3639

THE UNIVERSITY OF CHICAGO

DATE June 28, 1957

Griliches, Zvi — Author

September 12, 1930 — Birth Date

Hybrid Corn: An Exploration in Economics of Technological Change — Title of Dissertation

Department or School	Degree	Convocation
Economics	Ph.D.	August, 1957

Permission is herewith granted to the University of Chicago to make copies of the above title, at its discretion, upon the request of individuals or institutions and at their expense.

Date filmed	Number of pages	Signature of writer
8/20/1957	81	Zvi Griliches

Extensive Quotation or Further Reproduction of This Material by Persons or Agencies Other than the University of Chicago May Not Be Made without the Express Permission of the Writer.

SHORT TITLE: Hybrid Corn

IRREGULAR NUMBERING ☐

OVERSIZED SHEETS ☐

PAID { By cash ☐ / By thesis deposit ☐

DATE BILLED ____________

THE UNIVERSITY OF CHICAGO

HYBRID CORN: AN EXPLORATION IN ECONOMICS OF TECHNOLOGICAL CHANGE

A DISSERTATION SUBMITTED TO

THE FACULTY OF THE DIVISION OF THE SOCIAL SCIENCES

IN CANDIDACY FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

DEPARTMENT OF ECONOMICS

BY

ZVI GRILICHES

CHICAGO, ILLINOIS

AUGUST, 1957

ACKNOWLEDGMENT

I am indebted to T.W. Schultz for arousing my interest in this problem and for his continuous encouragement, to H.G. Lewis and A.C. Harberger for their valuable advice and guidance, and to the members of the Public Finance Workshop and other members of the Department of Economics of the University of Chicago, both faculty and students, for their suggestions and criticisms. I also want to acknowledge and thank the people directly connected with hybrid corn in the agricultural experiment stations, in the U.S.D.A., and in the private seed companies for their complete cooperation.

I wish to acknowledge the generosity of the Field Crop Statistics Branch of the AMS which made available to me a large part of the unpublished data used in this work.

This work was begun during my tenure as a Research Training Fellow of the Social Science Research Council. It has been supported by the Office of Agricultural Economics Research at the University of Chicago and by a generous grant from the National Science Foundation.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

CHAPTER I

INTRODUCTION

This study is an attempt to understand a body of data: the percentage of all corn acreage planted with hybrid seed, by states and by years. By concentrating on a single, major, well defined, and reasonably well recorded development--hybrid corn--I hope to learn something about the ways in which technological change is generated and propagated in U.S. agriculture.

> Hybrid corn is the product of a controlled, systematic crossing of specially selected parental strains called "inbred lines." These inbred lines are developed by inbreeding, or self-pollinating, for a period of four or more years. Accompanying inbreeding is a rigid selection for the elimination of those inbreds carrying poor heredity, and which, for one reason or another, fail to meet the established standards.[1]

> [The inbred lines] are of little value in themselves for they are inferior to open-pollinated varieties in vigor and yield. When two unrelated inbred lines are crossed, however, the vigor is restored. Some of these hybrids prove to be markedly superior to the original varieties. The development of hybrid corn, therefore, is a complicated process of continued self-pollination accompanied by selection of the most vigorous and otherwise desirable plants. These superior lines are then used in making hybrids.[2]

While the idea of increased vigor from hybridization dates back to Darwin and earlier, serious work on hybrid corn did not

---

[1] N.P. Neal and A.M. Strommen, Wisconsin Corn Hybrids (Madison: Wisconsin Agricultural Experiment Station, Bulletin 476, Feb., 1948), p. 4.

[2] R.W. Jugenheimer, Hybrid Corn in Kansas (Manhattan: Kansas Agricultural Experiment Station, Circular 196, Feb., 1939), pp. 3-4. See also the references in the next footnote.

begin until the first years of this century.[1] The actual idea of increasing the yield of corn through the development and crossing of pure inbred lines was first suggested by G.H. Shull in 1908. He and E.M. East, who had been working independently on the problem and had reached similar conclusions, advocated the crossing of pure inbred lines as a means of improving upon the existing corn varieties, but were quickly discouraged by the difficulties involved in the commercial production of such seed. The crossed seed would have had to be produced on one of the parent inbred lines, resulting in low yields, in seeds having undesirable kernel sizes, unsuited for use in ordinary corn-planting machinery, and in very expensive seed in general. However, several generations of East's students kept working away at the problem and in 1918 Donald F. Jones, working at the Connecticut Agricultural Experiment Station, suggested the use of the double-cross, a cross between two single crosses, as a solution to these difficulties.

> The double cross [was] . . . a brilliant and highly effective, and by no means purely accidental, solution to the most important problem in hybrid corn breeding, the problem of seed production. It [was] . . . an invention which marked an important turning point in the history of hybrid corn, for it made the difference almost at once between hybrid corn as a theoretical possibility and hybrid corn as a practical reality.[2]

---

[1]A popular history of hybrid corn can be found in A.R. Crabb, The Hybrid Corn Makers: Prophets of Plenty (Brunswick: Rutgers University Press, 1948). See also F.D. Richey, "The Lay of the Corn Huckster," Journal of Heredity, XXXIX (1948), 177-180; G.F. Sprague, "The Experimental Basis for Hybrid Maize," Biological Reviews, XXI (1946), 101-120; M.T. Jenkins, "Corn Improvement," in U.S. Department of Agriculture Yearbook, 1936 pp. 455-522; and H.A. Wallace and W.L. Brown, Corn and Its Early Fathers (East Lansing: Michigan State University Press, 1956).

[2]Mangelsdorf, op. cit., p. 178.

Beginning shortly after 1920 there was a very rapid expansion in the number of inbreeding projects. In 1922, at the instigation of Henry A. Wallace, whose father was then the Secretary of Agriculture, F.D. Richey was placed in charge of corn investigations in the Division of Cereal Crops and Diseases, Bureau of Plant Industry, U.S. Department of Agriculture. Richey believed in the future of hybrids and was instrumental in expanding the breeding activities of the U.S.D.A. and in co-ordinating those of the different experiment stations into a more definite national program. In 1925 a cooperative breeding program involving the U.S.D.A. and Corn Belt agricultural experiment stations was inaugurated with the support of Purnell Act funds. This program was very important. It co-ordinated the breeding programs of the various stations and provided channels for a very free interchange of information and breeding materials.

During the same period hybrid corn began to be discussed in the farm press. Of particular importance in arousing the interest of various people in hybrid corn were the articles published by Donald F. Jones and Henry A. Wallace. In an article for the Prairie Farmer, published on March 21, 1925, entitled "The Revolution in Corn Breeding," Henry A. Wallace summarized the state of research at that date and made an accurate prophecy about its future: "No seed company, or farmer, or experiment station has any inbred seed or cross of inbred seed for sale today. The revolution has not come yet, but I am certain that it will come within 10 or 15 years."

The revolution began gathering momentum in the early 1930's. While some hybrid seed was sold commercially in Connecticut in

1920, the revolution had to be transplanted to the heart of the Corn Belt before it caught fire. In 1924 the seed of a single cross was sold for the first time in Iowa. In 1926 the first company devoted exclusively to hybrid seed corn production, the Pioneer Hi-Bred Co., was founded by Henry A. Wallace. At the same time, several commercial producers of open pollinated seed began inbreeding and developing hybrids from their own and experiment station inbreds. Prominent among these were the Funk Bros. Seed Company of Bloomington, Illinois, the DeKalb Agricultural Association, and Lester Phister of El Paso, Illinois. However, none of these efforts matured until the 1930's. Only in 1931 did Pioneer distribute a hybrid which lasted more than a few years. Hybrids developed by the Iowa and Indiana Experiment Stations were first distributed to farmers in commercial quantities in 1933. In 1934 and 1935 other Corn Belt experiment stations entered the scene with their own hybrids. In 1935 the DeKalb Agricultural Association produced for the first time a substantial amount of hybrid seed (14,500 bushels) and found itself on the way to the top of the hybrid seed corn industry.

Once the development got started, it grew by leaps and bounds. More money was appropriated for research by various experiment stations. Stations began to release new hybrids almost every year. The number of commercial hybrid seed companies mushroomed, with almost everybody scrambling to get on the bandwagon. For example, in 1935 only five different producers of hybrid seed had entries in the sixth district of the Iowa Corn Yield Test. In 1938 there were 27; in 1940, 45; and in 1941 a peak was reached with 50 different firms submitting entries. During the same

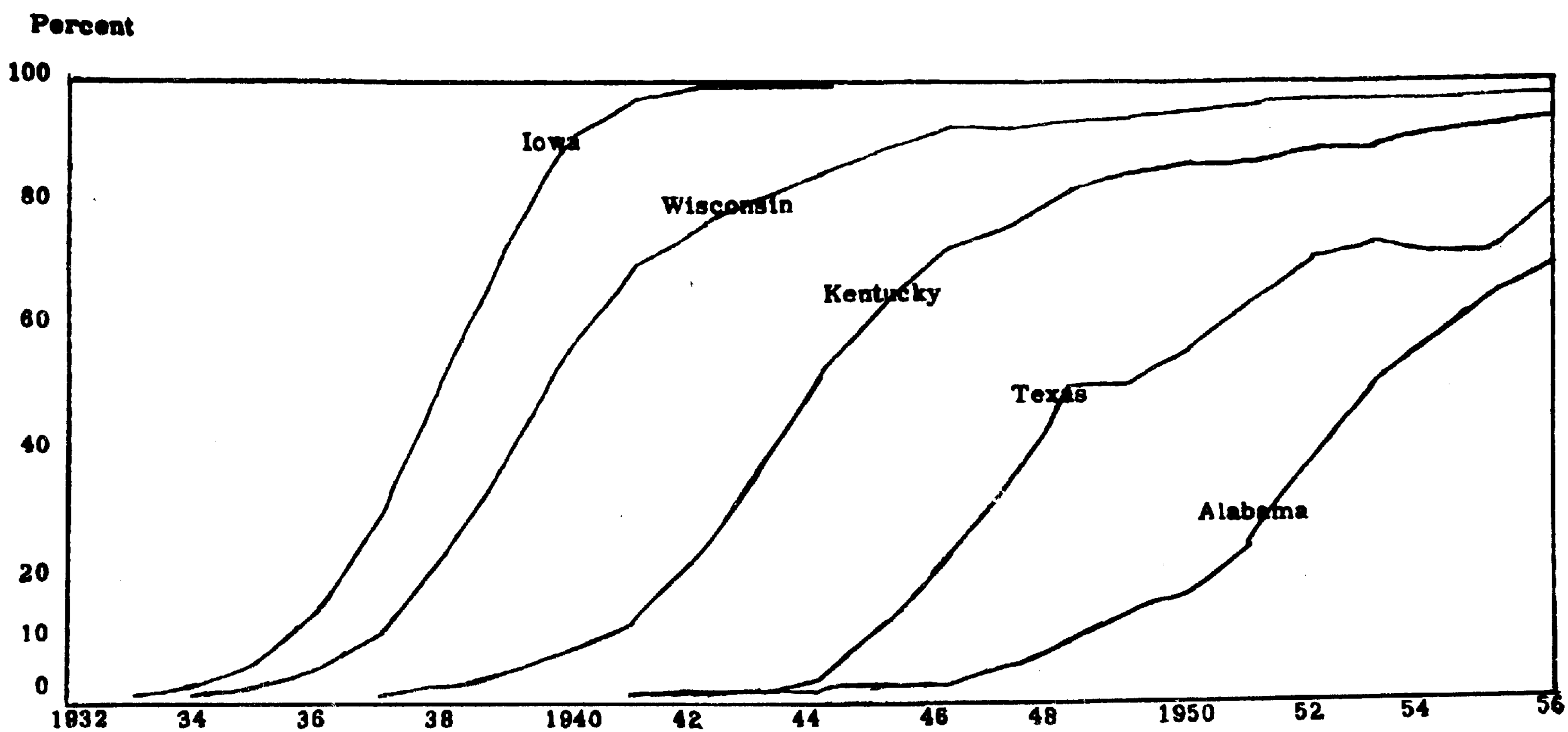

Fig. 1. PER CENT OF ALL CORN ACREAGE PLANTED WITH HYBRID SEED. SOURCE: AGRICULTURAL STATISTICS

period, the percentage of the total corn acreage planted with hybrid seed in Iowa rose from 6 per cent in 1935 to 90 per cent in 1940. The development in other Corn Belt states was similar to that of Iowa.

Figure 1 illustrates the development in Iowa and other states. While the spread of hybrid corn throughout the Corn Belt and the rest of the nation was quite rapid, there have been, nonetheless, marked geographic differences in this development. In particular, the southern states were substantially later in getting adaptable hybrids and also slower in accepting them. Hybrid corn was not a single invention immediately adaptable everywhere. It was an invention of a method of inventing, a method of breeding superior corn for specific localities. The actual breeding of adaptable hybrids had to be done separately for each area. Hence, besides trying to explain differences in the rate of adoption of hybrids by farmers--the "acceptance problem--I will also attempt to explain the lag in the development of adaptable hybrids for specific areas--the "availability" problem.

In the next chapter, the method used to summarize the data will be outlined. Essentially it will consist of fitting trend functions (the logistic) to the data by states and crop reporting districts, reducing thereby the differences among areas to differences in estimates of three parameters: "origins," "slopes" and "ceilings." In Chapter III it will be shown that the lag in the development of adaptable hybrids for particular areas and the lag in the entry of seed producers into these areas (differences in "origins") can be explained on the basis of a varying profitability of entry, "profitability" of entry being a function of market density, and innovation and marketing cost. Chapters IV and V will analyze the differences in the equilibrium levels of hybrid corn use ("ceilings") and the differences in the rates of approach to these levels ("slopes"), and show that these can be explained, at least in part, by differences in the profitability of the shift from open pollinated to hybrid varieties in various parts of the country. Finally, the results will be summarized in Chapter VI and an attempt will be made to draw some more general implications from them.

CHAPTER II

THE METHOD OF ANALYSIS

A graphical survey of the data by states and crop reporting districts along the lines of Figure 1 led to the conclusion that nothing would be gained by trying to explain each observation separately, as if it had no antecedent.[1] It became obvious that the observations are not points of equilibrium which may or may not change over time, but points on an adjustment path, moving more or less consistently towards a new equilibrium position. Hence we should phrase our questions in terms of the beginning of the movement, its rate, and its destination. This led to the decision to fit some simple trend functions to the data and concentrate on the explanation of the cross-sectional differences in the estimates of their parameters.

The choice of a particular algebraic form for the trend function is somewhat arbitrary. As the data are markedly S-shaped, several simple S-shaped functions were considered. The cumulative normal and the logistic are used most widely for such purposes.

---

[1]This conclusion was also supported by the results of an attempt to fit a model in which the year to year changes in the percentage planted to hybrid seed were to be explained by year to year changes in the price of corn, price of hybrid seed, the superiority of hybrids in the previous year or two, etc. The trend in the data was so strong that, within the framework of this particular model, it left nothing of significance for the "economic variables to explain. See Appendix B.

As there is almost no difference between the two over the usual range of data,[1] the logistic was chosen because it is simpler to fit and in our context easier to interpret. While there are some good reasons why an adjustment process should follow a path which is akin to the logistic, I do not want to argue the relative merits of the various S-curves.[2] In this work the growth curves serve as a summary device, perhaps somewhat more sophisticated than a simple average, but which should be treated in the same spirit.

The logistic growth curve is defined by

$$P = \frac{K}{1 + e^{-(a + bt)}}$$

where P is the percentage planted with hybrid seed, K the ceiling or equilibrium value, t the time variable, b the rate of growth coefficient, and a the constant of integration which positions the curve on the time scale. Several features of this curve are of interest: it is asymptotic to 0 and K, symmetric around the in-

[1]For a comparison, see C.P. Winsor, "A Comparison of Certain Symmetrical Growth Curves," Journal of the Washington Academy of Sciences, XXII (1932), 73-84; and J. Aitchison and J.A.C. Brown, The Lognormal Distribution (Cambridge University Press, 1957), pp. 72-75.

[2]It may be worthwhile to indicate why it is reasonable that the development should have followed an S-shaped growth curve. The dependent variable can vary only between 0 and 100 per cent. If we consider the development to be an adjustment process, the simplest reasonable time-path between 0 and 100 per cent is an ogive. While the supply of seed can increase exponentially, the market for seed is limited by the total amount of corn planted, and that will act as a damping factor. Also, if we interpret the behavior of farmers in the face of this new, uncertain development as if they were engaged in sequential decision making, the ASN curve will be bell-shaped, and the cumulative will again be S-shaped. See also H. Hotteling, "Edgeworth's Taxation Paradox and the Nature of Demand and Supply Curves," Journal of Political Economy, XL (1932), 577-616. The argument for the logistic is given by R. Pearl, Studies in Human Biology (Baltimore, 1924), pp. 558-583; and S. Kuznets, Secular Movements in Production and Prices (Boston: Houghton Mifflin, 1930), pp. 59-69.

flection point, and the first derivative with respect to time is given by

$$\frac{dP}{dt} = -b\left(\frac{P}{K}\right)(K-P).^{1}$$

The rate of growth is proportional to the growth already achieved and to the distance from the ceiling. It is this last property that makes the logistic useful in so many diverse fields.[2]

There are several methods of estimating the parameters of the logistic.[3] The method chosen involves the transformation of the logistic into an equation linear in a and b. By dividing both sides of the logistic by K-P and taking the logarithm, we get its linear transform, $\log \frac{P}{(K-P)} = a + bt$, allowing us to estimate the parameters directly by least squares.[4] The value

---

[1]For a more detailed description of the logistic and its properties, see Pearl, op. cit.

[2]Perhaps the simplest interpretation of the logistic is given by A. Lotka, Elements of Physical Biology (Baltimore: Williams and Wilkins, 1925), p. 65. We are interested in the general adjustment function, $dP/dt = F(t)$, which is some function of time. Using a Taylor Series approximation and disregarding all the higher terms beyond the quadratic we get a function whose integral is the logistic. The logistic is the integral of the quadratic approximation to the adjustment function.

[3]See Pearl, op. cit.; H.T. Davis, The Theory of Econometrics (Bloomington: Principia Press, 1941), chap. ii; and G. Tintner, Econometrics (New York: John Wiley and Sons, 1952), pp. 208-211 and the literature cited there.

[4]This is a simplification of a method proposed by Joseph Berkson. Berkson's method is equivalent to weighted least squares regression of the same transform with P(K-P) as weights. J. Berkson, "A Statistically Precise and Relatively Simple Method of Estimating the Bioassay with Quantal Response, Based on the Logistic Function," Journal of the American Statistical Association, XLVIII (1953), 565-599, and "Maximum Likelihood and Minimum Chi-Square Estimates of the Logistic Function," ibid., L (1955), 130-162. Berkson proposed this procedure in the context of bio-assay. It is not clear, however, whether the bio-assay model is applicable in our context, nor is it obvious, even in bio-assay, what system of weights is optimal. See also J. Berkson, "Estima-

of K, the ceiling, was estimated crudely by plotting the percentage planted to hybrid seed on logistic graph paper and varying K until the resulting graph approximated a straight line. After adjusting for differences in K, the logistic was fitted to the data covering approximately the transition from 5 to 95 per cent of the ceiling. The observations below 5 and above 95 per cent of the ceiling value were discarded because they are liable to very large percentage errors and would have had very little weight anyway in any reasonable weighting scheme. The period included in the analysis, however, accounts for the bulk of the changes in the data.

The procedure outlined above was used to calculate the parameters of the logistic for 31 states and for 132 crop reporting districts within these states.[1] The states used account for almost all of the corn grown in the U.S. (all states except the West and New England). Out of a total of 249 crop reporting districts only those were used for which other data by crop reporting districts were readily available. Districts with negligible amounts of corn and unreliable estimates of hybrid corn acreage

---

tion by Least Squares and by Maximum Likelihood," Proceedings of the Third Berkeley Symposium on Mathematical Statistics, I (Berkeley: University of California Press, 1956), 1-11. Hence no weights were used. In view of the excellent fits obtained, it is doubtful whether different weighting systems would have made much difference.

[1]Each state is usually divided into nine crop reporting districts numbered in the following fashion:

```
    N
  1 2 3
W 4 5 6 E
  7 8 9
    S
```

were also left out.[1]

The results of these calculations are presented in Tables 1 and 2. Table 1 summarizes the state results, Table 2 the results by crop reporting districts. Time is measured from 1940, and $\frac{(-2.2-a)}{b}$ indicates the date at which the function passed through the 10 per cent value.[2] This date will be identified below with the date of "origin" of the development. Several things are noteworthy about these figures: the high r^2's indicate the excellent fits obtained.[3] The b's, representing the slope of the transform or the rate of adjustment, are rather uniform, becoming lower as we move towards the fringes of the Corn Belt. The values of $\frac{(2.2-a)}{b}$, the date of "origin," indicate that the development started in the heart of the Corn Belt and spread, rather regularly, towards its fringes.[4] The ceiling--K--also declines as we move away from the Corn Belt.

In this section we have succeeded in reducing a large

---

[1]It should be noted that the sum of logistics is not usually a logistic. However, the logistic is also valid for an aggregate, as long as the components are similar in their development. See L.J. Reed and R. Pearl, "On the Summation of Logistic Curves," Journal of the Royal Statistical Society, XC (New Series, 1927), 729-746. How good the approximation is in fact is indicated by the results below.

[2]This is derived by solving the following equation for t^*:

$$\log_e \left(\frac{0.1}{1.0-0.1}\right) = a + bt^*. \text{ As } \log_e (0.1/0.9) = -2.2,$$

$$t^* = \frac{-2.2-a}{b}.$$

[3]These r^2's should be taken with a grain of salt. They are the r^2's of the transform rather than of the original function and give less weight to the deviations in the center. Also, they do not take into account the excluded extreme values. Nevertheless, an examination of the original data indicates that they are not a figment of the fitting procedure.

[4]"Origin" is measured from 1940. Hence, the "origin" in Iowa is placed approximately in 1936, and in Georgia in 1948.

Table 1

HYBRID CORN LOGISTIC TREND FUNCTIONS BY STATES

States	"Origin" $\frac{-2.2-a}{b}$	"Rate of acceptance" b	"Ceiling" K	r^2
N.Y.	- .89	.36	.95	.99
N.J.	-1.48	.54	.98	.90
Pa.	-1.29	.48	.95	.98
Ohio	-3.35	.69	1.00	.97
Ind.	-3.13	.91	1.00	.99
Ill.	-4.46	.79	1.00	.99
Mich.	-1.44	.68	.90	.98
Wis.	-3.52	.69	.91	.99
Minn.	-3.06	.79	.94	.99
Iowa	-4.34	1.02	1.00	.99
Mo.	-3.32	.57	.98	.98
N.D.	- .65	.43	.65	.96
S.D.	- .40	.42	.85	.95
Neb.	- .60	.62	.97	.99
Kan.	.42	.45	.94	.97
Del.	.21	.47	.99	.98
Md.	- .73	.55	.98	.97
Va.	1.60	.50	.92	.97
W. Va.	- .23	.39	.85	.98
N.C.	5.14	.35	.80	.89
S.C.	5.72	.43	.60	.96
Ga.	7.92	.50	.80	.99
Fla.	2.89	.38	.90	.93
Ky.	.08	.59	.90	.99
Tenn.	2.65	.34	.80	.97
Ala.	7.84	.51	.80	.99
Miss.	4.75	.36	.60	.98
Ark.	1.46	.41	.78	.99
La.	4.89	.45	.53	.99
Okla	3.57	.56	.80	.98
Tex.	3.64	.55	.78	.98

$P = \frac{K}{1+e^{-(a+bt)}}$; $\log \frac{P}{K-P} = a + bt$; $t_{1940} = 0$;

N - 6 to 16 Max S_b = .06

Origin = Date of 10 per cent = $\frac{-2.2-a}{b}$; measured from 1940. For example, -4.0 = 1936 and 7.0 = 1947.

Rate of acceptance = Slope = b
Ceiling = K

Table 2

HYBRID CORN LOGISTIC TREND FUNCTIONS BY CROP REPORTING DISTRICTS*

State and C.R. District		"Origin"	"Rate of Acceptance"	"Ceiling"	r^2
Pa.	1	.15	.41	.85	.99
	2	1.16	.49	.90	.99
	3	.76	.46	.91	.98
	4	.44	.47	.92	.99
	5	.11	.62	.95	.98
	6	-1.02	.55	.95	.99
	7	- .63	.40	.90	.98
	8	-1.04	.54	.96	.99
	9	-2.35	.60	.98	.97
Ohio	1	-3.22	1.25	1.00	.99
	2	-2.73	.99	1.00	.98
	3	-1.77	.75	.95	.98
	4	-3.00	.90	1.00	.98
	5	-3.19	.77	1.00	.98
	6	-3.14	.69	.95	.94
	7	-2.69	.88	1.00	.98
	8	-1.78	.60	.95	.99
	9	-1.80	.73	.95	.97
Ind.	1	-3.82	1.15	1.00	.99
	2	-3.60	1.10	1.00	.99
	3	-3.12	1.15	1.00	.99
	4	-3.24	.95	1.00	.99
	5	-2.85	1.07	1.00	.99
	6	-2.63	1.12	1.00	.99
	7	-1.67	.87	1.00	.96
	8	-1.57	.82	1.00	.98
	9	-1.88	.76	1.00	.98
Ill.	1	-4.81	1.13	1.00	.99
	3	-4.59	.98	1.00	.99
	4	-4.16	1.08	1.00	.99
	4a	-2.65	1.09	1.00	.99
	5	-4.68	1.17	1.00	.99
	6	-4.25	1.18	1.00	.99
	6a	-2.46	.91	1.00	.99
	7	- .81	.64	1.00	.97
	9	- .58	.78	1.00	.97

*I am indebted to the Field Crop Statistics Branch of the AMS for the unpublished data by crop reporting districts.

Table 2--Continued

State and C.R. District		"Origin"	"Rate of Acceptance"	"Ceiling"	r^2
Mich.	7	-1.12	.77	.92	.97
	8	-1.04	.89	.92	.98
	9	-1.70	.78	.92	.98
Wis.	1	-2.17	.81	.85	.99
	2	-2.22	.97	.70	.99
	3	-2.42	.93	.60	.99
	4	-3.24	.67	.95	.96
	5	-2.54	.61	.90	.98
	6	-3.03	.87	.78	.99
	7	-4.16	.89	.98	.99
	8	-3.55	.88	.95	.99
	9	-3.21	.72	.95	.98
Minn.	7	-3.08	1.36	1.00	.99
	8	-3.66	1.14	1.00	.99
	9	-3.04	1.01	1.00	.99
Iowa	1	-4.39	1.00	1.01	.99
	2	-4.78	1.05	1.00	.99
	3	-4.46	1.00	1.00	.99
	4	-3.71	1.12	1.00	.99
	5	-4.70	1.13	1.00	.99
	6	-5.15	1.09	1.00	.99
	7	-2.74	1.25	1.00	.99
	8	-3.61	1.07	1.00	.99
	9	-4.15	1.10	1.00	.99
Mo.	1	-1.37	1.19	1.00	.97
	2	-1.33	1.15	1.00	.95
	3	-1.27	1.15	1.00	.96
	4	-1.51	.66	.95	.98
	5	- .64	.78	.93	.99
	6	-1.11	.72	.97	.97
	7	.16	.46	.90	.99
	8	.63	.63	.87	.99
	9	- .94	.64	.97	.99
N.D.	9	.40	.74	.85	.96
S.D.	3	- .53	.57	.90	.99
	6	- .71	.85	.93	.99
	9	-1.72	.75	.95	.99
Neb.	3	-2.48	.90	1.00	.99
	5	.36	.82	.93	.99
	6	-2.18	.85	1.00	.99

Table 2--Continued

State and C.R. District		"Origin"	"Rate of Acceptance"	"Ceiling"	r^2
Neb.	7	2.33	.90	.95	.99
	8	1.60	.94	.95	.99
	9	- .77	.91	1.00	.97
Kan.	1	2.68	.41	.95	.95
	2	1.52	.66	1.00	.98
	3	- .88	.72	1.00	.99
	6	- .88	.68	.92	.99
	9	.73	.41	.95	.99
Md.	1	2.92	.37	.97	.94
	2	-1.12	.64	1.00	.97
	8	.88	.48	.98	.98
	9	.40	.60	1.00	.93
Va.	2	.87	.68	1.00	.99
	4	1.51	.61	.98	.93
	5	2.37	.68	.95	.99
	6	2.06	.63	.97	.96
	7	1.21	.29	.80	.85
	8	2.18	.40	.85	.90
	9	1.04	.50	.95	.96
Ky.	1	.67	.89	.95	.97
	2	- .42	.72	.98	.99
	3	.49	.61	.90	.97
	4	- .36	.83	.92	.99
	5	- .77	.78	.90	.99
	6	1.94	.62	.60	.98
Tenn.	1	.76	.29	.85	.97
	2	1.88	.33	.55	.99
	3	2.64	.39	.70	.97
	4	2.53	.43	.75	.96
	5	3.43	.35	.80	.91
	6	2.94	.33	.70	.95
Ala.	1	7.73	.56	.60	.98
	2	6.33	.57	.99	.99
	2a	8.80	.45	.90	.97
	3	7.68	.54	.95	.98
	4	7.45	.42	.50	.95
	5	8.08	.49	.70	.95
	6	8.15	.39	.60	.95
	7	7.84	.58	.85	.97
	8	8.24	.45	.70	.97
	9	8.53	.55	.90	.99

Table 2--Continued

State and C.R. District		"Origin"	"Rate of Acceptance"	"Ceiling"	r^2
Ark.	1	.41	.37	.75	.97
	2	1.98	.40	.82	.98
	3	.68	.50	.82	.99
	4	2.24	.42	.77	.99
	5	1.89	.35	.85	.94
	6	1.54	.35	.80	.99
	7	1.66	.32	.55	.93
	8	2.41	.37	.70	.92
	9	1.88	.33	.85	.99
Okla.	3	2.61	.49	.80	.97
	5	3.62	.55	.90	.97
	6	3.17	.52	.88	.93
	7	4.05	.39	.80	.97
	8	4.85	.67	.90	.98
	9	4.08	.52	.75	.95

mass of data to three sets of variables--"origins," "slopes," and "ceilings." "Thus on the basis of three numbers we are prepared, in principle, to answer all the questions the original data sheet can answer provided that the questions do not get down to the level of a single cell. . . . This is saying a great deal."[1]

The economic interpretation of the differences in the estimated coefficients will be developed in the following sections. The values of the different parameters are not necessarily independent of each other, but for simplicity will be considered separately. Variations in the date of origin will be identified with supply factors, variations in slopes with factors affecting the rate of acceptance by farmers, and variations in ceilings with demand factors affecting the long run equilibrium position. In each case we shall consider briefly the implicit identification problem.

[1]R.R. Bush and F. Mosteller, Stochastic Models for Learning (New York: John Wiley and Sons, 1955), p. 335.

CHAPTER III

THE SUPPLY OF A NEW TECHNIQUE

There is no unique way of defining the date of "origin" or of "availability." Hybrid corn was not a single development. Various experimental hybrids were tried until superior hybrids emerged. After a while, these were again superseded by newer hybrids. Nor is there a unique way of defining "origin" with respect to the growth curve. The logistic is asymptotic to zero; it does not have a "beginning." Nevertheless, it is most important to distinguish between the lag in "availability" and the lag in "acceptance." It does not make sense to blame the southern farmers for being slow in acceptance, unless one has taken into account the fact that no satisfactory hybrids were available to them before the middle forties.

I shall use the date at which an area began to plant 10 per cent of its ceiling acreage with hybrid seed as the date of "origin."[1] The 10 per cent date was chosen as an indicator that the development had passed the experimental stage and that superior hybrids were available to farmers in commercial quantities. The reasonableness of this definition has been borne out by a

[1]The date at which the fitted logistic passes through 10 per cent is given by $Y = (-2.2-a)/b$. As the variation of b is small relative to that of a, small changes in the definition of Y will be in the nature of an additive constant and will rarely change the ranking of the date of "origin" in different areas.

survey of yield tests in various states and has been supported by conversations with various people associated with developments in hybrid corn in the experiment stations and private seed companies.[1] The geographical spread of hybrid corn, as defined above, is illustrated in Figure 2.

"Availability" is the result of the behavior of agricultural experiment stations and private seed companies. If we include the growers of station hybrids in the general term--"commercial seed producers"--then availability is the direct result of the actions of seed producers with the experiment stations affecting it through the provision of free research results and foundation stocks. The activities of the experiment stations serve to reduce the cost of innovation facing the seed producers but the entry decisions are still their own. The date at which adaptable hybrids became available in an area is viewed as the result of seed producers ranking different areas according to the expected profitability of entry and deciding their actions on this basis.[2] The relative profitability of entry into an area

---

[1]This is essentially a definition of "commercial" availability. An attempt was made to measure the date of "technical" availability by going through yield tests and other official publications and noting the first year in which hybrids clearly outyielded the open-pollinated varieties. The rank correlation between this technical definition and the "10 per cent" definition was .93. The average lag between the technical and the commercial availability was approximately two years. Also, preliminary explorations with 1 and 5 per cent definitions, and with the rank of an area rather than the absolute date, indicated that the results are not very sensitive to changes in definition.

[2]Implicitly, we have assumed here that the lag between the entry decision and actual availability is approximately constant or at least independent of other variables under analysis.

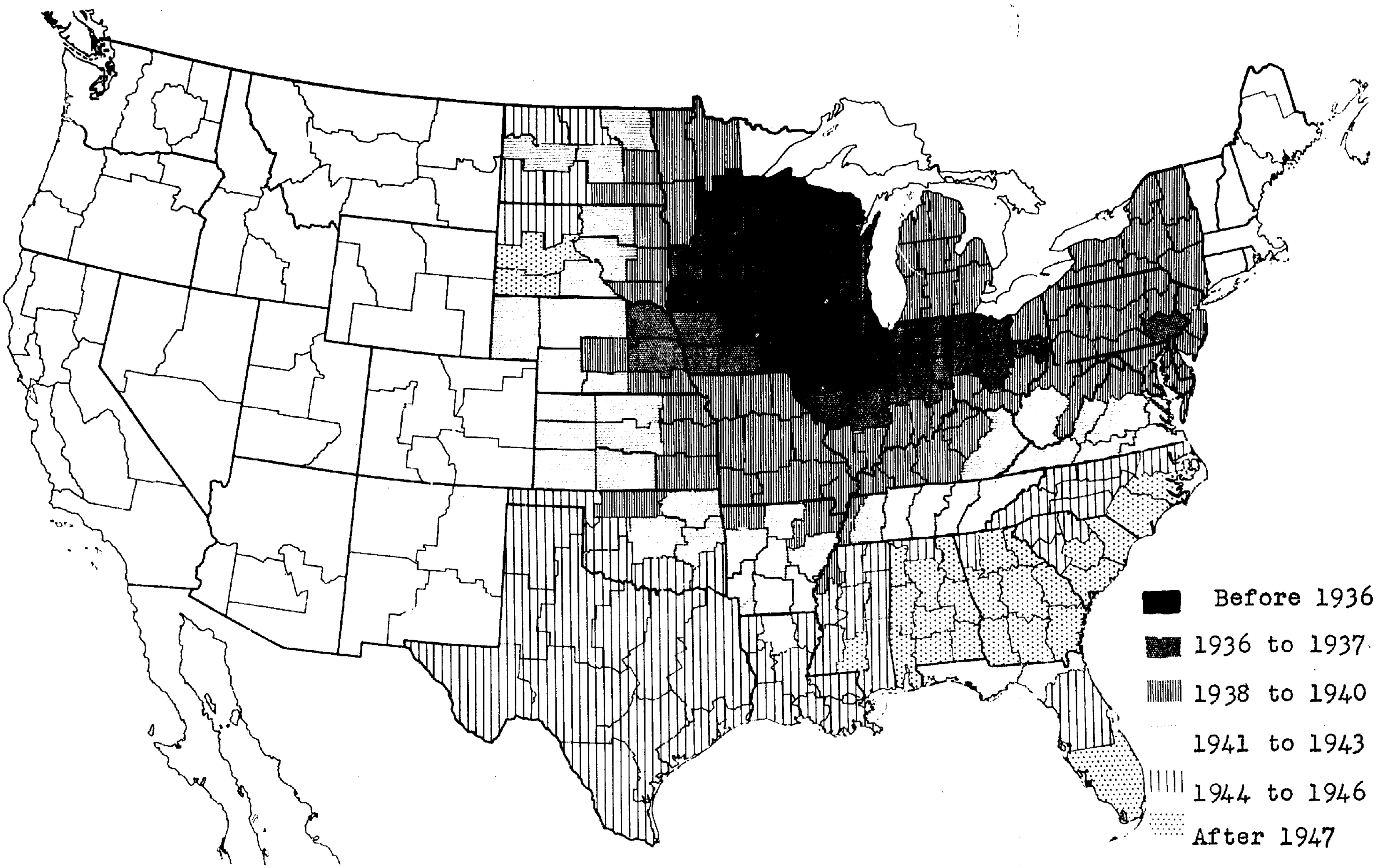

Fig. 2.--The Spread of Hybrid Corn: Estimated year in which 10 per cent of corn acreage was first planted with hybrids.

will depend on the size of the eventual market in that area, marketing cost, the cost of innovating for that area, and (given a positive rate of interest) the expected rate of acceptance.[1]

It is extremely difficult to define "market size" operationally. The definition is not independent of marketing cost or of the particular characteristics of the innovation (the area of adaptability of a particular hybrid) and is complicated by the arbitrariness of the political subdivisions used as the geographic units of analysis. The problem of the "right" geographic unit of analysis, however, will be postponed to the end of this section. As an approximation of the size of the market I shall use the average corn acreage in the area at about the time of entry, adjusted for differences in ceilings. That is, this acreage will be multiplied by .9 if that is the estimate of the fraction of the corn acreage which will be ultimately planted with hybrid seed. Because the political subdivisions are of various and sundry sizes, to make them more comparable the adjusted corn acreage was divided by total land in farms. The resulting variable--

$$X_1 = \frac{\text{(Average corn acreage) x}K}{\text{Total land in farms}}$$

is a measure of "market density" rather than of "market size."[2] If the areas are not too different in size and in the range of adaptability of their hybrids, market density will closely

---

[1]Throughout the paper it is assumed that the price of hybrid seed is given and approximately uniform in different areas. This is a very close approximation to reality and a result of a very elastic long run supply curve of seed.

[2]Differences in seeding rates have been disregarded here. There is, however, some evidence that the results would have been somewhat better if X_1 were adjusted for these differences.

approximate a relevant measure of market size. Also, in its own right, it is important as a measure of marketing cost, the relative cost of selling a given supply of seed in different areas. Under the name of "market potential," such a variable was, in fact, used by at least one of the major seed companies in its decision making process. Executives of the major seed companies, in my conversations with them, emphasized strongly that such a variable was one of the major determinants of their entry decisions.

The importance of marketing cost is underscored by the striking differences in marketing methods of hybrid seed in different parts of the country. While almost 90 per cent of all the seed in the Corn Belt is sold by individual salesmen who call on each farmer, almost all of the seed in the South is sold through stores where the farmer must come and get it. The small size of the corn acreage per farm, the relative isolation of the small farm, and the large proportion of corn on non-commercial farms make the type of marketing used in the Corn Belt prohibitively expensive in the South. The cost of selling a given amount of seed is quite different in various parts of the country, because to sell the same amount many more farmers have to be reached in one area than in another. As a measure of "average size of sale," I used average corn acres per farm reporting corn--X_3.

The estimated slope coefficient--b--was used as a measure of the expected rate of acceptance in different areas. This assumes that producers were able to predict reasonably well the actual rate of acceptance.

There is no good way of estimating the relative costs of innovation. It is probably true that there are no substantial differences in the cost of developing a hybrid from scratch for any corn growing area of the country and, if there were some, they would be swamped by the large differences in returns. A difficulty arises, however, from the fact that a hybrid may be adaptable in more than one area, allowing the cost of innovation to be spread over several areas, and because the experiment stations have borne a substantial part of the innovation cost by developing and releasing inbred lines and whole hybrids. That is, the actual cost of innovating for an area will depend on whether or not hybrids which have already been developed for other areas prove adaptable in this area, and on whether or not the experiment stations have produced and released inbred lines or hybrids adaptable to this area.

Figure 2 illustrates well some of these points. For example, the spread of hybrid corn was much faster in the East-West direction than it was in the North-South direction. One of the important things determining the range of adaptability of a particular hybrid is the length of the growing season and it is largely a function of latitude. Hence, the chances that the same hybrid will be adaptable east or west from the area for which it was originally developed are much higher than north or south of it. The earlier spread of hybrids to the North rather than to the South was at least in part due to the special contributions of the Minnesota and Wisconsin Agricultural Experiment Stations. They had started breeding early and contributed substantially beyond what one would have expected just on the basis of the relative importance of corn in their states. Similarly, one of the reasons that the Southwest had hybrid corn before the Southeast is the special contribution of the Texas and Louisiana Stations and the adaptability of some of the Corn Belt inbred lines there. The Southwest was not as different from the Corn Belt as the Southeast, nor did it suffer from the insect and disease problems which plagued the southeastern corn breeders.

Since most of the early research was done for the area known as the "Corn Belt," other areas benefited from the availability of these research results to a varying degree, depending on the adaptability of Corn Belt inbred lines to those areas. A measure of the degree to which other areas are different from the Corn Belt with respect to the adaptability of Corn Belt lines can be approximated by taking the published pedigrees of the recommended hybrids in different areas in 1956 and computing the percentage of all inbred lines represented by "Corn Belt" lines. An index of "Corn Beltliness"--X_4--was defined as the number of Corn Belt inbred lines in the pedigrees of the recommended hybrids for that area, divided by the total number of lines.[1] To take other aspects of the "complementarity" problem into account, another variable--X_{10}--was defined as the earliest date of entry ("origin") in the immediate (contiguous) neighborhood of the area under con-

[1] On the state level, a published list of recommended hybrids and their pedigrees was used, with Iowa, Illinois, Indiana, Ohio, and Wisconsin lines defined as "Corn Belt" lines. See C.B. Henderson, <u>Inbred Lines of Corn Released to Private Growers from State and Federal Agencies and Double Crosses Recommended by States, Second Revision</u> (Champaign: Illinois Seed Producers' Association, April 15, 1956). On the crop reporting district level, I used unpublished data from the Funk Bros. Seed Co., listing their hybrids by "maturity groups" and giving coded pedigrees.

sideration.[1] The variable X_{10} was introduced on the assumption that it may be cheaper, both from the point of view of the additional research needed and from the point of view of setting up a marketing organization, to enter an area contiguous to an area already entered even though the "market potential" there may be lower than in some other area farther away.

Using either the number of released inbred lines or hybrids or the reported research expenditures, several unsuccessful attempts were made to measure the relative contribution of the various experiment stations. To some extent, however, the impact of this variable is already accounted for by our measures of the "market." The contribution of the various experiment stations is strongly related to the importance of corn in the area. In the "good" corn areas the stations did a lot of work on hybrids, in the marginal areas, less.[2]

The simple correlation coefficients between these variables, on the state level and on the crop reporting district level, are presented in Tables 3 and 4, respectively. All of the correlation coefficients with Y have the expected sign and most of them are also significantly different from zero. However, the intercorrelation among the independent variables prevents us from successfully estimating their separate contributions from these

---

[1]This is analogous to the introduction of a lagged value of the dependent variable into the regression in time series analysis, except that the "lag" here is spatial rather than a time lag.

[2]There are a few exceptions to this statement. In the North, Connecticut, Wisconsin, and Minnesota contributed more than their "share," and so did Texas and Louisiana in the South.

Table 3

CORRELATION COEFFICIENTS ON THE STATE LEVEL--N = 31

	X_1	X_3	b	X_4	X_{10}
Y	-.44	-.35	-.62	-.89	.82
X_1		.52	.77	.55	-.39
X_3			.46	.28	-.36
b				.68	-.51
X_4					-.79

Table 4

CORRELATION COEFFICIENTS ON THE CROP REPORTING DISTRICT LEVEL
N = 132

	X_1	X_3	b	X_4	X_{10}
Y	-.55	-.35	-.70	-.73	.95
X_1		.69	.73	.57	-.56
X_3			.54	.40	-.36
b				.67	-.73
X_4					-.76

Y = Date of "origin." The date an area reached 10 per cent, computed. See Tables 1 and 2.

X_1 = Market density. For states: average corn acreage 1937-46 times K, divided by land in farms in 1945. Similar for crop reporting districts but averaged over different periods, depending on the availability of data. Source: Agricultural Statistics, Census of Agriculture, and published and unpublished materials from state agricultural statisticians.

X_3 = For states, average corn acres per farm, 1939. Source: Census of Agriculture. By crop reporting districts: the same average corn acreage as in X_1, divided by the 1939 or 1945 census number of farms reporting corn, depending on availability of data.

b = The slope of the logistic transform, computed.

X_4 = "Corn Beltliness." The proportion of all inbred lines accounted by "Corn Belt" lines in the pedigrees of recommended hybrid by areas. Source: C.B. Henderson, op. cit.

X_{10} = Earliest date of origin in the immediate neighborhood.

data.[1] Almost all sets and subsets of independent variables in these tables were tried without yielding more than one significant coefficient in each multiple regression.[2] These results are disappointing, particularly because the highest correlations are with the rather artificial variables X_4--"Corn Beltliness" and X_{10}--the "spatial trend."[3] Hence, another approach to the problem was sought.

The trouble with the above approach is that it does nothing about the problem of the "right" geographic unit of analysis. Considering only the "market density" variable, it is obvious that it does not always measure what we want. Markets are continuous. While some areas are poor by themselves, they may be a part of a larger market. Also, an area may be entered because it is a spring-board to other areas rather than on its own grounds. One way of taking these considerations into account is to define the "market potential" of an area as a weighted average of the "market densities" in _all_ areas, densities in other areas weighted inversely to the distance from the area under consideration.[4]

---

[1]Results of regressions of Y on various combinations of the independent variables are presented in Appendix C, Tables 11 and 12.

[2]Similar results were obtained when the logarithms rather than the actual values of the independent variables were used.

[3]The good performance of X_{10} is not surprising. The smaller the geographic unit of analysis, the better will be the relationship between Y and X_{10}. This can be seen by comparing the correlation coefficients on the state and crop reporting district levels. There is, however, another way of rationalizing the performance of X_{10}. See p. 28, n.

[4]See W. Warnz, "Measuring Spatial Association with Special Consideration of the Case of Market Orientation of Production," _Journal of the American Statistical Association_, LI (December, 1956), 597-604.

Given more than a few areas, however, the calculation of such a variable becomes impracticable.[1]

The trouble with our geographic units arises because states are too large while crop reporting districts are too small, and neither corresponds either to technical regions of adaptation of particular hybrids or to the decision units of the private seed companies. It is possible, however, to ask a more modest question: What were the characteristics of the areas entered in a particular year as compared with the characteristics of areas entered in another year? It is possible to aggregate areas according to the year of entry and test the "market potential" hypothesis on these aggregates. I shall define areas according to the year of entry, i.e., all districts with the "origin" in 1939 will make up one such area, and aggregate the data by crop reporting districts into such areas. Given our "10 per cent" definition of "origin," we have sixteen such areas, 1935 to 1950. Alternatively, we would like to define areas according to the adaptability of particular hybrids. However, most hybrids overlap geographically and there is almost no data on the geographical distribution of particular hybrids, but there are breakdowns

---

[1]It does suggest, though, a reason for the good performance of X_{10}. Consider a simple model in which the date of origin is a function of the "true" market measure, the "true" measure being a weighted average of the densities in all areas, weights declining with distance. This "true" measure can be approximated by the actual density in the area and the "true" measure in the immediate neighborhood. But the date of origin in the immediate neighborhood is a function of the "true" density there and can serve as its measure. This implies that X_{10} is another measure of the "market!" For a similar approach in a different context, see M. Nerlove, "Estimates of the Elasticities of Supply of Selected Agricultural Commodities," Journal of Farm Economics, XXXVIII (May, 1956), 500-503.

of the country into "maturity regions." A major seed company breaks down the U.S. and its line of hybrids into eleven "maturity groups," locating the areas of adaptation of these groups on a map. It is possible to aggregate the crop reporting districts into these "technical" regions and ask whether high market areas were entered earlier than others.

The results of these calculations are presented in Table 5. In the aggregation by year of origin, to simplify the calculations, the actual "10 per cent or more" year rather than the calculated date from the logistic was used. For the technical regions the computed origins by districts were used, weighted by the average corn acreage in the district and adjusted for differences in ceilings. That is, aggregate Y = $\frac{\Sigma YAK}{\Sigma AK}$, where A stands for average corn acres, and K is the estimated ceiling. Aggregate X_1 was defined as $\frac{\Sigma AK}{\Sigma L}$, where L stands for total land in farms. Because of the simplicity of the computations involved in this particular approach, ninety more crop reporting districts were added at this point to the analysis, raising the number of included districts to 222. Where separate logistic curves were not computed, Y was estimated by linear interpolation. As the technical regions overlap, each of the aggregates includes a few districts also included in the neighboring aggregates.[1]

To make the results comparable with those presented in Table 4, similar calculations were also performed on X_4 and X_{10}. For the aggregation procedure by "date of origin," X_{10} was

[1]Because one of the "maturity" areas is much larger than the others, it was divided into two on a north-south basis. Hence, we have twelve technical regions in our analysis.

Table 5

CORRELATION COEFFICIENTS BETWEEN THE AGGREGATES OF Y, X_1, X_4 and X_{10}

Aggregation		X_1	X_4	X_{10}
By "Date of Origin": All Areas N = 16				
	Y	-.82	-.98	.95
	X_1		.82	-.64
	X_4			-.93
By "Date of Origin": All Areas except the Southeast N = 13				
	Y	-.94	-.97	.96
	X_1		.90	-.86
	X_4			-.97
By "Technical Regions": N = 12				
	Y	-.69	-.82	.95
	X_1		.90	-.59
	X_4			-.78

defined as the earliest date of origin in the immediate neighborhood of the area defined by the procedure, and X_4 as a simple unweighted average for the districts included in the aggregate. For the aggregation by "technical regions," X_{10} was defined as the lowest weighted average date of origin among the neighboring "technical regions." No aggregation had to be performed on X_4, as it had been originally defined and computed for these regions.

The results presented in Table 5 indicate a strong association between the date of "origin" and average market density in the area, and suggest that the market density variable is more important than is implied by Tables 3 and 4.[1] This association is substantially higher if we exclude the Southeast from the aggregation procedure. This is explained by the relative lateness of the research contributions of the southeastern experiment stations and by the various obstacles put in the way of private seed companies there. Also, after we come down to a certain low level, it does not really pay to discriminate between areas on the basis of X_1; the differences are too small, and other factors predominate. This is brought out when we ask the same question about the association of Y with X_1 within each technical region separately. When regressions of Y on log X_1 were computed for each of the technical regions separately, nine had the expected sign and were significantly different from zero, while the other three were not significantly different from zero.[2] This result is significant on a sign test alone. More interesting, however, is the fact that the r^2's had a rank correlation of .66 with the mean values of X_1 in the respective regressions, indicating that the explanatory power of X_1 is low when X_1 itself is low.

Our aggregation procedure, besides indicating that X_1 is a better variable than is implied by Tables 3 and 4, also helps us with the collinearity problem. Before aggregation, the partial correlation coefficient of Y with X_1, holding X_{10} constant, was

---

[1]These results represent a linear approximation of the relationship between Y and X_1. There are indications that this relationship is of a more complex, non-linear form.

[2]See Appendix C, Table 14.

-.24 on the state level and only -.08 on the crop reporting district level. Now it becomes -.90 for the aggregates by date of origin, -.84 when we leave out the Southeast, and -.64 for the data by technical regions. The regressions of Y on X_1 and X_{10} are presented in Table 6. The coefficient of X_1 has the expected sign and is significantly different from zero for the aggregates by "date of origin" and is almost twice the size of its standard error for the aggregates by technical region.[1]

TABLE 6

REGRESSIONS OF Y ON X_1 AND X_{10}

Aggregation	Coefficients of		R^2
	X_1	X_{10}	
By "Date of Origin": All Areas	-17.8 (2.5)	1.02 (.07)	.982
By "Date of Origin": All Areas except the Southeast	-16.5 (3.4)	1.03 (.07)	.977
By "Technical Region":	-10.5 (5.6)	.88 (.12)	.925

The figures in parentheses are the calculated standard errors.

On the basis of the data presented earlier in this chapter, we found it difficult to discriminate directly between the hypothesis that hybrid corn was first introduced where it was

[1]Comparable regressions with the same variables on the state and crop reporting district level can be found in Appendix C, Tables 12 and 13.

most profitable to do so and the hypothesis that it just "spread." This difficulty arose because market density declines more or less uniformly as we move away from the center. Hence, X_{10}, the "geographic trend" or "proximity" measure will be highly correlated with any reasonable measure of market density. The apparent high correlation between the "date of origin" and X_{10} is then the result of <u>both</u>, real "proximity" variables operating through the cost side of entry and the high correlation of X_{10} with market density.

In spite of these problems, the regression results presented in Table 6 indicate that it is possible to separate the contributions of X_1 and X_{10} if we define our geographic unit correctly. Moreover, we have additional evidence indicating that "market density" is a variable with a contribution of its own: (1) The original invention occurred on the East Coast, in New York and Connecticut. Why did then the commercial development of hybrids begin in the heart of the Corn Belt rather than on the East Coast or someplace else? X_{10} does not explain it, while X_1 does. The development started in the center of the market, where the expected returns from innovation were highest. (2) The same pattern of development was reproduced again on a smaller scale, in Southeastern Pennsylvania and the surrounding area. Why did Pennsylvania crop reporting district 9 reach 10 per cent in hybrids before any of its neighbors? This is not explained by X_{10}, whereas X_1 does explain it. District 9 was a local market peak. It had a higher market density than any of the other districts in the immediate neighborhood.[1] (3) Almost

[1]The only other area which was earlier than all of its

all of the executives of the major seed companies with whom I spoke emphasized the importance of the "size of the market" (which we have tried to approximate by X_1) as one of the main determinants of their entry decisions.

While the quantitative results presented above may not be too conclusive, they, taken together with the information gathered in conversations with executives in the industry and a graphical survey of the data, leave little doubt in my mind that the development of hybrid corn was largely guided by expected pay-off, "better" areas being entered first, even though it may be difficult to measure very well all the variables entering into these calculations.

---

neighbors, district 3 of Florida, was not a local market peak. The early development there was due to the special contribution of the Florida Agricultural Experiment Station, or, more correctly, the contribution of one man--Dr. F.H. Hull. However, while the district was earlier than its neighbors, it did not serve as a center for another geographic wave of hybrid corn, the way district 9 of Pennsylvania did.

CHAPTER IV

THE RATE OF ACCEPTANCE

Differences in the "slope" or adjustment coefficient b will be interpreted as differences in the rate of adjustment of demand to the new equilibrium, and will be explained by variables operating on the demand side rather than by variables operating on the supply side.[1] Actually, the path traced out is an intersection of short run supply and demand curves. However, it is assumed that, while shifts on the supply side determine the origin of the development, the rate of development is largely a demand, or "acceptance," variable.[2] The usefulness of this assumption is due to a very elastic long run supply of seed and is

---

[1]The dimension of b, the adjustment coefficient may be of some interest. It indicates by how much the value of the logistic transform will change per time unit. A value of b = 1.0 implies that the development will go from, for example, 12 to 27 to 50 to 73 to 88 per cent from year to year; that is, the distance from 12 to 88 per cent will be covered in four years. A value of b = 0.5 would imply a path: 12, 18, 27, 38, 50, 62, 73, 82, 88, etc. It would take twice the time, eight years, to transverse the same distance. If one thinks in terms of the cumulative normal distribution positioned on a time scale, which is very similar to the logistic, then b is approximately proportional to $1/\sigma$. A low standard deviation implies that it will take a short time to go from, for example, 10 to 90 per cent, while a higher standard deviation implies a longer period of adjustment.

[2]Implicitly, we have the following model: the potential adjustment path of supply is an exponential function, which after a few years rises quickly above the potential adjustment function of demand. The demand adjustment function has the form of the logistic. The actual path followed is the lower of the two, which after the first few years is the demand path.

supported by the fact that only local and transitory seed shortages were observed. On the whole, the supply of seed was not the limiting factor.[1]

Differences in the rate of acceptance of hybrid corn, the differences in b, are due, at least in part, to differences in the profitability of the changeover from open pollinated to hybrid seed. This hypothesis is based on the general idea that the larger the stimulus the faster the rate of adjustment to it.[2] Also, in a world of imperfect knowledge, it takes time to realize that things have in fact changed. The larger the shift the faster will entrepreneurs become aware of it, "find it out," and hence they will react more quickly to larger shifts.[3]

My hypothesis is that the rate of acceptance is a function of the profitability of the shift, both per acre and total. Per acre profitability may be defined as the increase in yield due to

---

[1]"Clearly it would have been physically impossible for a large percentage of operators to have planted hybrids in the early thirties. There simply was not enough seed. It seems likely, however, that this operated more as a potential than an actual limitation upon the will of the operator, and that rapidity of adoption approximated the rate at which farmers decided favorably upon the new technique," B. Ryan, "A Study in Technological Diffusion," <u>Rural Sociology</u>, XIII (1948), p. 273. Similar views were expressed to the author by various people closely associated with the developments in hybrid corn.

[2]For example, "The greater the efficiency of the new technology in producing returns . . . the greater its rate of acceptance." "How Farm People Accept New Ideas," Special Report No. 15, Agricultural Extension Service, Iowa State College, Ames, November, 1955, p. 6.

[3]This is analogous to the situation in sequential analysis. The ASN (average sample number) is an inverse function of, among other things, the difference between the population means. That is, the larger the difference between the two things which we are testing, the sooner will we accumulate enough evidence to convince us that there <u>is</u> a difference. See A. Wald, <u>Sequential Analysis</u> (New York: John Wiley and Sons, 1947).

the use of hybrid seed, times the price of corn, and minus the difference in the cost of seed.[1] As there is very little relevant cross-sectional variation in the price of corn, the seeding rate, or the price of seed, these will be disregarded and only differences in the superiority of hybrids over open pollinated varieties taken into account.[2]

I shall use two measures of the superiority of hybrids over open pollinated varieties: (1) the average increase in yield in bushels per acre, based on unpublished mail questionnaire data collected by the AMS--X_7, and (2) the long run average pre-hybrid yield of corn--X_8. The latter measure was used on the basis of the widespread belief that the superiority of hybrids can be summarized adequately as a percentage increase.[3] A variation in pre-hybrid yields, given a percentage increase, will also imply a variation in the absolute superiority of hybrids over open pollinated varieties. Twenty per cent is the figure quoted most

---

[1]Of course, hybrids differed from open pollinated varieties not only in yield but also in improved standability, uniformity, and disease resistance. However, there are no good quantitative measures of the improvement in factors other than yield and, besides, most of these were strongly correlated with the yield increases.

[2]The apparent cross-sectional variation in the average price of hybrid seed is largely due to differences in the mix of "public" versus "private" hybrids bought by farmers. The "public" hybrids sell for about $2.00 less per bushel. The rank correlation between the average price paid for hybrid seed and the estimated share of "private" hybrids in 1956 was .73.

[3]The data from experiment station yield tests indicate that this is not too bad an assumption. See Sprague, op. cit., and the literature cited there. It is unfortunate that these data are not comparable between states and, hence, cannot be used directly in this study.

often for this superiority.[1]

Average corn acres per farm, X_3, were used to add the impact of total profits per farm.

As the value of b depends strongly on the ceiling K, to make them comparable between areas, the b's had to be adjusted for differences in K. Instead of b, b' = bK was used as the dependent variable, translating the b's back into actual percentage units from percentage of ceiling units. Alternatively, one could have adjusted the independent variables to correspond only to that fraction of the acres which will eventually shift to hybrids. But there are no data for making such an adjustment, hence b was adjusted to imply the same actual percentage changes in different areas.[2]

Linear and log regressions were calculated for the data from 31 states and 132 crop reporting districts. The results are presented in Table 7.[3] The figures speak largely for them-

---

[1] "If an average percentage increase in yield to be expected by planting hybrids as compared to open pollinated varieties were to be computed at the present it would probably be near 20 per cent. . . ." J.T. Swartz, "A Study of Hybrid Corn Yields as Compared to Open Pollinated Varieties" (Insurance Section, FCIC, Washington, April and May, 1942), unpublished manuscript.

"Experience in other corn-growing regions of the United States shows that increases of approximately 20 per cent over the open pollinated varieties may be expected from the use of adapted hybrids. Results so far in Texas are in general agreement with this figure." J.S. Rogers and J.W. Collier, Corn Production in Texas (Texas Agricultural Experiment Station, Bulletin 746, February, 1952), p. 7.

"Plant breeders conservatively estimate increase in yields of 15 to 20 per cent from using hybrid seed under field conditions. They expect about the same relative increases in both low--and high--yielding areas." U.S.D.A., Technology on the Farm (Washington, 1940), p. 22.

[2] This adjustment affects our results very little. See Table 8, p. 40.

[3] X_7 was not used on the state level because it was felt

Table 7

REGRESSIONS OF "SLOPES" ON "PROFITABILITY" VARIABLES

Regression	Coefficients of X_3	X_7	X_8	R^2
By states--N = 31:				
$b'=c_0+c_3X_3+c_8X_8$	.006 (.002)		.017 (.005)	.66
$\log b'=c_0+c_3\log X_3+c_8\log X_8$	.30 (.08)		.66 (.11)	.67
By crop reporting districts--N = 132:				
$b'=c_0+c_3X_3+c_7X_7$	.0073 (.0008)	.079 (.009)		.57
$b'=c_0+c_3X_3+c_8X_8$	.0076 (.0007)		.016 (.002)	.61
$\log b'=c_0+c_3\log X_3+c_7\log X_7$	.44 (.04)	.70 (.09)		.61
$\log b'=c_0+c_3\log X_3+c_8\log X_8$	.44 (.03)		.57 (.05)	.69

Figures in parentheses are the calculated standard errors.

X_3 - Average corn acres per farm reporting corn.

X_7 - The average difference between hybrid and open pollinated yields by districts, tabulated only from reports showing both and averaged over 4 to 10 years, depending on the overlap of the available data with the adjustment period (10 to 90 per cent). Based on unpublished AMS "Identicals" data.

X_8 - Pre-hybrid average yield. Usually an average for the ten years before an area reached 10 per cent in hybrids. Sometimes fewer years were used, depending on the available data. Source: States - Agricultural Statistics. Crop reporting districts--various published and unpublished data from the AMS and from State Agricultural Statisticians.

selves, indicating the surprisingly good and uniform results obtained. The log form and X_8 rather than X_7 did somewhat better but not significantly so. The similarity of the coefficients in comparable regression is striking. For example, compare the coefficients of X_7 and X_8 in the log regressions and all the coefficients in similar regressions on the state and crop reporting district levels. These results were also similar to those obtained in preliminary analyses using b rather than b' as the dependent variable.[1]

An attempt was made to incorporate several additional variables into the analysis. Rural sociologists have suggested that socio-economic status—or level-of-living is an important determinant of the rate of acceptance of a new technique.[2] The USDA

---

that the aggregation error would be too large. We want an average of differences while I could only get a difference between averages. For some states this difference exceeded the individual differences in all the crop reporting districts within the state.

[1]

Table 8

REGRESSIONS OF UNADJUSTED "SLOPE" ON "PROFITABILITY" VARIABLES

Regression	Coefficients of			R^2
	X_3	X_7	X_8	
$b=c_0+c_3X_3+c_7X_7$ N = 65	.005 (.001)	.06 (.01)		.40
$b=c_0+c_3X_3+c_8X_8$ N = 32	.005 (.001)		.022 (.002)	.75

These were calculated for sub-samples of 65 and 32 crop reporting districts.

[2]See "How Farm People Accept New Ideas," op. cit.; and E.A. Wilkening, "The Acceptance of Certain Agricultural Programs and Practices in a Piedmont Community of North Carolina" (unpublished Ph.D. thesis, University of Chicago, 1949); and Acceptance

level-of-living index for 1939, when added to the regressions by states, had a negative coefficient in the linear form and a positive coefficient in the logarithmic form. In neither case was the coefficient significantly different from zero.

A measure of the "importance" of corn--the value of corn as a percentage of the value of all crops--was added in the belief that the rate of acceptance may be affected by the relative importance of corn within the farmer's enterprise. However, its coefficient was not significantly different from zero. Nor was the coefficient of total capital per farm significantly different from zero. The latter variable was introduced in an attempt to measure the impact of "capital rationing."[1]

The rate of acceptance may also be affected by the "advertising" activities of the extension agencies and private seed companies. However, there are no data which would enable us to take it into account. There is also some evidence that the estimated rate of acceptance will be affected by the degree of aggregation and the heterogeneity of the aggregate. Heterogeneous areas imply different component growth curves and hence a lower aggregate slope coefficient. This is exhibited by the lower state values for b as compared to the values for the individual crop reporting districts within these states. No way has been found, however, to introduce this factor into the analysis.

<u>of Improved Farm Practices in Three Coastal Plain Counties</u> (Tech. <u>Bull. No. 98, North Carolina Agricultural Experiment Station</u>, May, 1952).

[1]The failure of the last two variables is due largely to their strong intercorrelation with the included variables. "Importance" is highly correlated with average yield and capital with corn acres per farm. When used separately, these two variables did as well on the state level as yield and corn acres per farm. See Appendix C, Table 14.

Nevertheless, our results do suggest that a substantial proportion of the variation in the rate of acceptance of hybrid corn is explainable by differences in the profitability of the shift to hybrids in different parts of the country.

CHAPTER V

THE EQUILIBRIUM LEVEL OF USE

I am interpreting the "ceilings" as the long run equilibrium percentages of the corn acreage which will be planted to hybrid seed. Differences in the percentage at which the use of hybrid seed will stabilize are the result of long run demand factors. It is assumed that in the long run the supply conditions of seed are the same in all areas, the same percentage increase in yield over open pollinated varieties at the same relative price.[1] However, this same technical superiority may mean different things in different parts of the country.

The ceiling is a function of some of the same variables which determine b, the rate of acceptance. It is a function of average profitability and of the distribution of this profitability. With the average above a certain value no farmer will be faced with zero or negative profitability of the shift to hybrids. With the average profitability below this level some farmers will be facing negative returns and hence will not switch

[1]This is not a necessary assumption, but it is forced upon us by the lack of consistent data on differences in the superiority of hybrids in different parts of the country (see Appendix A). A fixed percentage superiority is the most simple assumption one can make. Obviously, this assumption did not hold everywhere, and we would have done better if we had taken account of the deviations from it in our analysis. Nevertheless, even without taking them into account, we are still able to explain a substantial proportion of the spatial variability in ceilings.

to hybrids. In marginal corn areas, however, "average profitability" may become a very poor measure. Its components lose their connection with the concepts they purport to represent. Yield variability may overshadow the average increase from hybrids. The relevance of the published price of corn diminishes. In many marginal corn areas there is almost no market for corn off the farm. The only outlet for increased production is as an input in another production or consumption process on the farm. But on farms on which corn is a marginal enterprise, with little or no commercial livestock production, the use of corn is limited to human consumption, feed for draft animals, a cow and a few chickens. The farmer is interested in producing a certain amount of corn to fill his needs, having no use for additional corn. It will pay him to switch to hybrid corn only if he has alternative uses for the released land and other resources which would return him more than the extra cost of seed. But in many of these areas corn is already on the poorest land and uses resources left over from other operations on the farm. Also, there may already be substantial amounts of idle land in the area. All these factors may tend to make hybrids unprofitable although they are "technically" superior. Similarly, in areas where capital rationing is important, the recorded market rate of interest will be a poor measure of the opportunity costs of capital. While the returns to hybrid corn may be substantial, if corn is not a major crop, the returns to additional investments in other branches of the enterprise may be even higher.

Ceilings are not necessarily constant over time. Even without any apparent change in the profitability of the shift from open pollinated to hybrid corn, they may change as the result of a change in the relative profitability of corn growing, an improvement in the functioning of the market for corn, or an increase in storage facilities. Also, in areas where there are large year to year changes in the corn acreage, the percentage planted to hybrid seed may fluctuate as a result of the differential exit and entry of hybrid and open pollinated seed using farmers in and out of corn. These changes may occur without any "real" changes in the relative profitability of hybrids or in farmers' attitudes toward them. It is very difficult to deal statistically with a development composed of a series of adjustments to shifting equilibrium values.[1]

As a first approximation I shall ignore this problem. Only in the marginal corn areas is this of some importance. For most of the Corn Belt the assumption of an immediate ceiling of 100 per cent is tenable. In the fringe areas ceiling values somewhat lower than 100 per cent fit very well. There are some indications that in the South ceilings may have shifted over time, but I doubt that this is important enough to bias seriously our results.

In spite of all these reservations and the crudeness with which the ceilings were estimated in the first place, it is possible to explain a respectable proportion of their variation with the same "profitability" variables that were used in the analysis of "slopes." Because there is a ceiling of 1.00 to the possible

---

[1] I am aware of only one attempt in the literature to deal with this kind of problem. See C.F. Roos and V. von Szelisky, "Factors Governing Changes in Domestic Automobile Demand," particularly the section on "The Concept of a Variable Maximum Ownership Level," in Dynamics of Automobile Demand (New York: General Motors Corporation, 1939), pp. 36-38.

variation in K, the logistic function was used again, giving us logit $K = \log\left[\frac{K}{1-K}\right]$ as our dependent variable. As there were a substantial number of areas with K = 1.0, a value not defined for the transform, two approximations were used. On the state level all values of K = 1.0 were set equal to .99, while on the crop reporting level, where there was no problem of degrees of freedom, these values were left out of the analysis. X_3--average corn acres per farm, and X_8--pre-hybrid yield, were used as "profitability" measures, and X_{11}--capital per farm, was added to take "capital rationing" into account.

The results of these calculations are presented in Table 9. They indicate that differences in measures of average profitability, differences in average corn acres and pre-hybrid yields, can explain a substantial proportion of the variation in "ceilings," the long run equilibrium level of hybrid seed use. The proportion of the variation explained on the state level is substantially higher, indicating that additional variables which may be at work on the crop reporting district level may cancel out on the state level. For example, the coefficient of capital investment per farm, a measure of "capital rationing," is significantly different from zero on the crop reporting district level but not on the state level. Undoubtedly this analysis could be improved by the addition of other variables but I would not expect it to change the major conclusion appreciably.

Table 9

REGRESSIONS OF LOGIT K ON "PROFITABILITY" VARIABLES

Regression	Coefficients of X_3	Coefficients of X_8	Coefficients of X_{11}	R^2
By states--N=31:				
$c_0+c_3X_3+c_8X_8$	.03 (.01)	.11 (.02)		.71
$c_0+c_3\log X_3+c_8\log X_8$	1.94 (.56)	5.88 (.80)		.71
$+c_{11}\log X_{11}$	1.55 (.84)	5.25 (1.30)	.71 (1.14)	.72
By crop reporting districts--N = 86:				
$c_0+c_3\log X_3+c_8\log X_8+c_{11}\log X_{11}$	1.09 (.48)	2.22 (.61)	1.35 (.64)	.39

Figures in parentheses are the calculated standard errors.

X_3 - Average corn acres per farm.

X_8 - Pre-hybrid yield.

X_{11}- On the state level, value of land and buildings per farm, 1940. Source: Statistical Abstract of the United States, 1948, p. 600. On the crop reporting district level, total capital investment per farm, 1949. Computed from Table 11, E.G. Strand and E.O. Heady, Productivity of Resources Used on Commercial Farms (Washington: U.S.D.A., Technical Bulletin No. 1128, November, 1955), p. 45.

CHAPTER VI

SUMMARY AND CONCLUSIONS

The above analysis does not purport to present a complete model of the process of technological change. Rather the approach has been to break down the problem into manageable units and to analyze them more or less separately. I have concentrated on the longer run aspects of technological change, interpreting differences in the pattern of development of hybrid corn on the basis of the long run characteristics of various areas, and ignoring the impact of short run fluctuations in prices and incomes. This limitation is not very important in the case of hybrid corn because the returns from the changeover were large enough to swamp any short run fluctuations in prices and other variables.[1] It may, however, become serious were we to consider other technical changes requiring substantial investments, and not as superior to their predecessors as was hybrid corn. Nor can we transfer the particular numerical results to the consideration of other developments. Nevertheless, a cursory survey of trends in the number of cornpickers and tractors on farms, and of trends in the use of fertilizer, does indicate that it might be also possible to apply a version of our approach to their analysis.

---

[1]Estimates made for Kansas data indicate returns from 300 to 1000 per cent on the extra cost of seed.

While it is not original, I think it is useful to reiterate that technological changes in U.S. agriculture are gradual and follow a "growth curve" pattern. More important, this underlying pattern of change, at least for changes yielding large returns, is very little affected by short run fluctuations in prices and incomes. That is, nothing much more than "growth" is reflected in the series. Even in the case of such costly and long run investments as tractors, it took the deep depression years of 1932-34 to leave a visible ripple on the otherwise smooth pattern of growth. On the whole, this contradicts the arguments about technology "jumping" in response to higher prices advanced by Cochrane.[1]

Once an invention has occurred, a useful and feasible idea has been published, the process of innovation, of developing the new idea and adapting it to the particular conditions of various markets is largely a process which is guided by economic considerations. In particular, the decision as to which of the possible markets should be supplied first with the new technique is made on the basis of the relative profitability of entry into the various markets.

The rate at which an innovation is accepted is largely a function of the relative profitability of the innovation. Of course the particular characteristics of the innovation are also important, but the importance of relative profitability becomes apparent when we consider differential rates of acceptance of the same innovation in different parts of the country.

---

[1]W.W. Cochrane, An Analysis of Farm Price Behavior, Progress Report No. 50 (State College: Pennsylvania State College, May, 1951).

Similarly, the level at which the use of the new technique is stabilized is largely determined by economic variables. In areas of high average profitability, the changeover is profitable for everybody, and everybody will eventually switch to the new technique. In areas where the average profitability is low, there will be a fraction of the population to whom the switch is actually not worth the cost, and they will persist in the use of the old technique.

In this context one may say a few words about the impact of "sociological" variables. It is my belief that in the long run, and cross-sectionally, these variables tend to cancel themselves out, leaving the economic variables as the major determinants of the pattern of technological change. This does not imply that the "sociological" variables are not important if one wants to know which individual will be first or last to adopt a particular technique, only that these factors do not vary very much cross-sectionally. Partly this is a question of semantics. With a little ingenuity, I am sure that I can redefine 90 per cent of the "sociological" variables as economic variables. Also, some of the variables I used, e.g., yield of corn and corn acres per farm, will be very highly related cross-sectionally to education, socio-economic status, level-of-living, income, and other "sociological" variables. That is, it is very difficult to discriminate between the assertion that hybrids were accepted slowly because it was a "poor corn area" and the assertion that the slow acceptance was due to "poor people." Poor people and poor corn are very closely correlated in the U.S. Nevertheless, one may find a few areas where this is not so. Obviously, the slow acceptance of hybrids on the western fringes of the Corn Belt, in western Kansas, Nebraska, South Dakota, and North Dakota is not due to poor people, but, the result of "economic" factors, poor corn area.

There is an interesting implication in some of the above. Hybrid corn was an innovation which was more profitable in the "good" areas than in the "poor" areas. This, probably, is also a characteristic of many other innovations besides hybrid corn. Obviously, tractors contribute more on larger than on smaller farms, and so forth. Hence, there may be a tendency for technological change to accentuate regional disparities in levels of income and rates of growth. Moreover, this tendency is reinforced by the economics of the innovation process, which results in the new techniques being supplied to the "good" areas before they are supplied to the "poorer" areas, and also because these techniques are accepted faster in the same areas. As Haavelmo has shown, a lag of this sort by itself can cause long run regional differences in levels of income.[1] The kinds of inventions we get, and the process by which they are distributed, may lead to the aggravation of the already serious problem of regional differentials in levels of income and growth.[2]

Looking at the hybrid seed industry as a part of the specialized sector which provides us with technological change, it can be said that both private and public funds were allocated

---

[1] T. Haavelmo, A Study in the Theory of Economic Evolution (Amsterdam: North Holland Publishing Company, 1954).

[2] If, however, areas and people that are "good" with respect to one of the new techniques are "bad" with respect to another technique, then these tendencies may cancel out.

efficiently within that sector. Given a limited set of resources, the hybrid seed industry expanded according to a pattern which made sense, allocating its resources first to the areas of highest returns. In this context, the importance of the free interchange and dissemination of research results between experiment station workers and private corn breeders in the Corn Belt cannot be overestimated. It illustrates clearly the gains from a free exchange of knowledge. Not all was a bed of roses, however. In some areas, experiment stations tried to hang onto their research results, their inbred lines, restricting the distribution of "their" lines only to special sectors of the industry, the producers of "station" hybrids. Even more deleterious was the use, in some areas, of state police powers to support a local "infant" seed industry, and prevent the entry of "big," "foreign" private seed companies. Interference with the free market, in this as in other areas, was not particularly beneficial either to farmers or to the economy at large.

With respect to hybrid corn specifically, the contribution of this study is in the improved understanding of a body of data. What started out as a puzzle about the peculiar patterns in the data has been answered. The level of use of hybrid seed is largely determined by the date on which superior hybrids become available for an area, and the rate at which they are accepted by farmers. The date at which hybrids became available in various areas is the result of the relative contributions of the various experiment stations and the activities of the private companies guided by the profit motive. For example, the South was late in getting hybrids both because marketwise it was substantially inferior to other areas, and because (for similar reasons) the southern experiment stations did not produce anything of importance until the middle forties. Similarly, the rate of acceptance of hybrids by farmers in the Corn Belt was substantially higher than in the South because the absolute profitability of the changeover to hybrids was much higher there too.

The American farmer appears to have adjusted rationally to these new developments. Where the profits from the innovation were large and clear cut, the changeover was very rapid. It took Iowa farmers only four years to increase their acreages in hybrid corn from 10 to 90 per cent. In areas where the profitability was lower, the adjustment was also slower. On the whole, taking account of uncertainty and the fact that the spread of knowledge is not instantaneous, farmers have behaved in a fashion consistent with the idea of profit maximization. Where the evidence appears to indicate the contrary, I would predict that a closer examination of the relevant economic variables will show that the change was not as profitable as it appeared to be.

APPENDIX A

SOURCES AND LIMITATIONS OF DATA

I. The percentage of total corn acreage planted with hybrid seed.

The percentage planted with hybrid seed in each state and year since 1933, is given in various volumes of Agricultural Statistics. The figures by crop reporting districts have not been published, except by a few states. They were furnished to me by the Field Crops Statistics Branch of the AMS. These percentages are estimated by State Agricultural Statisticians and the Crop Reporting Board on the basis of answers to the annual mail questionnaire sent out to farmers.[1] The answers to the questionnaire are tabulated and then adjusted for non-response and other factors. Besides the usual limitations and biases common to most of our agricultural statistics, a particular bias, of special interest to us, is introduced into the data by the estimating procedure. In essence, the Board's instructions to the

[1]The general method of collecting these statistics is described in BAE, USDA, Agricultural Estimating and Reporting Services of the USDA, Misc. Pub. 703 (Washington, 1949). A relevant critique of these methods can be found in Ivan M. Lee, "A Critical Evaluation of Available Agricultural Statistics," reprinted in Crop Estimating and Reporting Services of the Department of Agriculture, Report and Recommendations of a special subcommittee of the Committee on Agriculture of the House of Representatives, 82nd Congress, 2nd Session (U.S. Government Printing Office, 1952), pp. 54-61. For our purposes, it will suffice to note that these estimates are based on returns from mail questionnaires sent to a biased list and characterized by a very high rate of non-response (about 0.7).

State Statisticians read: "When in doubt, use a growth curve to adjust the data." Hence, the good fits of the logistics may have been built into the data by the estimating procedure. However, having read through a file of correspondence between the Crop Reporting Board and the State Agricultural Statisticians, I do not think that this source of bias is very important. No specific growth curve was prescribed, just the general idea of one, and both from the correspondence and the actual data it appears that very high weight was given to the actual tabulations of the mail responses. Also, the actual data is full of cross-sectional and year to year variations, indicating that the "smoothing by a growth curve" did not play a major role in the estimating procedure. Nevertheless, because of this and also because the sample size is smaller, the crop reporting district estimates are less reliable than the state estimates. For similar reasons, the estimates of the percentage planted with hybrids in marginal and unimportant corn areas are less reliable than the estimates in other areas. On the whole, however, these data seem to reflect very well the actual history of hybrid corn, particularly in the major corn areas.

II. The superiority of hybrids over open pollinated varieties.

Three specific sources of data are available for this purpose:

(1) During 1939-41, as a part of its annual "Corn Production Data," the AAA collected data on yields of open pollinated and hybrid corn, by counties, in the "Commercial Corn Area." While it is not too clear from the records (the data were never

published) how these data were collected, my impression is that they were the result of an actual enumeration of all the farms in the Commercial Corn Area.

(2) Since 1939 the AMS has been collecting separate data on yields of open pollinated and hybrid corn. These data were never published and their collection was discontinued in 1953. They are a summary of answers to a judgment question added to the December General Schedule. These data were summarized by crop reporting districts and states for an increasing number of states. To eliminate bias a special tabulation was made including only those reports which reported both hybrid and open pollinated yields. It was felt by the AMS that this would give a better indication of the differential between hybrid and open pollinated yields. This special tabulation will be referred to as "Identicals."[1]

(3) Since the 1930's, beginning with the Iowa Agricultural Experiment Station, almost all of the agricultural experiment stations have conducted some kind of a yield trial comparing the yields of various hybrids and open pollinated varieties. The data from these tests will be referred to as AES material. These data raise several difficult problems. They represent results on one or several fields in a whole state, conducted under varying and better than average conditions. The relation between the experiment station results and what the farmer may expect on his own farm is not clear. In particular, this relation may not remain constant between different states. For example, while the

---

[1] A similar tabulation of the AAA data is available for Illinois only.

average yield in Iowa tests was around 80 bushels per acre at a time when the average yield for the state was around 40 bushels, the North Carolina tests averaged more than 100 bushels, but at the same time the average state yield was only around 30 bushels. This implies that one cannot compare directly the differential of hybrids over open pollinated varieties from the Iowa tests with those from the North Carolina tests. Nevertheless, within each state these data may give reasonably good rankings of subareas according to the magnitude of the differential.

The AES data could have been very useful in providing an "objective" estimate of the superiority of hybrids if the level of practices and yields were appropriate for the farmer. The AAA data cover too narrow an area and too short a period to be useful except as a check. The AMS "Identicals" represent farmers' estimates of superiority of hybrid seed--a desirable feature --and restrict themselves to data collected from farms that grow both hybrid and open pollinated corn. This is an advantage in the sense that it eliminates a large source of bias due to differences in the level and quality of other inputs. However, as the acreage of open pollinated corn becomes small, farmers that still grow both become rather unrepresentative. Furthermore, the AMS mailing list may be biased to a different degree in different parts of the country. But the biggest problem of all, as we shall see below, is the lack of relationship between any of these three series.

While no comprehensive comparison of the various yield series has been attempted, several "spot checks" were made with very disquieting results. For 1939-41 there was no apparent relationship between the AAA and the AMS data in Illinois and Indiana, with the AMS data substantially below the AAA data. Nor did the scatter of the AMS estimates against the AES data yield any significant relationship for Iowa 1939-41, Missouri 1939-42, and North Carolina 1949-53. In all cases the AES estimates were substantially above the AMS and there was even a faint suggestion of a negative relationship.

These results were very discouraging. We have three sets of data which purport to measure the same thing, the superiority of hybrid over open pollinated varieties, and there is no relationship among them. While they in fact measure different things and one would not have expected a perfect relationship, nevertheless an undertone of agreement should have been present. The counsel of despair is to abandon all these measures.[1] However, I felt that there is still some use to be had from these data. In particular the AMS data for the Corn Belt are reasonably stable and follow a pattern suggested by outside knowledge. Also, while they may not be any good for indicating year to year changes in the superiority of hybrids, longer run averages seem to indicate very well the cross-sectional differences in the superiority of hybrids over open pollinated varieties.

As an alternative measure of the longer run cross-sectional differences in the superiority of hybrids, I used the average pre-hybrid yield of corn. This is based on the widespread belief that hybrids represent a constant percentage gain

---

[1] See Oscar Morgenstern, "The Accuracy of Economic Observations," (Princeton, 1949. Mimeographed).

of about 15 to 20 per cent over open pollinated varieties. Hence, differences in average pre-hybrid yields will imply differences in the absolute gain in yield due to the use of hybrids.

Usually an average for the ten years before an area reached 10 per cent in hybrids was used as the estimate of the average pre-hybrid yield. For states, the source was Agricultural Statistics. For crop reporting districts, various published and unpublished data from the AMS and from State Agricultural Statisticians were utilized. Sometimes fewer years were included in the average, depending on the length of the available series.

III. Other data.

No major problem arises in the use of other data beyond those already discussed in the text. The sources of other data are given in the text and in the Bibliography.

APPENDIX B

A "STATIC" ANALYSIS OF KANSAS DATA

An attempt was made to use the following model to explain changes in the observed percentage of total corn acreage planted with hybrid seed: Everybody for whom it is profitable to do so plants hybrid seed. We have, however, data only on average and not on individual profitability of the shift to hybrids. Given a smooth, bell-shaped distribution function, the proportion of farmers for whom the profitability of the shift to hybrids is zero or less will be a decreasing function of average profitability. As an approximation to expected profitability, last year's profitability and a lagged two year average profitability were used. "Average profitability" was defined as the average increase in yield of hybrids over open pollinated varieties times the price of corn and minus the change in the cost of seed (one plus the rate of interest times the seeding rate times the difference between the prices of hybrid and open pollinated seed per bushel). A logarithmic time trend was added to take into account lags in adjustment and other factors which change slowly and uniformly over time. This model was fitted separately to the data for each crop reporting district in Kansas (except district 7) and in addition an over-all regression was computed by pooling the data from all the crop reporting districts.

The results of these computations are presented in Table 10. They are uniformly dismal. While the coefficient of "profitability" has the expected sign and is sometimes significantly different from zero, time trend alone accounts for approximately 90 per cent of the variation, with "profitability" contributing only an additional 1 to 4 per cent. A futile attempt was made to improve upon these results by separating "profitability" into its various components--increase in yield, price of corn, price of hybrid seed--and by trying first differences. The results were even worse.

These results should not be interpreted to imply that the general hypothesis is false. The hypothesis that hybrid corn is used if it is profitable to do so is almost a tautology. Its usefulness depends on our ability to find empirical counterparts to "expected profitability." Obviously, last year's or last two years' profitability was not the right choice. In fact, the data suggest that given a substantial "objective" profitability, there is a more or less uniform process which changes "subjective" profitability over time. This serves as a point of departure for the analysis presented in Chapter IV.

TABLE 10

REGRESSIONS OF THE PERCENTAGE PLANTED WITH HYBRIDS ON AVERAGE PROFITABILITY AND TRENDS, KANSAS, 1940-1954

	Crop Reporting Districts								All Districts
	1	4	2	5	8	3	6	9	
$C=a+b\pi_{t-1}+g \log t$ N = 15									
b	.59	.68	3.87	.53	-.08	1.12	.46	.68	1.41
S_b	1.38	1.34	1.46	1.23	1.16	.65	.79	1.26	.71
g	79.9	63.7	75.1	67.0	67.1	69.8	75.8	82.2	70.5
S_g	14.0	14.1	11.6	10.8	11.8	6.1	5.8	11.6	6.6
R^2	.843	.80	.92	.82	.81	.94	.95	.88	.64
$r^2_{C \log t}$	.841	.79	.88	.82	.81	.92	.95	.88	.62
$C=a'+b'\pi+g' \log t$ N = 14									
b'	.78	1.98	5.54	1.09	.35	1.12	.61	.71	2.24
S_b	1.37	1.39	1.84	1.15	1.32	.93	1.10	1.16	.94
g'	101.9	72.4	76.1	83.8	83.5	70.1	80.3	103.1	82.1
S_g	14.7	16.1	16.0	10.6	14.2	9.5	8.0	10.9	9.5
R^2	.92	.90	.96	.90	.88	.89	.94	.94	.64
$r^2_{C \log t}$	.92	.88	.92	.89	.88	.88	.94	.94	.62

C_{it} -- The percentage of total corn acreage planted with hybrid seed in district i in year t.

π_{it} -- The profitability of the shift from open pollinated to hybrid seed in district in year t. $\pi = P_c D-(1+i)S_r(P_{hs}-P_{op})$, where P_c is the average price of corn received by Kansas farmers, D is the average yield of hybrid corn minus the average yield of open pollinated corn on farms reporting both, i is the rate of interest paid by North Central farmers, and P_{hs} and P_{op} are the prices paid for hybrid and open pollinated seed respectively.

$$\bar{\pi}_t = (\pi_t + \pi_{t-1})/2$$

$$t_{1940} = 1$$

Sources: <u>Agricultural Statistics</u>, and for C_{it} and D_{it} unpublished data from the Field Crops Statistics Branch, AMS.

APPENDIX C

ADDITIONAL REGRESSIONS

TABLE II

"ORIGIN" REGRESSIONS BY STATES--PART I

Form of the Regression	Dependent Variable	Coefficients of X'_1	X_3	X_{16}	X_{17}	X_{18}	R^2
Linear in the independent variables	Y_1	.43 (.14)	.20 (.09)	.44 (.10)	.34 (.07)	-.19 (.13)	.81
"	"	.47 (.15)	.20 (.10)	.45 (.10)	.41 (.06)		.79
"	"	.36 (.18)	.04 (.10)		.38 (.07)		.65
Logarithms of all the independent variables except X_{18}	"	6.6 (4.6)	16.9 (15.4)	13.7 (9.0)	43.8 (14.0)	-.26 (.17)	.67
"	"	8.6 (4.7)	-2.6 (6.0)		60.3 (12.5)		.58
Linear in the independent variables	Y_2	- .17 (.07)	- .09 (.04)	- .21 (.05)	- .19 (.02)		.78
"	"	- .11 (.08)	.04 (.05)		- .18 (.03)		.61
Logarithms of independent variables	"	-2.6 (2.1)	2.5 (2.7)		-27.8 (5.7)		.54

The figures in parentheses are the calculated standard errors.

Y_1 is the rank of a state when states are ranked according to the date at which they reached 10 per cent in hybrid corn. Earliest states have the highest numerical value.

Y_2 is the year (measured from 1940) a state reached 10 per cent. Linear interpolation rounded to one decimal point.

We expect the coefficients to be opposite in sign when Y_2 rather than Y_1 is used as the dependent variable.

X_1 is the "market density" unadjusted for differences in ceilings: average corn acreage 1937-1946 over land in farms in 1945.

X_3 is the average number of corn acres reported per farm, 1939.

X_{16} is the per cent of all farms reporting 5 acres of corn or less, 1939.

X_{17} is the per cent of all corn on "commercial" (Class 1 - 5) farms, 1950.

X_{18} is the rank of a state according to the total number of inbreds released by the state experiment station, cumulative, as of 1955.

Sources: Agricultural Statistics, Census of Agriculture, and, for X_{18}, Henderson, op. cit.

TABLE 11

"ORIGIN" REGRESSIONS BY STATES--PART II--Y_3 AS THE DEPENDENT VARIABLE

Form of Regression	Coefficients of X_1	X_3	b	X_4	X_{10}	X_{16}	X_1''	R^2
Linear	.12 (.07)	-.04 (.02)	- 2.1 (3.0)	-.99 (.12)				.84
"	- .07 (.06)				.83 (.12)			.69
"			- 5.5 (2.4)		.74 (.12)			.73
"		-.03 (.02)		-.94 (.09)				.82
"				-.72 (.13)	.33 (.13)			.84
"					.51 (.15)	.17 (.05)	-.07 (.07)	.78
"	.08 (.13)	-.03 (.05)	-14.0 (14.9)					.39
Linear, excluding the Southeast and Louisiana N = 23	- .20 (.09)	.01 (.03)	- 1.5 (3.8)					.52
Logarithms of independent variables	.78 (2.39)	-.42 (2.53)	-17.6 (5.7)					.37

Y_3 = Y in Table 3. X_1'' equals X_1 multiplied by the seeding rate in 1944. (Source:

Agricultural Statistics) times 10. X_{16} is the rank of a state according to the number of different hybrids recommended by the state experiment station in 1941. (Source: *Report of the Fourth and Fifth Corn Improvement Conferences of the North Central Region*. Reported by M.T. Jenkins, (Washington, February 18, 1942).

See Table 3 for definitions of all other variables.

TABLE 12

"ORIGIN" REGRESSIONS BY CROP REPORTING DISTRICTS--DEPENDENT VARIABLE--Y_3--N = 132

Form of Regression	X_1	X_3	b	X_4	X_{10}	R_2
Linear	- .015 (.017)	.003 (.006)	.032 (.625)	- .001 (.004)	1.10 (.06)	.90
"	- .010 (.012)			- .001 (.004)	1.10 (.05)	.90
"	- .07 (.02)			- .051 (.006)		.56
"	- .011 (.012)				1.11 (.04)	.90
"	- .03 (.03)		- 7.8 (1.1)			.49
"	- .08 (.04)	.02 (.01)	7.4 (1.2)			.51
Linear, excluding Mich., Wisc., Minn. Tenn., Ala. N=101	- .04 (.01)				.90 (.05)	.88
"	- .12 (.02)			- .029 (.004)		.73
Logarithms of independent variables	-2.32 (1.30)	1.13 (1.03)	-11.9 (1.8)			.48
"	-1.55 (1.09)		-11.1 (1.7)			.48

For definitions of the variables, see Table 3. Y_3 = Y.

TABLE 13

REGRESSIONS OF "ORIGIN" ON "MARKET DENSITY" WITHIN EACH TECHNICAL REGION: $Y_{3i} = a_i + b_i \log X_{1i}$

Region	N	b	r^2	$\bar{X}_1$
0	10	-4.4	(-).43	.02
1	16	-4.1	(-).25	.08
2	20	-5.1	(-).52	.11
3	26	-5.4	(-).71	.19
4	23	-4.9	(-).58	.24
5	23	-3.5	(-).38	.25
6	45	-3.7	(-).53	.16
7	22	-2.1	(-).20	.10
8a	37	-1.9	(-).27	.08
8b	16	2.4	(+).11	.07
9	29	1.8	(+).05	.08
10	23	.1	(+).001	.05

The regions are based on Funk Bros. Seed Company's "Maturity Groups." They are more or less latitudinal divisions of the U.S. numbered from North to South. That is, Region 0 is on the Canadian border, while Region 10 is on the Mexican Gulf.

TABLE 14

SOME ADDITIONAL "SLOPE" REGRESSIONS--DEPENDENT VARIABLE: b' or LOG b'

Regression	Coefficients of					R^2
	X_3	X_8	X_9	X_{11}	X_{13}	
By states N = 32						
Linear	.006 (.002)	.017 (.005)	-.001 (.002)			.658
Logarithmic	.293 (.093)	.624 (.205)	.037 (.163)			.668
"	-.109 (.179)	-.110 (.315)		.471 (.219)	.430 (.148)	.749
"	.239 (.081)	.512 (.134)			.183 (.100)	.704
"				.371 (.071)	.373 (.070)	.744
By crop reporting districts N = 132						
Logarithmic	.419 (.052)	.541 (.069)			.053 (.083)	.688

X_3 and X_8 are defined in Table 7. X_{11} is defined in Table 9. X_9 is the farm-operator family level-of-living index for 1940. Source: U.S.D.A., BAE, Farm-Operator Family Level-of-Living Indexes for Counties of the U.S., 1930, 1940, 1945, and 1950 (Washington, May 1952). X_{13} is the value of corn as per cent of all crops, 1939. Source: 1940 Census of Agriculture.

BIBLIOGRAPHY

I. Hybrid Corn

Books

Crabb, A.R. The Hybrid Corn Makers: Prophets of Plenty. Brunswick: Rutgers University Press, 1948.

Henderson, C.B. Inbred Lines of Corn Released to Private Growers from State and Federal Agencies and Double Crosses Recommended by States. 2nd revision. Champaign: Illinois Seed Producers' Association, April 15, 1956. (Processed.)

Sprague, G.F. (ed.) Corn and Corn Improvement. (Advances in Agronomy, Vol. III) New York: Academic Press, 1955.

Wallace, H.A., and Bressman, E.N. Corn and Corn Growing. 5th ed. New York: John Wiley and Sons, 1949.

Wallace, H.A., and Brown, W.L. Corn and Its Early Fathers. East Lansing: Michigan State University Press, 1956.

Public Documents

Cunningham, J.C. Maize Bibliography for the Years 1888 to 1945, Inclusive. ("Contributions from Iowa Corn Research Institute," Vol. III, Nos. 1-3.) Ames: Iowa Agricultural Experiment Station, 1941, 1948, and 1951.

Jugenheimer, R.W. Hybrid Corn in Kansas. Circular 196. Manhattan: Kansas Agricultural Experiment Station, February, 1939.

Neal, N.P., and Strommen, A.M. Wisconsin Corn Hybrids. Bulletin 476. Madison: Wisconsin Agricultural Experiment Station, February, 1948.

Rogers, J.S., and Collier, J.W. Corn Production in Texas. Bulletin 746. College Station: Texas Agricultural Experiment Station, February, 1952.

Ryan, B., and Gross, N. Acceptance and Diffusion of Hybrid Corn Seed in Two Iowa Communities. Research Bulletin 372. Ames: Iowa Agricultural Experiment Station, January, 1950.

U.S. Department of Agriculture. Technology on the Farm. Washington: 1940.

Wilkening, E.A. Acceptance of Improved Farm Practices in Three Coastal Plain Counties. Technical Bulletin 98. Raleigh: North Carolina Agricultural Experiment Station, May, 1952.

Reports and Proceedings

Annual Hybrid Corn Industry--Research Conference, Proceedings. Chicago: American Seed Trade Association, 1946-1956.

Report of Seed Certified by State Certifying Agencies. Washington: U.S. Federal Extension Service, 1946-1956.

Report of the Corn Improvement Conference of the North Central Region. Beltsville: U.S. Department of Agriculture, Division of Cereal Crops and Diseases, various years, 1937-1954. (Mimeographed.)

Report of the Northeastern Corn Improvement Conference. Beltsville: U.S. Department of Agriculture, Division of Cereal Crops and Diseases, various years, 1945-1956. (Mimeographed.)

Report of the Southern Corn Improvement Conference. Beltsville: U.S. Department of Agriculture, Division of Cereal Crops and Diseases, various years, 1938-1956. (Mimeographed.)

Articles

Apodaca, A. "Corn and Custom: Introduction of Hybrid Corn to Spanish American Farmers in New Mexico," in Human Problems in Technological Change. ed. Spicer, E.H. New York: Russell Sage Foundation, 1952, pp. 35-40.

Dowell, A.A., and Jesness, O.B. "Economic Aspects of Hybrid Corn," Journal of Farm Economics, XXI (May, 1939), 479-88.

Jenkins, M.T. "Corn Improvement," in Yearbook of Agriculture, 1936. Washington: U.S. Department of Agriculture, 1936, pp. 455-522.

________. "Corn Hybrids for the South," in Yearbook of Agriculture, 1943-47. Washington: U.S. Department of Agriculture, 1947, pp. 389-94.

Mangelsdorf, P.C. "The History of Hybrid Corn," Journal of Heredity, XXXIX (1948), 177-180.

Richey, F.D. "The Lay of the Corn Huckster," Journal of Heredity, XXXIX (1948), 10-17.

Ryan, B., and Gross, N.C. "The Diffusion of Hybrid Seed Corn in Two Iowa Communities," Rural Sociology, VIII (1943), 15-24.

Ryan, B. "A Study in Technological Diffusion," Rural Sociology, XIII (1948), 273-85.

Sprague, G.F. "The Experimental Basis for Hybrid Maize," Biological Reviews, XXI (1946), 101-20.

Wallace, H.A. "The Revolution in Corn Breeding," Prairie Farmer, March 21, 1925.

________. "Corn and the Midwestern Farmer," Proceedings of the American Philosophical Society, C (1956), 455-66.

Unpublished Material

Swartz, J.T. "A Study of Hybrid Corn Yields as Compared to Open Pollinated Varieties," Washington: Federal Crop Insurance Corporation, Insurance Section, April and May, 1942. (Handwritten.)

Wilkening, E.A. "The Acceptance of Certain Agricultural Programs and Practices in a Piedmont Community of North Carolina." Unpublished Ph.D. dissertation, University of Chicago, 1949.

II. Growth Curves and Estimation Problems

Books

Aitchison, J., and Brown, J.A.C. The Lognormal Distribution. Cambridge: Cambridge University Press, 1957.

Bush, R.R., and Mosteller, F. Stochastic Models for Learning. New York: John Wiley and Sons, 1955.

Davis, H.T. The Theory of Econometrics. Bloomington: Principia Press, 1941.

Finney, D.J. Probit Analysis. Cambridge: Cambridge University Press, 1947.

Kuznets, S. Secular Movements in Production and Prices. Boston: Houghton Mifflin, 1930.

Lotka, A. Elements of Physical Biology. Baltimore: Williams and Wilkins, 1925.

Pearl, R. Studies in Human Biology. Baltimore: Williams and Wilkins, 1925.

Tintner, G. Econometrics. New York: John Wiley and Sons, 1952.

Articles

Berkson, J. "A Statistically Precise and Relatively Simple Method of Estimating the Bioassay with Quantal Response, Based on the Logistic Function," Journal of the American Statistical Association, XLVIII (1953), 565-99.

________. "Maximum Likelihood and Minimum Chi-Square Estimates of the Logistic Function," Journal of the American Statistical Association, L (1955), 130-62.

________. "Estimation by Least Squares and by Maximum Likelihood," in Proceedings of the Third Berkeley Symposium on Mathematical Statistics, I (Berkeley: University of California Press, 1956), 1-11.

Dodd, S.C. "Testing Message Diffusion in Harmonic Logistic Curves," Psychometrika, XXI (1956), 191-205.

Hotteling, H. "Edgeworth's Taxation Paradox and the Nature of Demand and Supply Curves," Journal of Political Economy, XL (1932), 577-616.

Reed, L.J., and Pearl, R. "On the Summation of Logistic Curves," Journal of the Royal Statistical Society, XC (New Series, 1927), 729-46.

Winsor, C.P. "A Comparison of Certain Symmetrical Growth Curves," Journal of the Washington Academy of Sciences, XXII (1932), 73-84.

III. Miscellaneous

Books

Haavelmo, T. A Study in the Theory of Economic Evolution. Amsterdam: North-Holland Publishing Co., 1954.

Morgenstern, O. "The Accuracy of Economic Observations." Princeton, 1949. (Mimeographed.)

Wald, A. Sequential Analysis. New York: John Wiley and Sons, 1947.

Public Documents

Cochrane, W.W. An Analysis of Farm Price Behavior. Progress Report 50. State College: Penn State College, May, 1951.

How Farm People Accept New Ideas. Special Report No. 15. Ames: Iowa State College, Agricultural Extension Service, November, 1955.

Strand, E.G., and Heady, E.O. Productivity of Resources Used on Commercial Farms. Technical Bulletin 1128. Washington: U.S. Department of Agriculture, November, 1955.

U.S. Department of Agriculture. Agricultural Statistics. Washington: various years.

________. Bureau of Agricultural Economics. Agricultural Estimating and Reporting Services of the U.S. Department of Agriculture. Misc. Pub. 703. Washington: 1949.

________. Farm-Operator Family Level-of-Living Indexes for Counties of the U.S., 1930, 1945, and 1950. Washington: May, 1952.

U.S. Department of Commerce. Census of Agriculture. Washington: 1940 and 1950.

________. Statistical Abstract of the United States. Washington: various years.

Articles

Lee, I.M. "A Critical Evaluation of Available Agricultural Statistics," in Crop Estimating and Reporting Services of the Department of Agriculture. Report and Recommendations of a Special Subcommittee of the Committee on Agriculture of the House of Representatives, 82nd Congress, 2nd Session, 1952, pp. 54-61.

Nerlove, M. "Estimates of the Elasticities of Supply of Selected Agricultural Commodities," Journal of Farm Economics, XXXVIII (1956), 496-509.

Roos, C.F., and von Szelisky, V. "Factors Governing Changes in Domestic Automobile Demand," in Dynamics of Automobile Demand. New York: General Motors Corporation, 1939.

Warnz, W. "Measuring Spatial Association with Special Consideration of the Case of Market Orientation of Production," Journal of the American Statistical Association, LI (1956), 597-604.