FARM SIZE AND REAPER DIFFUSION
IN THE ANTEBELLUM MIDWEST[*]

By

David Levine

I. Introduction

The first known reference to the use of horse-power to reap grain is by Pliny the Elder in 70 A.D.[1] The modern history of the horse-powered reaper can be dated from the first sale of the Hussey reaper in 1833.[2] In the following six years total known reaper production amounts to a mere forty-five reapers.[3] By 1848 known reaper production in the U.S. amounted to 2,277 reapers while in 1849 annual production jumped to 1,730 reapers. Beginning in 1854 reaper production increased even more dramatically, climbing from 3,358 in 1854 to 6,627 in 1855, and then to about 12,000 by 1859.

This remarkable pattern of reaper production raises a variety of important issues. Chief among these is the question posed by P. A. David:[4]

> "[W]hy only [in the mid-1850s] were large numbers of farmers suddenly led to abandon an old, labor intensive method of cutting their grain, and switch to the use of a machine available since its invention two decades earlier?"

Comparative static partial equilibrium analysis suggests that the introduction of reapers should depend on factor prices. Prior to the 1830s small grains were reaped by hand; either by sickle or by cradle.[5] Because the rental cost of both sickles and cradles is very small, the relevant

cost is that of hiring the labor to do the reaping. Beginning in the 1830s the more capital intensive reaper became available. Depending on the relative prices of reapers and labor, reapers would or would not be used. Unfortunately this comparative static partial equilibrium analysis does not explain why the shift should take over twenty years.

To remedy this defect David[6] proposed a very specific model of how factor prices in conjunction with farm sizes will cause gradual reaper diffusion. Briefly, the model is this. Assume that the small grain acreage of farms is determined exogenously; independent of the availability of reapers. If reapers are purchased rather than rented, the cost of using a reaper is (ignoring issues of wear and tear to the reaper) independent of the number of days it is used each year. Labor, on the other hand, can either be hired until the grain has been harvested, or else has alternative uses of commensurate worth. Since reapers are indivisible if they have no alternative uses and cannot be rented out, the full yearly cost of the reaper must be borne regardless of how many acres are harvested. As a result the least cost method of harvesting depends on farm size as well as on factor prices. Small farms may find it cheaper to use a small amount of labor rather than an entire reaper, while larger farms may use reapers. There will be some threshold farm size at which the farmer does not care which technique is used.[7]

The gradual expansion of reaper use is now explained

by two factors. Changing relative factor prices lowered the
threshold size, while farm sizes increased exogeneously.
This resulted in a larger proportion of farms over the thres-
hold size and consequently more reapers were purchased. In
fact, Jones[8] showed that between 1849 and 1859 the number of
farms over the threshold size (in Illinois) increased by
930% while the (total U.S.) reaper stock increased by 1,060%.

Although attractive for its great simplicity and abil-
ity to explain the growth in the reaper stock, the threshold
hypothesis has peculiar assumptions -- the foundations of the
model appear to be that reapers can be used on only one farm
each harvest and that small grain acreage will not be increased
to take full advantage of reaper technology. David[9] ration-
alized this by hypothesizing that transactions costs (of
different types) prevented both the movement of reapers and
changes in farm size. However, Olmstead[10] has presented per-
suasive evidence that transactions costs did not prevent
widespread movements of reapers between farms, nor joint pur-
chases of reapers by several farms.

A moment's reflection will show that while Olmstead
decisively rejected one of the assumptions of the threshold
model, his evidence has no bearing on the model's predictions.
Olmstead's evidence is entirely consistent with the threshold
model.

Suppose that farms purchase reapers only if they are
over the threshold size.[12] The threshold size is only about

1/3-2/3rd the optimal small grain acreage for a single reap-
er.[13] As a result a farm exactly at the threshold size will
have completed reaping its grain with a reaper after three
to six days of a ten day harvest period.[14] Rather than simply
leaving his reaper in the barn for the remaining 4-7 days,
it is surely sensible for the farmer to pick up some extra
money by hiring it out to his neighbors at the rate they would
pay to reap by hand. In other words, even if only·farms above
the threshold size own reapers, there might be extensive
renting and sharing of reapers.

This does not explain why only farms over the threshold
size purchase reapers. There are, however, a variety of
reasons larger farms might purchase reapers when small farms
don't. Because of the cost of sharing a reaper or renting it
out to neighbors, it is more profitable to use a reaper on
the owner's farm than elsewhere. Since the large farmer can
use his reaper at the higher rate of return on his own farm,
while the small farmers must depend in part on the lower rate
of return from renting or sharing, the large farmers may find
it profitable to own a reaper when his smaller neighbor does
not. In addition, bigger operations may have better access
to capital markets and consequently prefer relatively more
capital intensive techniques. They may also be run by better
entrepreneurs and more accurately evaluate the profibatility
of purchasing a reaper.

This argument is a rationalization of why the threshold

hypothesis may hold. Regardless of how convincing it seems, what the argument establishes is that there may be plausible reasons for believing the threshold hypothesis, regardless of what renting and/or reaper sharing actually occurred. To test the hypothesis some other type of evidence is required.

The distinctive feature of the threshold model is that it not only predicts how many farms will have reapers, it also predicts which farms will have reapers. If reapers are scattered over farms of all sizes, and not concentrated in farms over the threshold size, we have no choice but to reject the threshold model. In this case the aggregate reaper stock implications of the threshold model must be dismissed as mere coincidence; or we may less graciously wish to observe that given the wide uncertainty about what the threshold size actually is, it is not too surprising that some values can be found that will be consistent with the growth of the reaper stock.[15] It is also possible that causality runs the other way; that the growth in farm sizes was due to the introduction of reapers.

Although the question of which farms have the reapers is crucial, evidence is hard to come by, and the only evidence presented to date are two quotations cited by David.[16] The first is from "western agricultural journals during 1846 and 1847" to the effect that no farmer "'who has not at least fifty acres of grain'" should buy a reaper, while the second is by "a reliable contemporary witness, Lord Robert Russel"

who claimed that in 1853 the cereals were all cut by reaper on the larger farms on the prairies. Since this cannot be considered conclusive (to say the least) stronger evidence is desirable.

This paper uses cross sectional data on farm sizes and reaper sales in 1850-1860 Illinois to examine the implications of the threshold model. In the next section a model of the joint reaper purchase/farm expansion decision is presented and aggregated to the county level. The following section discusses estimation of the aggregative model. The empirical results are then presented and examined, and the paper wraps up with a summary.

II.   A Profit Maximizing Model of Reaper Purchases and Farm
      Expansion

     In this section the decision of a farmer to purchase a
reaper and expand the size of his farm is examined. The
resulting demand functions are then aggregated to the county
level to permit estimation. Hypotheses concerning these de-
mand functions which are relevant to the threshold model are
then examined.

     A farm prior to the availability of reapers is charac-
terized by its size s and a vector λ that describe other as-
pects of the farm or farmer. These can include, though need
not be limited to, entreprenurial skill, risk aversion, access
to capital markets, perceptions of the profit potential in
using reapers, ability to maintain and repair machines and
local market characteristics such as the wage paid hired hands.
Without a reaper the farmer faces an opportunity cost of
$W(\lambda)$ per acre harvested. This represents the cheapest of the
following three alternatives: hired labor, the opportunity
cost of the farmer's own labor, or the rental cost of a reap-
er. Alternatively the farmer can purchase a number of reapers,
k, for an annual cost of $R(\lambda)$ each. Each reaper can cut $M(\lambda)$
acres in one harvest season of about ten·days. If the farmer
wants to rent out his reaper, he can charge the prevailing
local wage rate $W(\lambda)$,[18] but he must pay out of this the var-
ious transactions costs of negotiating with his neighbors and

moving the reaper from farm to farm. As a result he receives only a fraction $b(\lambda)$ of the rental he is payed.

In addition to purchasing a reaper in the next period, the farmer may also wish to expand his small grain acreage by an amount ΔS to avoid the transactions costs from renting out his reaper. This costs an amount $C(\Delta S, \lambda)$. Beyond the harvest and expansion costs, the farmer receives a margin of $p(\lambda)$ above other costs for each acre of output. The farmer's profits as a function of the decision variables k and ΔS are then:

(1) $\Pi(k, \Delta S) = p(S + \Delta S) - C(\Delta S) - kR - w(S + \Delta S - kM)$
$$\text{if } S + \Delta S \geq kM$$

$$p(S + \Delta S) - C(\Delta S) - kR + bw(kM - S - \Delta S)$$
$$\text{if } S + \Delta S < kM$$

where the first expression represents profits if the farmer purchases fewer reapers than required, and the second more than required. The argument λ is suppressed for notational simplicity.

Maximization of the profit function (1) yields demand equations for reapers and farm expansion:[19]

(2) $$k = k(\lambda, S)$$

$$\Delta S = \Delta S(\lambda, S)$$

Since the form of these demand equations are not restricted in any way, profit maximization is not a necessary assumption for their derivation. Any theory that yields demand as a

function of the exogenous variables describing a farm will suffice, and the reader is welcome to replace the profit maximizing derivation above with her own pet theory.

To aggregate the demand equations to the county level requires the joint distribution of characteristics λ and farm sizes S. Denote by $g(S|Z_S)$ the number of farms of size S in a county with size distribution parameters Z_S. Let $f(\lambda|S, Z_\lambda)$ be the fraction of farms with size S that have characteristics λ in a county with parameters Z_λ. The county demand functions are given by:

$$(3) \quad K = \int k(\lambda,S)f(\lambda|S,Z_\lambda)g(S|Z_S)d\lambda dS = K(Z_\lambda,Z_S)$$

$$\Delta S = \int \Delta S(\lambda,S)f(\lambda|S,Z_\lambda)g(S|Z_S)d\lambda dS = \Delta S(Z_\lambda,Z_S)$$

This shows that county reaper purchases and total small grain acreage expansion are both functions of the parameters determining the size distribution of farms and the conditional density of farm characteristics in a county.

The object of deriving county demand functions is to test the threshold model of reaper diffusion. To do so it is convenient to use a different but equivalent pair of reduced form demand equations. The particular implications of the threshold model is that the number of reapers per acre should increase with the proportion of farms above the threshold level. As a result it is convenient to divide the first equation of (3) by total county small grain acreage $S + \Delta S$ to get:

(4)
$$\overline{K} = \frac{K}{S + \Delta S} = \frac{K(Z_\lambda, Z_s)}{S + \Delta S(Z_\lambda, Z_s)} = \overline{K}(Z_\lambda, Z_s)$$

where S is omitted from the final expression since it is implicit in the parameters Z_s. In the second equation we divide by S to get:

$$(5) \qquad \overline{\Delta S} = \frac{\Delta S}{S} = \frac{\Delta S(Z_\lambda, Z_s)}{S} = \overline{\Delta S}(Z_\lambda, Z_s)$$

The threshold model suggests three hypotheses concerning the reduced form demand equations (4) and (5). First, it implies that if increases in components of Z_s are measured such that the fraction of farms above the threshold size is increasing in Z_s, then $\partial \overline{\Delta K}/\partial Z_s > 0$. Second, it presumes that growth in farm sizes ΔS is exogenous to reaper purchases. We cannot estimate what the growth of farm sizes would have been in the absence of reaper availability. However, if the growth of farm sizes is influenced by reaper availability we would expect that variables affecting reaper purchases would also affect growth in farm sizes. In other words we would not expect that for all variables Z in $Z_\lambda \frac{\partial \overline{K}}{\partial Z} \frac{\partial \overline{\Delta S}}{\partial Z} = 0$, since this would mean there are no joint influences. This is a weak test however since there could easily be joint influences even if the growth of farm sizes was not affected by reaper availability. An even stronger version of the threshold model would suggest that the only difference between counties that would explain differential reaper purchases is differing farm sizes. This suggests the hypothesis that $\frac{\partial \overline{K}}{\partial Z} = 0$ for all parameters Z and Z_λ.

In the following section we discuss how to estimate the equations (4) and (5) simultaneously. Based on the statistical model developed we show how to test the three hypotheses of relevance to the threshold model.

III. The Statistical Model

The objective of this section is to describe how to estimate a cross-sectional aggregate model of reaper purchase and growth of farm size. The cross-sectional data are county aggregates from the state of Illinois for the period 1849-1859. Data from 1849 can be considered exogenous to the purchase/growth decision, since very few reapers were available prior to that date. Data on final farm size is computed from the 1860 census, and refers to the 1859 harvest, well after the widespread introduction of reapers in Illinois (McCormick moved his center of operations to Chicago in 1853). The reaper stock is computed from data on McCormick reaper sales in 1849-57, based on a 20% rate of reaper depreciation. This is the figure put forth by Olmstead.[21] The model was also run using a 10% rate of depreciation, as suggested by David.[22] This did not affect the results presented here.[23]

The distribution of farm sizes within a county was assumed lognormal, and the moments estimated by county as described in Appendix A. Farm size should be taken as a synonym for small grain acreage, which is the variable of interest. The parameters Z_λ posed a greater difficulty. Three variables were included as likely to influence the distribution of farm cost function characteristics. Two dummy variables were created to indicate the availability of transportation. The variable D, takes the value one if a

county has a river or other body of water in or adjacent to
it, and zero otherwise. This is likely to have an important
influence both on the cost of shipping grain and of shipping
reapers, especially before the widespread completion of rail-
way lines in 1855. The variable D_2 takes the value one if
a railroad existed in the county in 1850, and zero otherwise.
This reflects not only transportation costs, but the degree
of development of the county. As the extent to which a county
was settled and developed in 1859 can be expected to have
influenced initial cost parameters, population density in
1849 was included as a third measurable parameter in Z_λ. To
allow for the possibility of joint unobserved influences
affecting both equations, it was assumed that the covariance
between the equations was not necessarily zero. The variables
used in this study are reported in Table 1 and described in
greater detail in Appendix A.

The variables were entered linearly into the two equa-
tions as a simple approximation, except for the farm size
parameters in equation one, where the threshold hypothesis
suggests an alternative relationship. Let Φ denote the cum-
ulative normal distribution and ϕ its density, and suppose
farm sizes in a county are lognormally distributed with

parameters μ and σ. The threshold model predicts that all farms above a threshold size T will buy approximately S/M reapers where M is the number of acres harvested by a reaper in a season, and S is the size of the farm. If n is the total number of farms in the county we can compute the number of reapers per acre by:

$$(6) \quad \frac{\int_T^\infty \frac{S}{M} \frac{n}{\sigma S} \phi \left(\frac{\log S - \mu}{\sigma}\right) dS}{\int_0^\infty S \frac{n}{\sigma S} \phi \left(\frac{\log S - \mu}{\sigma}\right) dS} = \frac{1}{M} \phi \left(\frac{\mu + \sigma^2 - \log T}{\sigma}\right)$$

Consequently the basic model estimated is given by equations (7) and (8) where ϵ_r and ϵ_d are stochastic errors and the remainder of variables are described in Table 1:

$$(7) \quad r_{20} = \mu_r + \gamma_{1r} D_1 + \gamma_{2r} D_2 + \gamma_r D + \beta_{1r} \log \phi \left(\frac{z - \beta_{2r}}{\sigma}\right) + \epsilon_r$$

$$(8) \quad d = \mu_d + \gamma_{1d} D_1 + \gamma_{2d} D_2 + \gamma_d D + \beta_{1d} \mu + \beta_{2d} \log \sigma + \epsilon_d$$

The correlation coefficient between the two errors ϵ_r and ϵ_d is assumed equal to ρ an unknown parameter.

Because of the cross-sectional nature of the model, heteroskedasticity is likely to be present in the residuals. However the scale free form of the left-hand size variables makes it unclear what form this will take. To allow for a general heteroskedastic process it is assumed that the variance of the errors is related to county size (total crop acreage) H_2 by equation (9).

Table 1

Summary of Variables

r_{10} — log of reaper stock in 1857 with 10% depreciation per small grain acre in 1859

r_{20} — log of reaper stock in 1857 with 20% depreciation per small grain acre in 1859

d — percentage growth of average farm size

D — logarithm of population density per acre

D_1 — presence or absence of navigable body of water

D_2 — presence or absence of railroad in 1850

H_1 — number of farms in 1859

H_2 — total farm acreage (all types) in 1849

μ — estimated mean of logarithm of small grain acreage

σ — estimated standard error of logarithm of small grain acreage

$Z \quad = \mu + \sigma^2$

$$(9) \qquad V(\varepsilon_i) = \exp(\delta_i + \eta_i \log H_2) \qquad i = r,d$$

A variant in which number of farms replaced total acreage was also tried without any particular effect on the results.

In order to estimate the system by maximum likelihood, the errors were assumed normally distributed. To test for robustness an alternative exponential distribution using the absolute errors to the 1.1 power in place of the absolute errors squared was used. The results did not change substantially, although, as might be expected when outliters are omitted, the standard errors increased. The following section reports results from the basic model, while the results from the variants mentioned are in Appendix B.

Two hypotheses were tested in addition to the quasi-t tests for each coefficient. The first of these is for the absence of joint effects other than farm sizes between the equations. The null hypothesis is

$$H_o \quad \gamma_{1r}\gamma_{1d} = 0 \qquad \gamma_{2r}\gamma_{2d} = 0 \qquad \gamma_r\gamma_d = 0 \qquad \rho = 0.$$

This was tested by a Wald test based on the estimated variance-covariance matrix of coefficients.

We turn now to the empirical results.

IV. Empirical Results

The cross-sectional data on Illinois provides strong support for the hypothesis that farm size was an important consideration in reaper purchases. However the data suggest that the relationship is due to very small farms not purchasing reapers, rather than only very large farms purchasing reapers.

The major results are reported in Table 2. In the first equation we see that the presence of navigable water raises the number of reapers per acre by 73%, while counties with a railroad in 1850 had 16% more reapers. As we would expect lower transportation costs are associated with more reapers. Somewhat surprisingly, more heavily populated counties had fewer reapers -- doubling the population density causes the reapers/acre to fall by 36%. Overall, however, these parameters and the correlation coefficient are not individually significant. This lends support to the threshold hypothesis as it suggests that only farm size matters.

Turning to the farm size parameters, we see that β_{1r} is 1.9, while the theory predicts it should be equal to one. However, the standard error is such that we can hardly reject the hypothesis that it is equal to one. The estimated

Table 2

Empirical Estimates of Threshold Model

$$r_{20} = -7.6 + 0.55D_1 + 0.15D_2 - 0.36D + 1.9 \log \Phi \left(\frac{Z - 1.9}{\delta}\right) + \varepsilon_r$$

se	(2.9)	(0.64)	(0.62)	(0.79)	(1.2)	(0.70)
tst	(2.7)	(0.86)	(0.23)	(0.45)	(1.5)	(2.8)

$$d = 0.1D + 0.25D_1 + 0.24D_2 - 0.12D - 0.44\mu - 1.1 \log \sigma + \varepsilon_d$$

se	(0.51)	(0.14)	(0.80)	(0.11)	(0.068)	(0.64)
tst	(0.20)	(1.8)	(0.31)	(1.1)	(6.4)	(1.7)

$$V(\varepsilon_r) = \exp(17.0 - 1.4 \log H_2) \qquad V(\varepsilon_d) = \exp(-1.5 - 0.020 \log H_2)$$

se	(3.6)	(0.35)		(3.0)	(0.28)
tst	(4.6)	(4.1)		(0.49)	(0.072)

$$\rho(\varepsilon_r, \varepsilon_d) = 0.17$$

se	(0.13)
tst	(1.3)

$\lambda = -84.44$ (log-likeihood value excluding inessential constant)

$d = 7.8 \times 10^{-5}$ (convergence criterion for Berndt-Hall-Hall-Hausman)

$W(H_o, 4) = 2.4$ (Wald test for H_o)

threshold size (the exponential of β_{2r}) is very low -- only 6.7 acres (the smallest estimate made on theoretical grounds is 35.1 acres).[24] However this is based on 1849 sizes, by 1859 the average county grew 29%. This would mean that a 6.7 acre farm in 1849 would be 8.6 acres in 1859. The standard error associated with the 8.4 figure is 0.40 acres, so this is a relatively precise estimate. It strongly suggests that the absence of small farms rather than the presence of large farms is the crucial determinant in reaper purchases.

Opponents of the threshold hypothesis may claim that since β_{1r} is not significantly different than zero, farm size is not in fact important. Because of the non-linear nature of the model this does not necessarily follow. As a simple test of whether farm size matters at all, the threshold formulation was replaced with a version in which the lognormal moments enter linearly. Since the t-statistic on the mean farm size is 3.5, this indicates that the hypothesis that farm size doesn't matter can be rejected quite handily. Results from this version of the model can be found in Appendix B -- no other results are changed.

The second equation is of interest mainly insofar as joint effects in the two equations are important. Since the 10% level for a x^2 test with 4 degrees of freedom is 7.8, the Wald test value of 2.4 clearly shows the absence of significant joint effects. The estimated effects do have the expected sign however: variables that tend to raise rates of farm

growth, tend to raise reaper intensity as well.

The farm size parameters in the second equation are somewhat puzzling. The mean is decisively significant, which strongly contradicts Gibrat's law, so well beloved by industrial organization economists. I would speculate that the negative signs are due to small underdeveloped counties rapidly growing towards an optimal farm size. The effect is quite strong: halving the geometric mean farm size causes the rate of growth to jump by 44 percentage points.

V. Summary

The size of farms seems to be an important determinant of the intensity of reaper usage -- larger farms tend to use more reapers per acre. There is also evidence that the relationship takes the specific form that farms over a certain threshold size buy reapers while those below it don't. Despite this the evidence suggests that David's threshold model is wrong -- it is the extreme unprofitability of reapers on small farms, rather than the inability of medium sized farms to profitably rent and share reapers that is important.

Overall, while the model does not prove that exogenous changes in farm size led to gradual reaper diffusion, the absence of strong effects jointly moving reaper intensity and farm growth provides the first empirical evidence for this point of view.

APPENDIX A -- THE DATA

The raw data is from the following sources:

RS49 - McCormick reaper sales in 1849 from Hutchinson insert pp. 468-469

RS50   "   1850   "

RS54   "   1854   "

RS55   "   1855   "

RS56   "   1856   "

RS57   "   1857   "

AL59 - Acres of improved land in 1859 from 1860 Census of Agriculture

N3 - Number of farms in 1859 of 3-10 acres   "

N10   "   10-20   "

N20   "   20-50   "

N50   "   50-100   "

N100   "   100-500   "

N500   "   500-1000   "

N1000   "   1000+   "

CW59 - Bushels of wheat in   1859   "

CR59   "   rye   "   "

CO59   "   oats   "   "

N49 - Number of farms in   1848   from 1850 Census

AL49 - Acres of improved land   "   "

CW49 - Bushels of wheat   "   "

CR049 - Bushels of rye and oats   "   "

PP - Population   "   ".

TA - Acres of land in county   from 1945 Census of Agriculture

Small grain acreage was estimated using Parker and Klein's crop yield estimates as:

$$AS49 = \text{acres of small grains in } 1849 = \frac{CW49}{16} + \frac{CRO49}{40}$$

$$AS59 = \text{"} \quad \text{in } 1859 = \frac{CW59}{16} + \frac{CR59 + CO59}{40}$$

The standard error of the logarithm of small grain acreage was estimated using the following assumptions:

(1) Small grain acreage is proportional to acres of improved land.

(2) The standard error did not change between 1849 and 1859. Under these assumptions the standard error was estimated by maximum likelihood in a truncated lognormal model:

$$k_1 = 3 \quad k_2 = 10 \quad k_3 = 20 \quad k_4 = 50 \quad k_5 = 100 \quad k_6 = 500 \quad k_7 = 1000$$

Nk_i number of farms in size category k_i.

log-likelihood function =

$$\sum_{i=1,6} Nk_i \log \frac{\Phi(\frac{\log k_{i+1} - \lambda_1}{\lambda_2}) - \Phi(\frac{\log k_i - \lambda_1}{\lambda_2})}{1 - \Phi(\frac{\log 3 - \lambda_1}{\lambda_2})}$$

$$+ N_{1000} \log \frac{1 - \Phi(\frac{\log 1000 - \lambda_1}{\lambda_2})}{1 - \Phi(\frac{\log 3 - \lambda_1}{\lambda_2})}$$

where Φ is the cumulative normal.

This was estimated for each county using the Berndt, Hall, Hall and Hausman likehood maximiation algorithm, programmed for use on the MIT Multics system by Jerry Hausman, with modifications by Hank Farber, David Levine and Bill Dickens.

The final data used in estimation was:

$$r_{10} = \log \left(\sum_{t=49,50,54-57} RS_t (.9)^{57-t} / AS59 \right)$$

$$r_{20} = \log \left(\sum_{t=49,50,54-57} RS_t (.8)^{57-t} / AS59 \right)$$

$$d = \log \left(AS59 / \sum_{t=1,7} N_{k_i} \right) - \log \left(AS59 / N49 \right)$$

$$D = \log (PP/TA)$$

$$D_1 = \begin{array}{l} 1 \text{ river in or adjacent to county} \\ 0 \text{ otherwise} \end{array}$$

$$D_2 = \begin{array}{l} 1 \text{ railroad in county in 1850} \\ 0 \text{ otherwise} \end{array}$$

from map in Hutchinson insert between pp. 468-469

$$H_1 = N49$$

$$H_2 = AL49$$

$$\mu = \log (AS49/N49) - (1/2)(\hat{\lambda}_2)^2$$

$$\sigma = \hat{\lambda}_2$$

$$z = \mu + \sigma^2$$

Three new counties were formed between 1849 and 1859: Kankakee County from Iroquois County, Douglas County from Coles County and Ford County from previously unincorporated land. As a result Ford County was omitted and 1859 observations on

Kankakee County were combined with Iroquois County and from
Douglas County with Coles County.   Information on the forma-
tion of these counties was provided by Diane Wilhelm, refer-
ence liabrarian of the Illinois State Historical Library,
in a personal communication to the author.

Means and standard errors of the variables are:

	Mean	Standard Error
r_{10}	-8.0	3.5
r_{20}	-8.2	3.5
d	0.29	0.61
D	-3.8	0.58
D_1	0.40	0.49
D_2	0.061	0.24
$\log H_1$	6.5	0.60
$\log H_2$	10.6	0.75
μ	1.7	0.89
σ	0.74	0.087
z	2.2	0.90

APPENDIX B - VARIANTS OF THE STATISTICAL MODEL

All models were estimated by maximum likelihood using the Berndt, Hall, Hall and Hausman algorithm, programmed for the MIT Multics system by Jerry Hausman, with the assistance of Hank Farber, David Levine and Bill Dickens. The convergence criterion was either inability to significantly increase the likelihood function over five steps, or else $d'Vd \leq 10^{-4}$ where d is the gradient of the likelihood function, and V is the estimated variance-covariance matrix.

Linear Version

$$x_{20} = -10.0 + 0.64D_1 + 0.20D_2 - 0.44D + 1.6\mu + 2.5 \log \sigma + \epsilon_r$$

se   (4.6)   (0.70)   (0.80)   (1.0)   (0.47) (1.9)

tst   (2.5)   (0.92)   (0.25)   (0.43)   (3.4)   (1.3)

$$d = 0.12 + 0.26D_1 + 0.25D_2 - 0.12D - 0.45\mu - 1.1 \log \sigma + \epsilon_d$$

se   (0.54) (0.15)   (0.80)   (0.12)   (0.071) (0.67)

tst (0.22) (1.7)   (0.31)   (1.1)   (6.3)   (1.7)

$$V(\epsilon_r) = \exp(16.0 - 1.3 \log H_2) \qquad V(\epsilon_d) = \exp(-1.3 - 0.036 \log H_2)$$

se   (4.3)   (0.42)   (2.9)   (0.26)

tst   (3.6)   (3.2)   (0.45)   (0.13)

$$\rho(\epsilon_r, \epsilon_d) = 0.16$$

se   (0.13)

tst   (1.2)

$$\lambda = -94.97$$

$$d = 5.2 \times 10^{-5}$$

$$W(H_o, 4) = 2.6$$

Ten Percent Depreciation Version

$$r_{10} = -7.1 + 0.60D_1 + 0.19D_2 - 0.28D + 2.0 \log \phi(\frac{Z - 1.0}{\sigma}) + \epsilon_r$$

se   (2.8)   (0.62)   (0.59)   (0.77)   (1.3)   (0.66)

tst   (2.5)   (0.96)   (0.33)   (0.36)   (1.6)   (2.9)

$$d = 0.10 + 0.25D_1 + 0.24D_2 - 0.12D - 0.44\mu - 1.1 \log \sigma + \epsilon_d$$

se   (0.51)   (0.14)   (0.83)   (0.11)   (0.068)   (0.64)

tst   (0.20)   (1.8)   (0.30)   (1.1)   (6.4)   (1.7)

$$V(\epsilon_r) = \exp(18.0 - 1.5 \log H_2) \qquad V(\epsilon_d) = \exp(-1.5 - 0.21 \log H_2)$$

se   (3.5)   (0.34)   (3.0)   (0.28)

tst   (5.1)   (4.6)   (0.49)   (0.074)

$$\rho(\epsilon_r, \epsilon_d) = 0.17$$

se   (0.13)

tst   (1.3)

$$\lambda = -83.79$$

$$d = 5.1 \times 10^{-5}$$

$$W(H_o, 4) = 2.7$$

Version with Alternative Measure of County Size

$$r_{20} = -7.5 + 0.40D_1 + 0.033D_2 - 0.40D + 1.2 \log \Phi(\frac{z-2.5}{\tilde{\sigma}}) + \varepsilon_r$$

se   (3.6)   (0.77)   (1.2)   (0.10)   (0.63)   (0.79)

tst   (2.1)   (0.52)   (0.028)   (0.40)   (1.9)   (3.2)

$$d = 0.10 + 0.25D_1 + 0.24D_2 - 0.12D - 0.44\tilde{\mu} - 1.1 \log \tilde{\sigma} + \varepsilon_d$$

se (0.51) (0.14)   (0.79)   (0.11)   (0.069)   (0.63)

tst (0.20) (1.8)   (0.30)   (1.1)   (6.3)   (1.8)

$$V(\varepsilon_r) = \exp(9.0 - 1.2 \log H_1) \quad V(\varepsilon_d) = \exp(-1.4 - 0.040 \log H_1)$$

se   (2.2) (0.34)   (2.2)   (0.33)

tst   (4.0) (3.3)   (0.65) (0.12)

$$\rho(\varepsilon_r, \varepsilon_d) = 0.15$$

se   (0.14)

tst   (1.1)

$$\lambda = -92.95$$

$$d = 3.0 \times 10^{-3}$$

$$W(H_o, 4) = 1.6$$

Non-normal Version

$$r_{20} = -9.6 + 0.90D_1 + 0.31D_2 - 0.84D + 3.6 \log \Phi(\frac{Z-1.3}{\delta}) + \epsilon_r$$

se	(5.0)	(0.92)	(0.91)	(1.4)	(18.0)	(3.4)
tst	(1.9)	)0.97)	(0.34)	(0.59)	(0.20)	(0.38)

$$d = -0.039 + 0.22D_1 + 0.22D_2 - 0.17D - 0.41\mu - 1.1 \log \sigma + \epsilon_d$$

se	(0.44)	(0.098)	(0.25)	(0.096)	(0.055)	(0.38)
tst	(0.089)	(2.2)	(0.89)	(1.7)	(7.5)	(2.9)

$$V(\epsilon_r) = \exp(16.0 - 1.5 \log H_2) \quad V(\epsilon_d) = \exp(-0.25 - 0.29 \log H_2)$$

se	(4.7)	(0.46)	(3.5)	(0.32)
tst	(3.4)	(3.4)	(0.073)	(0.90)

$$\rho(\epsilon_r, \epsilon_d) = 0.24$$

se	(0.089)
tst	(2.7)

$$\lambda = -8.53$$

$$d = 8.6$$

$$W(H_o, 4) = 65.0$$

NOTES

[1]Hutchinson, p. 55.

[2]Rogin, p. 73.

[3]Rogin, p. 73.

[4]David, p. 201.

[5]Rogin, pp. 69-72.

[6]David, pp. 210-211.

[7]David, pp. 220-232.

[8]Jones, p. 454.

[9]David, p. 208.

[10]Olmstead, pp. 334-344.

[11]No footnote.

[12]We ignore the issue of joint ownership, the implications of which are unclear. Although Olmstead on p. 337 states that cases where one owner was identified as a financial backer were counted as single ownership, we have no way of knowing in how many cases one of the owners was a financial backer, but not identified as such in the McCormick ledgers from which Olmstead's data comes. In addition David in a private communication indicates much joint ownership was by farms owned by members of the same family, which suggests that it might not be that important.

[13]Threshold estimates for 1854-1857 range from 35.1 to 67.6 acres (Jones, p. 451), while the optimal small grain acreage implicit in David's and Olmstead's calculations is 110 acres, based on 11 acres per reaper day (David, p. 223) and a 10 day harvest period (Rogin, p. 130).

[14]Rogin, p. 130.

[15]In fairness to Jones, we note that he used Olmstead's threshold estimates.

[16]David, p. 215.

[17]However, more data may be available in McCormick ledgers.

[18]Because farmers can potentially hire out to their neighbors this must be the same for all farms in a sufficiently small locale.

[19]The demand correspondences are assumed to be single valued.

[20]No footnote.

[21]Olmstead, p. 331.

[22]David, p. 225.

[23]It may be objected that reapers other than McCormick reapers may have been sold in Illinois. However this cannot be too serious (although it undoubtably biases the constant term). Jones gives total U.S. reaper production in 1856 as 9,995 (p. 452), while Hutchinson gives McCormick reapers production then as 4,095 (p. 369). However McCormick sold mainly in Illinois, while many other producers sold primarily or solely on the East Coast, especially New England and the Atlantic States. As a result McCormick must have sold well more than half the reapers in Illinois.

[24]Jones, p. 451.

REFERENCES

Berndt, Hall, Hall and Hausman, "Estimation and Inference in Non-Linear Structural Models," Annals of Economic and Social Measurement, 3/4, 1974.

David, Paul A., "The Mechanization of Reaping in the Ante-Bellum Midwest" in Technical Choice, Innovation and Economic Growth (London: Cambridge University Press 1975), pp. 195-232.

Hutchinson, William T., Cyrus Hall McCormick, Vol. 1, Seed-Time 1809-1856, (New York: The Century Company, 1930).

Jones, Lewis R., "The Mechanization of Reaping and Mowing in American Agriculture: Comment," Journal of Economic History 37 (June 1977), pp. 451-455.

Kennedy, Joseph C.G., Agriculture of the United States in 1860 Compiled from the Original Returns of the Eighth Census (Washington, D.C.: Government Printing Office, 1864).

Olmstead, Alan L., "The Mechanization of Reaping and Mowing in American Agriculture," Journal of Economic History 35 (June 1975), pp. 327-352.

Parker, William N., and Judith L.V. Klein, "Productivity Growth in Grain in the United States, 1840-60 and 1900-10," in Output, Employment and Productivity in the United States after 1800 (New York: Columbia University Press for N.B.E.R., 1966.

U.S. Department of Commerce, Bureau of the Census, U.S. Census of Agriculture, 1945, Volume 1, Part 6, pp. 18-38.