g^r(X_{n+1}^r,\ X_{n+2}^r,\ \ldots\ X_p^r)$$

$$D_t^r = F_d^r(X_{p+1}^r,\ X_{p+2}^r,\ \ldots\ X_s^r)$$

$$\mu_t^r$$

where the G_t^r factors denote regional trend elements, the D_t^r factors are regional short-run forces, and the μ_t^r stochastic elements indigenous to the regions. The term, μ_t, is randomly drawn and is independent of μ_t^r; similarly, the μ_t^r are random and independent of each other.

The basis upon which the G factors are distinguished from the D components is a familiar one, but must be made explicit in order that the procedure followed is clear. In Schumpeter's words:

> "If trend-analysis is to have any meaning it can derive it only from previous theoretical considerations which must not only guide us in interpreting results, but also in choosing the method. Failing this, a trend is no more than a descriptive device summing up past history with which nothing can be done. It lacks economic connotation. ."[9]

In order to endow our method with an economic meaning, and thereby gain methodological meaning, we have adopted an outlook defined by Frickey:

> "One view is that we can best give rationality to the statistical process of separating the various elements in a time series--referred to as the statistical decomposition of the series--by thinking of this process

---

[9] J. A. Schumpeter, "Mitchell's Business Cycles," *Quarterly Journal of Economics*, XLV (1930), 167.

> as one of analysing the effects of particular groups of causes. This general conception would, of course, lead to different schemes for time-series decomposition according to the beliefs held as to the nature of the causal influences."[10]

As the economy moved through time, the groups of trend forces brought about a monotonic increase in steel production, as a result of the irreversible impacts of these forces. The key characteristic of these forces is not their continuity, for they may act discretely; nor is it their graduality, for they may be quite strong as measured by their influence upon production. It is, rather, the *a priori monotonicity* of the path they enforce upon steel production, whether national or regional. If a systematic factor, X, does not possess this quality of affecting Y_t or Y_t^r monotonically, we shall group it with the D_t or D_t^r subsets. Specifically, in an intertemporal analysis, if a variable changes in period t and induces a rise (fall) in Y_t but a fall (rise) in Y_{t+1}--in the manner, for example, of some of Schumpeter's innovations--it is not considered one of the G_t or G_t^r subsets.

Correlation is the existence of an ordered sequence in one variable when it is ordered with reference to a second variable. The construction above is designed to produce a positive or negative correlation of the values of the variables in the growth subsets with time. They may be looked upon as related to the causally-bogus time variable, t, as follows:

$$\begin{aligned} &\underline{X_1 = F_1(t, \upsilon_1)} \\ [5] \qquad &\underline{X_2 = F_2(t, \upsilon_2)} \\ &\quad \vdots \\ &\underline{X_i = F_i(t, \upsilon_i)} \end{aligned}$$

and

---

[10] E. Frickey, *Economic Fluctuations in the United States*, Cambridge, Mass., 1942, 37.

[6]
$$\begin{gathered}\underline{X^r_{n+1} = F^r_{n+1}(t,\ \upsilon^r_{n+1})}\\ \underline{X^r_{n+2} = F^r_{n+2}(t,\ \upsilon^r_{n+2})}\\ \vdots\\ \underline{X^r_p = F^r_p(t,\ \upsilon^r_p)}\end{gathered}$$

where the $\underline{\upsilon}$ factors are stochastic terms drawn at random from independent, normally-distributed universes.

Two considerations dictate that this common dependence upon time be exploited and a "net resultant" function be sought rather than estimates of individual functions of the types in [5] and [6]. First, the basic inability of our analytical tools to specify all of the systematic variables having the characteristics catalogued above, whose impacts upon $\underline{Y}_t$ and $\underline{Y}^r_t$ are non-negligible, must be admitted. Second, the existence of a set of variables related to the same independent variable in the same functional way (as would probably be true) leads to troublesome identification problems. Therefore, we postulate the relationships,

[7]
$$\underline{G}_t = \underline{F_g(X_1,\ X_2,\ .\ .\ .\ X_i) = F_h(t,\ \upsilon_1,\ \upsilon_2,\ .\ .\ \upsilon_i) = F_j(t, \upsilon\ )}$$

[8]
$$\underline{G}^r_t = \underline{F^r_g(X^r_{n+1},\ .\ .\ .\ X^r_p) = F^r_h(t,\ \upsilon^r_{n+1},\ .\ .\ .\ \upsilon^r_p) = F^r_j(t,\ \upsilon^r)}$$

It follows from the discussion of trend factors that to the extent the "short-run" subsets contributing components $\underline{D}_t$ and $\underline{D}^r_t$ do not possess these characteristics, they are uncorrelated with a monotonic variable. A harmonic movement, say of the sine curve type, will, over long periods, decline for one half of the observations and rise for the other half, and therefore be uncorrelated with a monotonic curve. Similarly, it is of the nature of irregular short-term movements to be uncorrelated with time. Of course, the extremely short time period of our analysis may lead to the spurious correlations between time series whose explanation by Yule has become classic.[11] Once more, however, we must appeal to the *a priori* nature of the movements to furnish the basis for separating them, not to their sample behavior.

The methods of economic analysis are such as to force us to place our sole reliance for prediction upon the ability to isolate the dependence of $\underline{G}_t$ and $\underline{G}_t^r$ upon $\underline{t}$ and to postulate its stability as a transformation unique to a linear degree of the net resultant of the underlying relationships defined in [3] and [4]. Given the rigorous meaning we confer upon a trend line and our assertion of the hypothesis that it measures the impact of the set of variables we seek to isolate unique to a linear transformation, this study will treat time, $\underline{t}$, not as a bogus variable in the ordinary sense, but as one which has a meaningful, if indirect, role in the causal process which generated our data.

At the same time, our methods will debar us from accurate prediction of $\underline{D}_t$ or $\underline{D}_t^r$ and therefore $\underline{Y}_t$ and $\underline{Y}_t^r$. Of course, ideally, the task of the analysis consists in isolating these $\underline{D}_t$ and $\underline{D}_t^r$ from the sample data and proceeding to determine a trend line in such a way as to reduce the residual variation to randomness; moveover, because our data are time series and, consequently, of the high probability of autocorrelation in these error terms, this test for randomness should take into explicit account the temporal sequence of the errors.

However, the very logic of our method is the result of recognizing the general inability of our analytical tools to isolate the short-run factors

---

[11] G. Udny Yule, "Why Do We Sometimes Get Nonsense-Correlations Between Time Series?", Journal of the Royal Statistical Society, LXXXIX (1926), 1-64.

with precision. Therefore, we must recognize that our estimates will be of the form:

$$[9] \qquad \underline{\bar{Y}_t - \bar{D}_t{}^* = F(\bar{G}_t, \lambda_t, \bar{\mu}_t)}$$

and

$$[10] \qquad \underline{\bar{Y}^r_t - \bar{D}_t{}^* - \bar{D}^r_t{}^* = F^r(\bar{G}_t, \bar{G}^r_t, \lambda_t, \lambda^r_t, \bar{\mu}_t, \bar{\mu}^r_t)}$$

where the λ factors are residual systematic variation resulting from faulty elimination of the D components, the bars denote estimates of population parameters, and the asterisks denote the faulty estimates of these short-run components. Since these short-run movements will be highly autocorrelated, the probability that the λ factors will also be autocorrelated is quite high, and, therefore, such residuals when combined with the stochastic element, $\bar{\mu}$, will bias the combination toward non-random behavior sequentially through time. A more realistic outlook upon the goal of our analysis of the short-term components then, is to reduce these λ factors to small enough mangitudes that their combination with the truly stochastic elements, $\bar{\mu}$, does not depart greatly from randomness.[12]

---

[12] That this is a more realistic picture of the task of the economist dealing with the statistical end of his profession is brought out well by C. Christ, "Aggregate Econometric Models", *American Economic Review*, XLVI (1956), 386. This point has a peculiar relevance for the present analysis, since in the correlation of variables involving time the non-existence of auto-correlated error terms is extremely important for reliable results. However, as a result of the work of Orcutt and James, it is possible to conclude: "...Except for the fact, which our experiments demonstrate, that nearly optimum results can be achieved if the error term is only a rough approximation to a random series, solution of the problem would seem rather hopeless for series of only twenty items." See "Application of Least Squares Regression to Relationships Containing Autocorrelated Error Terms", *Journal of the American Statistical Association*, 44 (1949), 53.

d. The Isolation of the Short-Run Components. In the period under consideration, three sets of systematic variables dominated the $\underline{D}_t$ and $\underline{D}_t^r$ components:

1. labor strife internal and external to the steel industry;
2. wars of total or limited variety in 1943-1945 and 1950-1953;
3. cyclical downturns in 1948-1950 and 1953-1954.

This catalogue excludes the development of cold war with the adoption of the Truman Doctrine and the enunciation of the Marshall Plan in 1947; the development of an extensive military aid program after the outbreak of the Korean War; and other measures of this nature, all of which mark structural changes in the environment in which foreign policy must be designed. These cold war forces have been subsumed as elements in the subset of variables contributing to $\underline{G}_t$, as having effected, from the time of their inception, a monotonic upward pressure upon steel production, and, under the basic set of assumptions in section 2b above, have been assumed to continue to do so in the period for which projections are sought.

1. Effects of Work Stoppages. Major interruptions in the production of steel on the account of labor stoppages have occurred during general and non-general but major strikes within the industry and bituminous coal strikes outside the industry. Major occurrences of this type, together with data on actual production during the periods and estimates of potential production, are listed in Table 1:

Table 1

Major Work Stoppages, Production, and Estimated Production Losses, 1943 - 1955[13]

Year	Cause	Actual Production (net tons)	Potential Production (net tons)	Tonnage Lost (net tons)
1943	---	-------	------	----
1944	---	-------	------	----
1945	Bituminous Coal Strike, Sept. 21 - Oct. 17	11,579,300	13,340,300	1,761,000
1946	General Steel Strike, Jan. 21 - Feb. 17	11,774,300	19,563,300	7,789,000
	Bituminous Coal Strike, April 1 - June 3	15,559,600	19,788,600	4,229,000
	Bituminous Coal Strike, Nov. 20 - Dec. 9	12,218,300	13,673,300	1,455,000
	Total, 1946	39,552,200	53,025,200	13,473,000
1947	Bituminous Coal Strike, June, followed by Interplant Railroad Workers Walkout, July - Sept.	27,309,400	28,826,400	1,517,000
1948	Bituminous Coal Strike, March and April	12,444,700	14,059,200	1,614,500
	Bituminous Coal Controversy, June 28 - July 26	8,336,400	8,670,000	333,600
	Total, 1948	20,781,100	22,729,200	1,948,100
1949	General Steel Strike, Oct. 1 - Nov. 15	3,709,800	12,879,300	9,169,500

---

[13] Data from the American Iron and Steel Institute compiled from data furnished by producers.

Table 1 (continued)

Year	Cause	Actual Production (net tons)	Potential Production (net tons)	Tonnage Lost (net tons)
1950	Bituminous Coal Strike, Jan. 9 - Mar. 6	11,458,500	12,530,000	1,071,500
	Other Stoppages		455,732	455,732
	Total, 1950		12,985,732	1,527,232
1951	Work Stoppages		867,583	867,583
1952	Work Stoppages		110,509	110,509
	General Steel Strike, April - August		17,900,166	17,900,166
	Total, 1952			18,010,675
1953	Work Stoppages		703,100	703,100
1954	Work Stoppages		8,200	8,200
1955	Work Stoppages		277,700	277,700

The employment by the American Iron and Steel Institute of "normal" production data before the strike dates to obtain estimates of the potential production lost during the stoppages probably introduces an upward bias into the calculations of Table 1. The general steel and bituminous coal strikes are foreseeable by consumers of steel and lead to abnormal production and inventory build-ups before the strikes in times of strong demand. In instances of general slackening of demand such stoppages may serve merely as a method of inventory liquidation. Consequently, the projection of production rates just anterior to the strikes gives unrealistically high representations of the impacts of such strikes. This is doubly true in that the method does not take into account the possibility of compensating for "lost" production after settlement.

It is easier to identify these biases than to eliminate them. The results of using the data of Table 1 to adjust production are credible insofar as the years of great labor strife -- 1946, 1949, and 1952 -- yield production figures after adjustment which are below the previous year's adjusted data in the first two cases and above it only for 1952, a year which saw the culminating effects of the Korean War build-up. These unadjusted data for national steel production and such data adjusted by adding the estimates of tonnage lost through labor strife are listed in Table 2:

Table 2

Steel Production, Unadjusted and Adjusted for Work Stoppages, United States, 1943-1955

Year	Steel Production (net tons)	Steel Production Adjusted (net tons)
1943	88,743,637	88,743,637
1944	89,461,731	89,461,731
1945	79,719,451	81,480,451
1946	66,590,604	80,063,604
1947	84,783,981	86,300,981
1948	88,533,729	90,481,829
1949	77,868,353	87,037,853
1950	96,696,769	98,224,001
1951	105,134,553	106,002,136
1952	93,156,375	111,167,050
1953	111,609,719	112,312,819
1954	88,311,652	88,319,852
1955	117,036,085	117,313,785

Unfortunately, the breakdown of the work stoppage data by regions is not available. Economic factors and the mores of the industry introduce a great deal of regional regularity in steel production, and we shall take advantage of this regularity in our regional adjustments. In Table 3 are plotted the troughs in steel production in the nation as a whole and in the six regions as percentages of the peak most recently attained before the downturn:

Table 3

Troughs in Steel Production as Percentages of Previous Peaks

Regions	1946	1949	1952	1954
Eastern	71.3 o/o	89.4 o/o	87.2 o/o	78.7 o/o
Pittsburgh-Youngstown	71.8	84.9	87.9	74.3
Cleveland-Detroit	81.3	93.1	93.1	78.9
Chicago	76.4	89.2	87.2	85.1
Southern	74.4	90.4	91.6	88.7
Western	83.4	90.4	93.6	88.7
United States	74.4	88.0	88.6	79.1

This table can be regarded as a slice through the two economic dimensions of time and space. If we use variance analysis techniques to partition the data into variation ascribable to time, to space, and to error, we find about 67 per cent of the variation associated with the time factor and only about 30 per cent ascribable to space. Of course, no attempt to argue on probabilistic grounds will be made on the strength of these data, but the result does give us some idea of the relative strength of the $\underline{D}_t$ as compared with the $\underline{D}_t^r$ components.[14]

This marked tendency for the regional production data to move together is revealed by a further study of Table 3. Of the 24 regional percentages, 15 are

---

[14] The model used adopted the assumption that each observation, $\underline{X}_{ij\alpha}$, subtracted from the grand mean, $\underline{\overline{X}}$, equals a deviation from its column mean, $\underline{X}_{i..}$, plus a deviation from its row mean, $\underline{\overline{X}_{.j.}}$, plus an unexplained residual.

within ± 4 percentage points of the respective national figure, although of the 9 wider deviations the Western and Cleveland-Detroit regions account for 3 each. However, when the plenitude of differential regional factors is considered, this is a striking degree of similarity. Three factors seem to dominate in the explanation of this phenomenon:

1. the nature of the demand for the product in terms of delivered price is such that it can bear long haulages between regions, and market interpenetration is a result;

2. the oligopolistic character of the industry, together with its history of price leadership and frequent excess capacity, has resulted in a discriminatory pattern of pricing which makes effective the latent possibilities of (1) above;

3. the existence of a strong, industrial union, whose practice is industry-wide bargaining on a pattern basis, imposing a great deal of similarity of timing and relative amplitude upon shorter-run movements in all regions.

Therefore, in the absence of information about the distribution of production lost for the nation among the 6 regions, we distributed a given year's loss between regions on the basis of each region's percentage of total steel produced <u>for the previous year</u>. This results in the data of Table 4:

Table 4

Estimates of Regional Production Lost Through Work Stoppages, Net Tons

(000 omitted)

Year	Eastern	P - Y	C - D	Chicago	South	West
1943	----	---	---	----	---	---
1944	----	---	---	----	---	---
1945	310	727	144	389	81	70
1946	2,654	5,389	1,132	3,072	620	593
1947	293	607	137	343	70	68
1948	380	791	164	429	90	99
1949	1,788	3,677	825	1,990	449	486
1950	302	592	137	337	76	84
1951	168	337	83	187	43	49
1952	3,548	7,060	1,675	3,818	865	1,009
1953	136	273	69	147	34	44
1954	2	3	1	2	--	--
1955	56	98	28	65	16	17

The regional data, adjusted for these estimates, are listed in Table S-1, Appendix, Column 2.

2. Total and Limited War Periods. The period chosen for analysis contains two periods during which production of steel was strongly affected by war: 1943-45 and 1950-52. Moreover, the transition periods immediately following the termination of active hostilities felt the direct impacts of the war, while later periods' production data continued to be affected by these events.

It is next to impossible to eliminate this vastly complicated component of $\underline{D}_t$ and $\underline{D}_t^r$. The series are not long enough to use a moving average smoothing technique for even the crudest of methods to eliminate the bias. Some reasons tend to mollify the extent of disturbance to our estimates introduced by this set of considerations. As pointed out above, evidence exists that World War II marked a structural change in the demand by the American economy for durable goods: after the war, the proportion of gross national product expended on such goods never returned to its prewar levels. The Korean War, launching the American state into a period of highly probable limited wars for the future, enhanced the trend, and projected it forward with military aid to our allies. Some basis exists, therefore, for arguing that at least some of this short-term movement will enter into the G_t component.

Nevertheless, the net impact of these movements should be to bias our trends upward, particularly because the wars occur at either end of the series used. *The influences of these events are the single largest group of unisolable factors causing the greatest degree of misgivings about the method employed, and, therefore, the results obtained.*

3. *Cyclical Movements.* We are fortunate in the selection of a time period to the extent that cyclical forces have been subordinate to irregular movements. The two recessions mirrored in the production data adjusted for work stoppages are those which occurred in 1948-49 and 1954. The data dip only slightly in 1949 before recovering and responding to the Korean boom, while the much greater dip in 1954 reflects the non-cyclical reductions in Korean war spending.

Two decisions were made in the face of these cyclical downturns. First, the terminal year of the series, 1955, was one of rampant prosperity, while that of 1954 was depressed. The selection of the former meant that the trend line began and ended in prosperous phases, which, while it might bias the level of the trend lines upward, would eliminate the slope-bias which selection of the year 1954 would have introduced. Moreover, to have dropped 1955's observation would have been to squander a precious degree of freedom. Given the stress the study places upon the conservation of these nominal degrees of freedom, it will not be surprising that the series analyzed ends in 1955.

Second, as will be explained below, in Sets 2 and 3 of the estimates the data for 1954 were adjusted in the interest of conservatism by substituting for them the average of 1953 and 1955, adjusted for work stoppages, in lieu of this latter figure.

No other adjustment was made to eliminate the cyclical forces, either nationally or regionally.

To sum up the isolation of short-term factors: our estimates of the data adjusted for short-term fluctuation, $[\overline{Y}_t - \overline{D}_t^*]$ and $[\overline{Y}_t^r - \overline{D}_t^* - \overline{D}_t^{r*}]$, are, with the exception of the 1954 observation in Sets 2 and 3, the original data adjusted for work stoppages.

e. *Fitting the Models Statistically.* After plotting the adjusted data on arithmetic and semi-logarithmic grids, we verified the hypotheses that the exponential function would give reasonably good visual fits in all cases. Specifically, the forms of the estimating equations adopted are:

$$\log[\overline{Y}_t - \overline{D}_t^*] = \log a + \log b\,[t] + \log \lambda_t + \log \bar{\mu}_t \tag{11}$$

$$\log[\overline{Y}_t^r - \overline{D}_t^* - \overline{D}_t^{r*}] = \log a^r + \log b\,[t] + \log b^r\,[t] + \log \lambda_t + \log \bar{\mu}_t + \log \lambda_t^r + \log \bar{\mu}_t^r \tag{12}$$

Three sets of analyses to fit these relationships were made, and are discussed below. For convenience, we shall denote the sum of the last four terms in [12] E_t^r.

1. *Estimates, Set 1.* A straight-forward least-squares method of fitting these linear functions to the logarithms of the data of Column 2, Table S-1, Appendix, was applied, with one simple modification. The term, *log b [t]*, was subtracted from each of the regional observations before fitting, in the hope that this simple transformation would aid in reducing E^r_t to non-significant auto-correlation. Therefore, in fitting the regional models, the observations were transformed to:

$$[13] \quad [\log \bar{Y}^r_t - \bar{D}^*_t - \bar{D}^{r*}_t] - \log b\,[t] = \log a^r + \log b^r\,[t] + E^r_t$$

For convenience, we define,

$$[14] \quad \text{antilog}[\log(Y_t - D_t)] = Z_t$$

$$[15] \quad \text{antilog}[\log(Y^r_t - D_t - D^r_t) - \log \beta[t]] = Z^r_t$$

and let $\bar{Z}_t$ and $\bar{Z}^r_t$ be the corresponding estimates made from the trend equations.

Using a subscript before the symbol to denote the set number of the estimates, we obtained these estimating equations:

Table 5

Estimating Equations, Set 1

Region	Equation
Eastern	$\log {}_1\bar{Z}^e_t = 7.20263 + .00120\,t$
Pittsburgh-Youngstown	$\log {}_1\bar{Z}^{py}_t = 7.53753 - .00593\,t$
Cleveland-Detroit	$\log {}_1\bar{Z}^{cd}_t = 6.82044 + .00820\,t$
Chicago	$\log {}_1\bar{Z}^c_t = 7.25746 - .00020\,t$
Southern	$\log {}_1\bar{Z}^s_t = 6.55549 + .00770\,t$

Table 5 (continued)

Region	Equation
Western	$\log {}_1\bar{Z}^w_t = 6.49244 + .02176\, t$
United States	$\log {}_1\bar{Z}_t = 7.91252 + .01042\, t$

Estimates of regional production based upon these equations,[15] and estimates of long-term national production, will be found in column 3 of Table S-2, Appendix, and residuals from the equations can be found in column 6 in absolute form and column 9 in percentage terms.

The following hypothesis was framed for testing:

Hypothesis 1: the prior elimination of short-term movements from the data to which these equations were fitted, as well as the trend analysis itself, has reduced the systematic variation present in the residuals to the extent that λ_t and λ^r_t can be ignored.

To test this hypothesis with the added rigor of introducing time sequence into the test, we shall employ von Neumann's ratio. Let the term,

$$[16] \qquad \delta^2 = \frac{\sum_{i=1}^{n-1} (x_{i+1} - x_i)^2}{n-1}$$

be the mean-square successive difference of the residuals, x_i, ordered through time. Let s^2 be the variance of these residuals (calculated with a demoninator of n). Then, the ratio

$$[17] \qquad \frac{\delta^2}{s^2}$$

---

[15] That is, the antilogarithms of $\log {}_1\bar{Z}^r_t + \log b[t]$.

from randomly drawn samples from normal universes will have an expected value of *2n/n-1* and will be distributed symmetrically about this value.[16]

For a sample size of 12 the probability that the ratio is less than 1.25 is .05407, and, by transformation from the distribution of Young's statistic, the .05 level of significance for a sample size of 13 is 1.252. We shall adopt the .05 level of significance for a one-tailed test of the hypothesis. The sample von Neumann's ratios are given below in Table 6:

---

[16] See J. von Neumann, R.H. Kent, H.R. Bellinson, and B.I. Hart, "The Mean Square Successive Difference", *Annals of Mathematical Statistics*, XII (1941), 153-62, and J. von Neumann, "Distribution of the Ratio of the Mean Square Successive Difference to the Variance", *Annals of Mathematical Statistics*, XII (1941), 367-95. Values of the statistic have been computed for varying sample size and probability levels in B.I. Hart, "Tabulation of the Probabilities for the Ratio of the Mean Square Successive Difference to the Variance", *Annals of Mathematical Statistics*, XIII (1942), 213.

A similar statistic developed by L.C. Young, "On Randomness in Ordered Sequences", *Annals of Mathematical Statistics*, XII (1941), 293-302, is more convenient for our purposes because the .05 level is given for a sample size of 13, while it is available only for a size of 12 in Hart's table of von Neumann's ratio.

Young's statistic is defined as follows:

$$C = 1 - \frac{\sum_{i=1}^{n-1} (x_i - x_{i+1})^2}{2 \Sigma (x_i - \bar{x})^2}$$

with an expected value

$$E[C] = - \frac{1}{n}$$

von Neumann's ratio is then related to Young's statistic in the following manner:

$$\frac{\delta^2}{s^2} = \frac{2n}{n-1} (1 - C)$$

Table 6

von Neumann's Ratios

Estimates, Set 1

Region	Ratio
Eastern	2.01
Pittsburgh-Youngstown	1.40
Cleveland-Detroit	2.12
Chicago	1.99
Southern	1.84
Western	2.22
United States	2.08

On the basis of our tests we could accept the hypothesis in every case. However, a closer study of the data reveals that these high values of the ratio were obtained solely because of the extremely high residuals in the year 1954. As a consequence, the hypothesis was rejected in every case, and a new analysis begun to obtain random residuals with a corrected value for 1954's adjusted data.

2. Estimates, Set 2. As a first step in this new analysis, for each of the seven sets of data, the 1954 adjusted value in column 2 of Table S-1, Appendix, was changed to the mean of the 1953 and 1955 values. These substitutions are given at the foot of the table, and subsequent results for the year 1954 will be on the basis of this material.

Since we must expect the residuals of straight-forward least-squares regressions to be autocorrelated, we sought to find a Cochrane-Orcutt transformation of our adjusted data for the national figures which would yield an estimating equation with random residuals. It was hoped that the new log b estimate, when subtracted from the regional data adjusted for short-term factors, would lead to estimating equations for the regions on a straight-forward basis that would yield random residuals.

For the national model, therefore, a simple Markoff process which has become common as a representation of autocorrelation in economic time series was used as a hypothesis:

[18] $$\log \lambda_t = \log \lambda_{t-1} + \log \upsilon_t$$

where υ_t is a true stochastic factor.

Assuming that $\upsilon_{t=1} = 0$, and substituting [18] into [11] and [14],

[19] $$\log \bar{Z}_t - \log \overline{Z_{t-1}} = \log b + \log \alpha$$

where

[20] $$\log \alpha = \log \bar{\mu}_t - \log \overline{\mu_{t-1}} + \log \upsilon_t$$

which, as the sum of random variables, is also random. Therefore, if the adjusted national data were generated by such a model, the residuals should be random.

Fitting a least squares equation to the first differences of $\log \bar{Z}_t$, we obtained a value of $\log b = .01063$, the mean of the first differences. Substituting, the estimating equation would become:

[21] $$\log \bar{Z}_t = \log \overline{Z_{t=0}} + .01063\ t$$

Because of the unrepresentative nature of the origin year in our series, however, a different locational constant was obtained by a constrained least-squares procedure. That is, the normal equations were solved for $\log a$ subject to the constraint that $\log b = .01063$. The resulting equation is:

[22] $$\log {}_2\bar{Z}_t = 7.91953 + .01063\ t$$

Once more, Hypothesis 1 was framed for the national data and von Neumann's ratio adopted as a statistic to test the randomness of *the scatter of the differences about their estimate*, $\log b$. The value was 2.02, and therefore we accept the hypothesis that $\log \alpha$ is random.

Estimates made from [22] will be found in column 4 of Table S-1, Appendix, under the United States sector.

Next, this new estimate of $\log b$ was substituted in [15] and linear regressions calculated for each region. The results are found in Table 7:

Table 7

Estimating Equations, Set 2

Region	Equation
Eastern	$\log {}_2\bar{Z}^{e}_{t} = 7.19146 + .00448\ t$
Pittsburgh-Youngstown	$\log {}_2\bar{Z}^{py}_{t} = 7.52603 - .00254\ t$
Cleveland-Detroit	$\log {}_2\bar{Z}^{cd}_{t} = 6.80717 + .01187\ t$
Chicago	$\log {}_2\bar{Z}^{c}_{t} = 7.24513 + .00222\ t$
Southern	$\log {}_2\bar{Z}^{s}_{t} = 6.54586 + .00999\ t$
Western	$\log {}_2\bar{Z}^{w}_{t} = 6.48450 + .02403\ t$
United States	$\log {}_2\bar{Z}_{t} = 7.91953 + .01063\ t$

Estimates made for each region from these equations may be found in column 4, Table S-1, Appendix, absolute residuals in column 7, and percentage errors in column 10.

Hypothesis 1 was tested for each region by using von Neumann's ratio at the .05 level of significance. The following values were obtained:

Table 8

von Neumann's Ratios

Estimates, Set 2

Region	Ratio
Eastern	.59
Pittsburgh-Youngstown	.88
Cleveland-Detroit	.74
Chicago	.56
Southern	1.07
Western	2.07
United States	2.02

The hypothesis was rejected for every region but the Western. We therefore move to a third and final set of estimates.

3. **Estimates, Set 3.** This analysis consists in retaining the estimate of log b obtained in Set 2, and therefore the log ${}_{2}\bar{Z}_t^{r}$ of that analysis, but of adopting the expedient of fitting regression equations to the first differences of these transforms. Once more the locational constants in the equations were obtained by constrained least-squares. These results are given below in Table 9:

Table 9

Estimating Equations, Set 3

Region	Equation
Eastern	$\log {}_{3}\bar{Z}_t^{e} = 7.21180 + .00109\ t$
Pittsburgh-Youngstown	$\log {}_{3}\bar{Z}_t^{py} = 7.55243 - .00694\ t$

Table 9 (continued)

Region	Equation
Cleveland-Detroit	$\log {}_3\overline{Z}_t^{cd} = 6.82727 + .00852\, t$
Chicago	$\log {}_3\overline{Z}_t^{c} = 7.25878 - .00005\, t$
Southern	$\log {}_3\overline{Z}_t^{s} = 6.55911 + .00778\, t$
Western	$\log {}_3\overline{Z}_t^{w} = 6.53256 + .01605\, t$

Estimates based upon these equations, adjusted for the national trend, are to be found in column 5 of Table S-1, Appendix, absolute residuals are in column 8, and percentage errors in column 11.

Once more, Hypothesis 1 was tested for each region by using von Neumann's ratio calculated from the residuals of the differences about the estimate of $\log b^r$. The .05 level for a one-tailed hypothesis was adopted, and the following ratios computed:

Table 10

von Neumann's Ratios

Estimates, Set 3

Region	Ratio
Eastern	1.43
Pittsburgh-Youngstown	2.01
Cleveland-Detroit	2.18
Chicago	1.69
Southern	2.53
Western	1.48

Each of these lies within the region of acceptance, and, consequently, the hypothesis is adopted that the $\underline{\lambda}_t$ and $\underline{\lambda}_t^r$ have been sufficiently reduced by the transformations to allow us to assume randomness in the residuals.

We shall, then, adopt the regional equations of Table 10 and the national equation of Table 7 as those from which estimates will be made. The estimates of steel production in 1970 derived from them are given in Table 11:

Table 11

Estimates of Steel Production, 1970, from Equations, Set 3, and Sets 1 and 2

Region	Estimated Production, Net Tons, Final	Set 1	Set 2
Eastern	33,748,000	32,838,000	39,758,000
Pittsburgh-Youngs.	44,880,000	45,580,000	55,521,000
Cleveland-Detroit	22,196,000	21,047,000	25,982,000
Chicago	35,030,000	34,150,000	39,090,000
Southern	11,381,000	7,304,000	12,665,000
Western	17,903,000	10,953,000	26,322,000
TOTAL.	165,138,000	151,872,000	199,338,000
United States	160,090,000	156,264,000	160,090,000

As an indication of the goodness of fit of these equations, the coefficients of determination[17] for these analyses and the average absolute percentage errors are given in Table 12.

---

[17] The specific interpretation given the variable, time, in this analysis, will be recalled.

Table 12

Coefficients of Determination for Equations, Set 3, and Average Absolute Percentage Errors

Region	Coefficient of Determination	Error
Eastern	.738	7.4
Pittsburgh-Youngstown	.379	6.6
Cleveland-Detroit	.884	7.1
Chicago	.785	5.3
Southern	.982	5.0
Western	.896	7.3
United States	.751	6.0

In interpreting the mean absolute percentage errors, it must be remembered that they are based on deviations of the equations from the data adjusted for short-term fluctuation, not the original data.

Lastly, the rates of growth implied by our equations, regional rates of growth being interpreted as deviations from that of the United States', are given in Table 13.

Table 13

National and Regional Annual Rates of Growth Implied by Equations, Set 3

Region	Growth Rate (Deviation from National)	Actual Growth Rate
Eastern	+ .25 o/o	+ 2.73 o/o
Pittsburgh-Youngstown	- 1.63	+ .85
Cleveland-Detroit	+ 2.03	+ 4.51
Chicago	- .01	+ 2.47
Southern	+ 1.75	+ 4.23
Western	+ 3.86	+ 6.34
United States	0	+ 2.48

4. Interpretation and Qualifications. Our statistical analysis leads us to expect the long-term component in United States steel production to advance output by a compound rate of about 2.5 per cent per year to a total of about 165 million tons in 1970. This represents an advance from 117 million in 1955, and, therefore, a prediction that the long-run component of output will increase by about 50 million tons in 15 years' time. This is an average annual increase in output of about 3.3 million tons. If we use the percentage of production to capacity in 1955--a quite high one for the postwar period generally--of 93 per cent[18], this would imply an average increase in capacity of about 3.5 million tons per year. The trade journal, Steel, reported in 1956 that firm commitments for new steel capacity in the period 1957 through 1959 totalled 14.3 million tons, or 4.7 tons of new capacity per year.[19] This is quite large relative to non-war period increases in capacity for the series under consideration, so that short-period forces may well have played important roles in these plans. The discrepancy, therefore, is not considered disturbing.

Bound as we are to reject any rigorous type of inference, if we apply the average percentage error for the national data to this total of regional production estimates for 1970 we obtain a range of about 155 million to 175 million tons as the long-run growth component in steel production.

In the breakdown of this national total by regions we encounter the imponderables of differential regional growth factors of a wide variety. It was not possible to undertake an intensive study of the economic factors

---

[18] American Iron and Steel Institute, Annual Statistical Report, 1955, 51.

[19] "Steel Expansion Rolls Along", 139 (September 17, 1956), 134-35.

involved, but, fortunately, to have done so would have been merely to duplicate work by other students. The nature of their analyses and conclusions may be sketched briefly here.

The regional location of steel plants--as opposed to their site locations, where the availability of large amounts of cooling water and the existence of flat land for one-story plants may become dominant factors--is transport-oriented. Location, as an aspect of profit-maximization, is the result of absolute differentials in cost factors in production at alternative sites, including the assembling of materials, and shipment of the final product to markets. Absolute differentials in production costs between alternative locations tend to be outweighed by such differentials in transfer costs.[20] Consequently, the minimization of transport costs by varying location tends to minimize total costs.

During the 19th century it became a tenet of economic geography---- "perhaps the most commonly held principle in the whole field..."[21]---- that, on the basis of transport costs, iron moves to coal. It was the contribution of Hartshorne to point out that technological advances in the use of coal had led to a change in the general rule, so that any combination of two factors from coal, iron ore, or market could lead to the location of a steel industry. That is, if we assume transport costs

---

[20] In Weberian terminology, the differential in production costs per unit weight of product between two alternative sites usually falls outside the critical isodapane drawn about the point of minimum transport cost. See A. Weber, Theory of the Location of Industries, (trans. C.J. Friedrich), Chicago, 1928.

[21] R. Hartshorne, "Locational Factors in the Iron and Steel Industry", Economic Geography, 4 (1928), 247.

to be proportionate to weight and distance, and if the sites of iron ore, coal, and market form a triangle in space, then location at one of the corners of the triangle will generally occur only if one of the three weights used per ton of product--amount of ore, amount of coal, and the unit weight of product--exceeds the sum of the other two.[22] The relationship between the weights under present conditions is such, Hartshorne argues, that any two of the three elements will dominate the other, so that if market and coal, market and ore, or coal and ore coincide, the possibility of a steel mill arises.[23]

Isard, in a series of articles, pointed out the relative rise in the pull of the market because of coal-saving technological advances, the increasing use of scrap located at markets, and the higher freight rates on finished products.[24] These long-run technological advances, therefore,

---

[22] Isard has shown that when conditions of weights and relative locations of raw materials and market are such as to dictate a minimum transport point outside the triangle, using a common method of determination of such a point, location will occur at one of the material sites or the market. See W. Isard, Location and the Space-Economy, New York, 1956, 122.
We are implicitly using the index of location developed by W.H. Dean, Jr., The Theory of the Geographic Location of Economic Activities, Ann Arbor, 1938.

[23] Hartshorne, op. cit., 241-52. Without detracting from the contribution of Hartshorne and Isard, we should point out that the impact of the market upon steel location was not totally neglected. For example, M. Keir pointed out in his study of steel that the rise of an iron and steel industry in Pittsburgh began about 1860, at which period the construction of railroads centered about that city. Granted the availability of Connellsville coke in the Monongahela Valley, there also existed a very large market for the product--a consideration often neglected. See M. Keir, Manufacturing Industries in America, New York, 1920, 128-9.

[24] See W. Isard, "Some Locational Factors in the Iron and Steel Industry since the Early Nineteenth Century," Journal of Political Economy, LVI(1948), 203-17; W. Isard and J. Cumberland, "New England as a Possible Location for an Integrated Iron and Steel Works," Economic Geography, 26 (1950).

led to the expectation of a drift of new steel capacity away from coal to the market. This was reinforced by the increasing tendency of production cost differentials between regions to decline, so that transport orientation was enhanced.[25]

Since the Atlantic Coastal Plain from Maine to Virginia accounted for about 20 per cent of the nation's steel consumption in 1946, and since it was a deficit area in terms of production, the expectation was that more capacity would be located in the area after World War II. This was doubly true since tidewater plants could ship to Gulf Coast and Pacific Coast markets at an advantage over Middle West centers.[26] Although Isard and Cumberland brought out the important point that a modern integrated steel plant needed a sufficiently diversified demand for semi-finished steel products as well as an overall demand level of sufficient height, no doubt existed of this characteristic for East Coast demand as a whole.

Two other markets seemed to argue for greater indigenous capacity: the Gulf Coast and the Pacific Coast. The great disadvantage of the latter, even given the existence of integrated capacity at Fontana and Geneva due to the government's desire in World War II to disperse capacity, was its

---

[25] Thus, the point made by L. White, in "The Iron and Steel Industry of the Birmingham, Alabama, District," Economic Geography, 4 (1928), that the production costs in this region being the lowest of all regions would lead to the district becoming a major producer, did not prove true; in the words of P. Craig, "Location Factors in the Development of Steel Centers," Ohio State University, 1956, mimeographed; "The growing markets of the South and Southwest have brought the expansion which low assembly costs alone could not". (7)

[26] M. Barloon, "Steel: the Great Retreat," Harper's, 195 (1947), 145-55. So great is the advantage of water transportation over rail that, in 1947, Geneva, Utah, could ship to Los Angeles at only 14.5 cents per net ton less than Sparrows Point.

lack of large amounts of good quality coal. This disadvantage, coupled with the ability of tidewater plants on the East Coast to reach the market cheaply, may prove the high growth estimates of our study (6.34 per cent per year) overly sanguine. On the other hand, the Texas-Birmingham capacity does not have this penalty and should continue rapid growth.

A second characteristic which seemed to argue for the relative decline of the Midwest after World War II was the approaching depletion of Mesabi ores. In a classic article written just after the war, Barloon predicted, primarily on this ground coupled with the need to import more of our ores from South America and Labrador, that "a new steel industry will have to be built along the Atlantic and Gulf Coasts."[27] Prophet that he was, Barloon failed to judge accurately two developments which should slow down the long-term relative decline in the older steel centers. This first was the economic feasibility of beneficiating taconite. In his recent study, Craig estimates the costs of principal ores delivered per unit of iron at the following levels:

Table 14

Costs of Principal Ores Delivered Per Iron Unit, 1953-54

Ore Source	Lake Erie	Pittsburgh	East Coast
Mesabi	$12.5	$16.6	----
Taconite	12.2	15.6	----
Labrador	13.3	17.2	----
Venezuela	----	15.7	10.9

---

[27] Ibid., 145.

These estimates for Labrador are based upon freight rate estimates of the St. Lawrence Seaway when completed.

When it is recalled that Pittsburgh's ore costs have always been above the East Coast's, these data are quite comforting to Middle West producers. As a matter of fact, Craig estimates that the delivered cost of ore and coal alone is $19.89 per ton of pig iron using taconite and $22.28 using Labrador ore in Pittsburgh, while it is $21.37 using Venezuela ores in Baltimore.[28] This reflects the delivered price per net ton of coal at both places: in 1950, a $.40 per ton charge in Pittsburgh, but a $4.63 charge in Baltimore. An additional advantage is the smaller amounts of coal needed to smelt taconite, due to its high ore content and superior furnace qualities, than any other ore.[29]

The second development that Barloon failed to foresee accurately was the impact of the St. Lawrence Seaway upon ore delivery costs. His belief that costly transshipping charges to transfer Latin American ores from ocean-going vessels to *canallers* would prohibit the entry of these cheap ores may prove to be irrelevant, given the Labrador ores and the ability to use larger ore-carriers when the Seaway is complete.

We shall not, therefore, seek to adjust our estimates for any additional relative penalties on the Pittsburgh-Youngstown, Cleveland-Detroit, or Chicago regions on the account of rising ore costs. It seems clear at the moment that such prognoses cannot be substantiated.

A third possibility that was believed capable of exercising an independent impact upon location was the 1948 Supreme Court decision (FTC vs.

---

[28] P. Craig, *op. cit.*, 11-13.

[29] *Ibid.*, 13.

Cement Institute, 333 U.S. 683) outlawing the basing-point pricing system. It has become a tenet of location theory, however, that, theoretically, and in the long run, a basing point system, while it can provide existing plants which are poorly located with an umbrella, should not interfere with the principles of transport cost minimization in the location of new steel mills, nor can its removal affect these principles. Whether or not basing point systems are operating, plants should still tend to be located at the minimum transportation cost point in order to maximize profits.

It is, then, primarily to the possible shifts in regional production caused by the increasing pull of the market that we should look in qualifying the results of our statistical analysis. In the absence of a thoroughgoing economic analysis of these demand changes, our conclusions can only be provisional. However, our conclusion that this set of factors is, indeed, the major set whose influence must be estimated is borne out by a recent study which sought by questionnaire methods addressed to the management of most United States and Canadian steel firms to judge the relative pull of the market. The results are clearly favorable to our hypothesis.[30]

The percentage distribution of steel production by regions for 1943, 1950, and 1955, as well as those implied by our estimates of 1970, is given in Table 15:

---

[30] See A. Doerr, "A Quantitative Analysis of the Locational Factors in the Integrated and Semi-Integrated Iron and Steel Industry of the United States and Canada," Journal of Geography, LIII (1954), 393-402.

Table 15

Percentage Distribution of Outputs by Regions, Selected Years, and for Estimates of 1970 Output

Region	1943	1950	1955	1970
Eastern	20.3 o/o	19.4 o/o	21.1 o/o	20.4 o/o
Pittsburgh-Youngstown	41.8	38.8	35.1	27.2
Cleveland-Detroit	8.3	9.6	10.4	13.4
Chicago	22.1	21.6	22.4	21.2
Southern	4.3	5.0	5.3	6.9
Western	3.2	5.6	5.6	10.8

Although the Western region's growth in the future may be overestimated, in the absence of more thorough research into the lack of coal as a limitational factor, the competitive position of eastern and Gulf production points in the Pacific Coast markets, and the growth of the market, we feel that the projection should not be changed.

Therefore, these estimates of steel production for 1970 should be accompanied by several intensive spatial analyses before extreme confidence can be put in them. They are working estimates and nothing more.

Table S-1

Steel Production, Adjusted for Short-Term Movements, in the United States and Six Regions, 1943-1955, and Estimates from Three Estimating Equations with Absolute and Percentage Errors of Estimate

(In thousands of net tons)

1. Eastern

Year	Adjusted Prod.	Estimated			Absolute Errors			Percentage Errors		
		Set 1	Set 2	Set 3	Set 1	Set 2	Set 3	Set 1	Set 2	Set 3
1943	18,019	15,945	15,540	16,286	-2,074	-2,479	-1,733	-11.5	-13.8	- 9.6
1944	17,737	16,378	16,090	16,731	-1,359	-1,648	-1,006	- 7.7	- 9.3	- 5.7
1945	16,052	16,822	16,660	17,188	+ 770	+ 608	+1,136	+ 4.8	+ 3.8	+ 7.1
1946	15,510	17,278	17,250	17,659	+1,768	+1,740	+2,149	+11.4	+11.2	+13.9
1947	16,634	17,746	17,861	18,142	+1,112	+1,227	+1,508	+ 6.7	+ 7.4	+ 9.1
1948	17,629	18,228	18,493	18,638	+ 599	+ 864	+1,009	+ 3.4	+ 4.9	+ 5.7
1949	17,202	18,722	19,148	19,148	+1,520	+1,946	+1,946	+ 8.8	+11.3	+11.3
1950	19,075	19,230	19,826	19,672	+ 155	+ 751	+ 597	+ .8	+ 3.9	+ 3.1
1951	20,849	19,751	20,528	20,210	-1,098	- 321	- 639	- 5.3	- 1.5	- 3.1
1952	21,582	20,287	21,255	20,763	-1,295	- 327	- 819	- 6.0	- 1.5	- 3.8
1953	22,735	20,837	22,007	21,330	-1,898	- 728	-1,405	- 8.3	- 3.2	- 6.2
1954	17,783	21,404	22,786	21,915	+3,621	+5,003	+4,132	+20.4	+28.1	+23.2
1955	24,911	21,982	23,593	22,513	-2,929	-1,318	-2,478	-11.8	- 5.3	- 9.9
1956		22,576	24,428	23,129						
1970	------	32,838	39,758	33,748	------	------	------	-----	-----	-----

2. Pittsburgh-Youngstown

Year	Adjusted Prod.	Estimated			Absolute Errors			Percentage Errors		
		Set 1	Set 2	Set 3	Set 1	Set 2	Set 3	Set 1	Set 2	Set 3
1943	37,105	34,477	33,566	35,681	-2,628	-3,539	-1,424	- 7.1	- 7.1	- 3.7
1944	36,940	34,835	34,208	35,985	-2,105	-2,732	- 955	- 5.7	- 7.4	- 2.6
1945	32,599	35,198	34,850	36,303	+2,639	+2,251	+3,804	+ 8.1	+ 6.9	-11.7
1946	32,023	35,563	35,506	36,602	+3,540	+3,483	+4,579	+11.1	+10.9	+14.3
1947	35,045	35,933	36,173	36,914	+ 888	+1,128	+1,869	+ 2.5	+ 3.2	+ 5.3
1948	36,316	36,306	36,853	37,229	- 10	+ 537	+ 913	- 0.0	+ 1.5	+ 2.5
1949	33,820	36,683	37,547	37,547	+2,563	+3,727	+3,727	+ 7.6	+11.0	+11.0
1950	38,124	37,065	38,253	37,867	-1,059	+ 129	- 257	- 2.8	+ .3	- .7
1951	41,535	37,450	38,972	38,190	-3,345	-2,563	-4,085	- 9.8	- 6.2	- 8.1
1952	43,266	37,840	39,705	38,516	-5,426	-3,561	-4,750	-12.5	- 8.2	-11.0
1953	41,244	38,233	40,405	38,845	-3,011	- 839	-2,399	- 7.3	- 2.0	- 5.8
1954	30,459	38,630	41,212	39,176	+8,171	-10,753	-8,717	+26.8	-35.3	-28.6
1955	41,190	39,030	41,986	39,510	-2,160	+ 796	-1,680	- 5.2	+ 1.9	- 4.1
1956		39,436	42,776	39,847						
1970	------	45,580	55,521	44,880	------	------	------	-----	-----	-----

3. Cleveland-Detroit

Year	Adjusted Prod.	Estimates			Absolute Errors			Percentage Errors		
		Set 1	Set 2	Set 3	Set 1	Set 2	Set 3	Set 1	Set 2	Set 3
1943	7,365	6,614	6,415	6,718	- 751	- 950	- 647	-10.2	-12.9	- 8.8
1944	7,366	6,903	6,756	7,021	- 463	- 610	- 345	- 6.3	- 8.3	- 4.7
1945	6,879	7,206	7,115	7,509	+ 327	+ 236	+ 630	+ 4.8	+ 3.4	+ 9.2
1946	7,120	7,521	7,494	7,669	+ 401	+ 374	+ 549	+ 5.6	+ 5.3	+ 7.7
1947	7,261	7,851	7,892	8,014	+ 590	+ 631	+ 753	+ 8.1	+ 8.7	+10.4
1948	7,670	8,195	8,311	8,376	+ 525	+ 641	+ 706	+ 6.8	+ 8.4	+ 9.2
1949	7,809	8,554	8,753	8,753	+ 745	+ 944	+ 944	+ 9.5	+12.1	+12.1
1950	9,382	8,929	9,219	9,148	- 453	- 163	- 234	- 4.8	- 1.7	- 2.5
1951	9,842	9,320	9,709	9,566	- 522	- 133	- 276	- 5.3	- 1.4	- 2.8
1952	10,760	9,728	10,225	9,991	-1,032	- 535	- 769	- 9.6	- 5.0	- 7.1
1953	11,211	10,154	10,778	10,442	-1,057	- 433	- 469	- 9.0	- 3.9	- 4.2
1954	8,788	10,599	11,342	10,913	+1,811	-2,554	-2,125	+20.6	+29.1	+24.2
1955	12,172	11,063	11,945	11,405	-1,109	- 227	- 767	- 9.1	- 1.9	- 6.3
1956		11,548	12,580	11,918						
1970	------	21,047	25,982	22,196	------	------	------	-----	-----	-----

4. Chicago

Year	Adjusted Prod.	Estimates			Absolute Errors			Percentage Errors		
		Set 1	Set 2	Set 3	Set 1	Set 2	Set 3	Set 1	Set 2	Set 3
1943	19,608	18,091	17,584	18,146	-1,517	-2,024	-1,462	- 7.7	-10.3	- 7.5
1944	19,733	18,522	18,113	18,593	-1,211	-1,620	-1,140	- 6.1	- 8.2	- 6.1
1945	18,564	18,963	18,657	19,052	+ 399	93	+ 488	+ 2.1	+ .5	+ 2.6
1946	18,155	19,415	19,215	19,522	+1,260	+1,060	+1,367	+ 6.9	+ 5.8	+ 7.5
1947	18,989	19,877	19,793	20,003	+ 112	+ 804	+1,014	+ .6	+ 4.2	+ 5.3
1948	19,682	20,350	20,388	20,497	+ 668	+ 706	+ 815	+ 3.4	+ 3.6	+ 4.1
1949	19,171	20,835	21,000	21,000	+1,664	+1,829	+1,829	+ 8.7	+ 9.5	+ 9.5
1950	21,259	21,333	21,631	21,520	+ 74	+ 372	+ 261	+ .3	+ 1.7	+ 1.2
1951	22,521	21,838	22,280	22,050	- 683	- 241	- 471	- 3.0	- 1.1	- 2.1
1952	23,303	22,358	22,950	22,594	- 945	- 353	- 709	- 4.1	- 1.5	- 3.0
1953	24,526	22,890	23,641	23,151	-1,636	- 885	-1,375	- 6.7	- 3.6	- 5.6
1954	20,748	23,435	24,359	23,723	+2,687	+3,611	+2,975	+13.0	+17.4	+14.3
1955	26,267	23,993	25,080	24,318	-2,274	-1,187	-1,949	- 8.7	- 4.5	- 7.4
1956		24,565	25,833	24,907						
1970	------	34,150	39,090	35,030	------	------	------	-----	-----	-----

5. Southern

Year	Adjusted Prod.	Estimates			Absolute Errors			Percentage Errors		
		Set 1	Set 2	Set 3	Set 1	Set 2	Set 3	Set 1	Set 2	Set 3
1943	3,849	3,593	3,514	3,624	-256	-335	-225	- 6.7	- 8.7	- 5.8
1944	4,109	3,746	3,685	3,780	-363	-424	-329	- 8.8	-10.3	- 8.0
1945	3,787	3,906	3,865	3,944	+119	+ 78	+157	+ 3.1	+ 2.1	+ 4.1
1946	3,665	4,072	4,052	4,116	+407	+387	+451	+11.1	+10.5	+12.3
1947	4,004	4,246	4,250	4,293	+242	+246	+289	+ 6.0	+ 6.1	+ 7.2
1948	4,398	4,427	4,456	4,479	+ 29	+ 58	+ 81	+ .7	+ 1.3	+ 1.8
1949	4,344	4,615	4,673	4,673	+271	+329	+329	+ 6.2	+ 7.6	+ 7.6
1950	4,895	4,812	4,900	4,875	- 83	+ 5	- 20	- 1.7	+ .0	- .4
1951	5,051	5,017	5,138	5,085	- 34	+ 87	+ 34	- .7	+ .2	+ .7
1952	5,454	5,231	5,388	5,306	-223	- 66	-148	- 4.1	- 1.2	- 2.7
1953	5,874	5,454	5,650	5,536	-420	-224	-338	- 7.2	- 3.8	- 5.8
1954	5,178	5,686	5,925	5,776	+508	+747	+598	+ 9.8	+14.4	+11.5
1955	6,261	5,928	6,213	6,026	-333	- 48	-235	- 5.3	- .7	- 3.8
1956		6,181	6,515	6,287						
1970	-----	7,304	12,665	11,381	----	----	----	-----	-----	-----

6. Western

Year	Adjusted Prod.	Estimates			Absolute Errors			Percentage Errors		
		Set 1	Set 2	Set 3	Set 1	Set 2	Set 3	Set 1	Set 2	Set 3
1943	2,798	3,108	3,051	3,408	+ 310	+ 253	+ 610	+11.1	+ 9.0	+21.8
1944	3,577	3,347	3,305	3,625	- 230	- 272	+ 48	- 6.4	- 7.6	+ 1.3
1945	3,599	3.604	3,580	3,854	+ 5	- 19	+ 255	+ .1	- .5	+ 7.1
1946	3,577	3,881	3,877	4,098	- 304	- 300	+ 521	- 8.5	- 8.4	+14.6
1947	4,370	4,180	4,199	4,358	- 190	- 171	- 12	- 4.3	- 3.9	- .3
1948	4,793	4,501	4,548	4,634	- 292	- 245	- 159	- 6.1	- 5.1	- 3.3
1949	4,738	4,848	4,926	4,928	+ 110	+ 188	+ 190	+ 2.3	+ 4.0	+ 4.0
1950	5,490	5,220	5,335	5,240	- 270	- 155	- 250	- 4.9	- 2.8	- 4.6
1951	6,204	5,622	5,778	5,572	- 582	- 426	- 632	- 9.4	- 6.9	-10.2
1952	6,767	6,054	6,258	5,925	- 713	- 509	- 842	-11.8	- 7.5	-12.4
1953	6,722	6,520	6,778	6,300	- 202	+ 56	- 422	- 3.0	+ .8	+ 6.3
1954	5,363	7,021	7,341	6,699	+1,658	+1,978	+1,336	+30.9	+36.9	+24.9
1955	6,515	7,561	7,951	7,124	+1,046	+1,436	+ 609	+16.1	+22.0	+ 9.3
1956		8,143	8,612	7,575						
1970	-----	10,953	26,322	17,903	------	------	-----	-----	-----	-----