

PRELIMINARY REPORT ON SUPPLY AND DEMAND FUNCTIONS

ESTIMATED BY EPEA/CFP

Gordon W. Smith

March, 1967 -

I. INTRODUCTION

This is a preliminary report on attempts made by EPEA/CFP to estimate supply and demand functions for the principal products supported by the minimum price program. It has two purposes. The first, to divulge the supply estimates made thus far. The demand functions are still unsatisfactory and will be presented at a later date. Just as important, given the fact that the particular supply model (the Nerlove adjustment model) has already been applied by other researchers in Brazil,* this report provides a summary critique of:

- (1) the estimation procedures in the presence of high multicollinearity between independent variables,
- (2) the probable bias of the structural coefficients in the model and
- (3) the basic data which must be used for estimates outside of São Paulo State.

As has been shown elsewhere, a forward equilibrium pricing system would be a considerable improvement over previous minimum price policy.** Accurate supply and demand functions would be extremely helpful in determining minimum prices and their probable effects within this policy framework. They would permit:

- (1) Forecasts of internal equilibrium prices at the farm level, excluding the influence of foreign trade.

* See the works of Sérgio Brandt et al on supply functions in São Paulo in mimeograph (Divisão de Economia Rural do Estado de São Paulo) and in several issues of Agricultura em São Paulo and Del-fim Netto et al, Agricultura e Desenvolvimento Econômico no Brasil, Estudos ANPES Nº.5, 1966.

** See EPEA, The Minimum Price Program, (mimeo), 1966.

- (2) Forecasts of potential exports and imports at prices differing from internal equilibrium levels.
- (3) Forecasts of purchases required by CFP, given minimum prices and harvest forecasts.
- (4) Knowledge of the impact of prices on the production of the minimum price products.

With these advantages in mind, attempts were made by EPEA/CFP to estimate supply and demand functions for five minimum prices products: cotton, peanuts, rice, corn and mandioca.* Because of the quite different consumption and production characteristics in the North-Northeast and the Center-South, supply and demand functions should be estimated regionally. Moreover, the minimum price program acts much more vigorously and directly with the farmer in the Center-South than in the North-Northeast. However, the lack of interregional trade statistics, regional disposable income data and reliable regional population estimates restricted the demand estimates to Brazil as a whole, while supply functions were estimated for the two regions separately.

This makes determination of equilibrium prices (our original goal), either regionally or nationally, impossible since supply and demand functions apply to different market areas. It is hoped that future work may overcome these obstacles, perhaps by using a time trend in regional demand functions as a proxy for the combined influence of population and income effects.

The report proceeds as follows. First, the supply model is presented. Then the statistical results are analyzed, followed by a critique of the results, the particular model used (the Nerlove adjustment model) and the accuracy of the data employed.

* Beans was not attempted because of the extremely poor quality of the available data. Here were estimates made for soybeans, given the crop's rapid evolution in Rio Grande do Sul, in good part due to technological factors for which we have no quantitative information (e.g. rotation with wheat).

From the results presented below, several supply functions appear "satisfactory" judging by their R^2 and the signs and significance of their structural coefficients. Farmers do appear to respond to price stimuli, even for rice and mandioca in the North-Northeast region. The greatest price elasticities, as might be expected, were found for peanuts in São Paulo and cotton in the Center-South, two completely commercial crops. Significant, but considerably lower, price response was also found for rice and mandioca in both the Center-South and the North-Northeast. Previous work with supply functions in Brazil has obtained similarly satisfactory results.* This all indicates that government price policy must take into account the impact of its support prices upon production.

It will be argued, however, that the quantitative usefulness for the minimum price program of these and similar supply functions estimated with the Nerlove adjustment model using SEP data is very limited. First, it will be shown that the coefficient of adjustment and long-run supply estimates derived from the model are probably biased, perhaps considerably so in some cases. And finally we will see that the basic underlying data are of very dubious quality. Further and more refined econometric work with the SEP data is not warranted. If supply analysis is to improve and the minimum price program to be put on a more solid empirical basis, a complete revision of data gathering methods is required.

II. ESTIMATES OF SUPPLY FUNCTIONS

A. The Model

In the present analysis one of the distributed lag supply models proposed by Marc Nerlove was used, the so-called adjustment model.** This model, in common with traditional supply analysis,

* See the works of Sérgio Brandão et al and Delfim Netto et al, cited above.

** The basic work is, Nerlove, Marc, The Dynamics of Supply: An Estimation of Farmers Response to Price, Johns Hopkins, 1958.

assumes that long-run equilibrium output in time t is a function of farmers' expected price in time t . That is (linear or linear in logs):

$$(1) \quad Q_t^{1*} = a + bp_t^{1*} + cp_t^{2*} + dt + u_t$$

were

Q_t^{1*} = equilibrium output in time t of product 1.

p_t^{1*} = expected price of product 1 in time t .

p_t^{2*} = expected price of a substitute product in time t .

t = a time trend to capture shifts in the supply function not included explicitly in the model.

U_t = a disturbance term following a normal distribution

The model further assumes, as does traditional analysis, that farmers expected price in t equals the market price in the previous year, $t-1$:

$$(2) \quad p_t^{1*} = p_{t-1}^1$$

$$p_t^{2*} = p_{t-1}^2$$

Substitution of (2) into (1) yields:

$$(3) \quad Q_t^{1*} = a + bp_{t-1}^1 + cp_{t-1}^2 + dt + u_t$$

The fundamental assumption of the model is that farmers do not in one year adjust fully to long-run equilibrium production levels. Rather, because of rigidities in reallocating resources, e.g. investment lags, only a proportion of the desired long-run change is achieved. That is,

$$(4) \quad Q_t^1 - Q_{t-1}^1 = \delta (Q_t^{1*} - Q_{t-1}^1)$$

where Q_t^1 is actual output of product 1 in time t and δ is the coefficient (in logs, the elasticity) of adjustment. Substitution of (4) into (3) yields, after grouping and simplifying terms:

$$(5) \quad Q_t^1 = a\delta + b\delta p_{t-1}^1 + c\delta p_{t-1}^2 + d\delta t + (1-\delta)(Q_{t-1}^1) + \delta u_t$$

This may be rewritten as

$$(6) \quad Q_t^1 = \pi_1 + \pi_2 p_{t-1}^1 + \pi_3 p_{t-1}^2 + \pi_4 t + \pi_5 (Q_{t-1}^1) + e_t$$

This is the reduced form of the system (3) and (4). Least squares estimates of the coefficients of (6) enable us indirectly to determine the short-run supply adjustment to prices (π_2 and π_3), the coefficient of adjustment, $1 - \pi_5$, and the long-run elasticities, $\frac{\pi_2}{1-\pi_5}$ and $\frac{\pi_3}{1-\pi_5}$.

Although far from perfect (especially in its assumptions about price expectations), this is a more useful model for use in minimum price programs, than traditional supply analysis. It would better permit an evaluation of the impact of stable prices over a period of 3-5 years. In traditional models, price stability would have no additional impact above and beyond the level of prices fixed in the first year.*

B. A Statistical Problem: multicollinearity.

Multicollinearity between the time trend and lagged output (area) is often very high in Brazil, due to the rapid expansion of the output of several products. Multicollinearity tends to create insignificant and unstable coefficients for one or both of the two variables, and makes it difficult to separate their respective influences. In some cases multicollinearity may, in addition, give the misleading impression of significance of coefficients which are in fact not significant at all.**

* Several more complex distributed lag supply models have been proposed by Nerlove, relaxing the assumption that farmers fully project last year's prices and combining both adjustment lags and expectations coefficients. These models involve more complicated estimation procedures. We decided to use the simpler model as a first approximation, extending the analysis later to other models depending on the quality of initial results. Most published work in Brazil and the United States has used the adjustment model, for which Nerlove himself has expressed a preference. See Journal of Farm Economics, August, 1955 pp. 723-728.

** On this point see, Johnston, John, Econometric Methods, New York, 1963, pp. 204-205.

Thus it is possible for the adjustment coefficient to appear significant merely because of correlation with the time trend^o and even though farmers do not behave as posited by the model.

The problem of multicollinearity often requires estimation in first differences as a check of the stability of the coefficients . Transformation into first differences eliminates time as a variable, but may, however, introduce auto-correlation in the residuals. In any case, previous researchers in Brazil have either ignored this multicollinearity or, while recognizing its possible effects, have not attempted to test the stability of the coefficients in the first difference transformation.

III. STATISTICAL RESULTS.

Supply functions were estimated for two separate regions: the North-Northeast, including Bahia and Sergipe northward, and the Center-South -- Espírito Santo, Minas Gerais, Goiás and Mato Grosso southward. For peanuts, estimates were made only for São Paulo State, which has produced 92-93% of total output in recent years according to SEP. In addition to the advantages given above of separate regional estimates, recent work by Delfim Netto et al^o at an aggregative level concluded that farmers in the Northeast do not respond to price stimuli. Decomposition into individual products may show, however, a significant response to price in the Northeast also.^{***}

SEP price, output and area data were used except for peanuts. All prices were deflated by Conjuntura Econômica's index N° 2, with 1947 = 100. In all products, both output and area response functions were estimated both in linear and exponential (linear in logs) forms. Initially the prices of several probable substitute crops were included. Their coefficients were generally not significant at the 10% level (one-tail test) nor even larger than their standard errors. The functions were then re-estimated excluding the non-significant variables.

^o See Section IV. A. below for the rationale of the time trend variable and the possible results of its misspecification.

^{**} op. cit.

^{***} This was actually the case with rice and mandioca.

Estimates were made for the period 1947-1964, except where otherwise noted.

A. Center-South

1. Cotton

The prices of two probable substitute crops, corn and peanuts, were included initially in the adjustment model, but their estimated coefficients were both smaller than their standard errors. The linear estimate gave the superior fits. Since the explanatory value of the output function was higher in the final form ($R^2 = .79$ as compared with .61), only the former is presented here. The t statistic appears in parenthesis below the coefficient.

Prod = Production in tons

P = Prices in 1941 Cr\$/ton

t = year - 1948 = 0

$$(1) \text{ Prod}_t = 273,963.19 + 223.25^* (P \text{ cotton}_{t-1}) + .24^* (\text{Prod}_{t-1}) + 27,527.36^* (t)$$

$$(4.46) \quad (1.43) \quad (4.20)$$

$$R^2 = .79^*$$

$$d = 1.41$$

Notes: * significant at 1% level.
 ** significant at 5% level.
 *** significant at 10% level.

One-tail tests were used. That is, $H_0: B \leq 0$ for coefficients which theory suggests are positive and $B \geq 0$ for those which should be negative. The more common $H_0: B = 0$, is a less powerful test.

To eliminate the possible effects of multicollinearity between Prod_{t-1} and t ($R = .62$)^{*} and possible auto-correlation in the residuals, the function was also estimated in first differences.

$$(2) \Delta \text{ Prod}_t = 12,877.91 + 209.00^* (\Delta P \text{ cotton}_{t-1}) - .18 (\Delta \text{ Prod}_{t-1})$$

$$(5.34) \quad (1.23)$$

$$R^2 = .79^*$$

$$d = 2.48$$

The explanatory power of both functions is the same: $R^2 = .79$. The implications, however, are quite different. The coefficient of

* See Section I.B above.

$Prod_{t-1}$ in (1) is positive and significant at the 10% level (one-tail test). In (2), however, it is negative (implying $\delta > 1$, which is absurd) and not significant. Thus it seems likely that either adjustment is immediate or that the model does not represent the real structure of supply. It also appears that multi-collinearity in (1) between t and $Prod_{t-1}$ may have created the illusion of significance in the coefficient of $Prod_{t-1}$ in (1).

The first difference estimation is not satisfactory. The constant term, which should represent the time trend coefficient, is negative. This is probably explained econometrically by the fact that only year-to-year changes are being measured. If the annual trend is erratic, the constant term in first differences may not be significant and could be even negative (tests for the significance of the constant were not made).

It appears, then, that the cotton supply function should be re-estimated in the original linear form excluding the lagged production variable.

Short-run price elasticities of cotton output calculated at the means of 1962-64 are .60 for (1) and .56 for (2), showing a substantial response to price stimuli. The "long-run" elasticity in (1) is .78, although of doubtful validity.

The predictive value of the two functions was about the same for 1965 output. (1) underestimated output in 1965 by 6.5% and (2), re-estimated by omitting the change in $Prod_{t-1}$ variable, by 6.7%

2. Rice

With rice only area functions were estimated, since upland rice, which predominates in São Paulo and the Estados Centrais, is subject to large yield variations from fluctuations in rainfall. The correlation between lagged area and time was .98 in the exponential form. Hence, the coefficients were estimated in first differences. No substitute product tested -- corn, beans or cotton -- was significant.

A = Area in ha.

P = Price in 1947 Cr\$

$$(1) \Delta \ln A_t = .0340 + .2055^* (\Delta \ln Price_{t-1}) + .2773^{***} (\Delta \ln A_{t-1})$$

(3.6937) (1.3840)

$$R^2 = .52$$

$$d = 2.33$$

The lagged area coefficient is significant at only the 10% level and implies an elasticity of adjustment of .72.* The short-run price elasticity is .21 and the long-run, only .28. These low response coefficients are very close to those estimated by Delfim Netto for the same region without first differences, although the significance of the price coefficients is higher in the first differences form presented here. In this case the high multicollinearity does not appear to have disturbed the estimates.

Estimates made by EPEA for rice in São Paulo and Rio Grande do Sul also confirm a substantially lower price elasticity for rice than for cotton or peanuts (see Tables II and III below). The São Paulo function, with better data and less aggregation, showed corn and perhaps cotton as significant substitutes. In Rio Grande do Sul, irrigated rice has no short-run substitute. As would be expected with irrigated rice, which requires additional investment for production expansion, the coefficient of adjustment was significant at the 1% level and was .44. The Rio Grande function should be re-estimated in first differences ^{however} to test the stability of the coefficients, since the correlation between lagged area and time was .86.

* In the log form the coefficients themselves are the elasticity estimates.

T A B L E II

= 10 =

RICE SUPPLY, SÃO PAULO - 1950-63, excluding 1954 and 1957

A = Area in 1,000 ha.

P = Prices in 1948/52 Cr\$ deflated by price index of 23 products paid to São Paulo farmers, excluding coffee. Corn, rice: 60 kg. Cotton: 15 Kg.

All variables in natural logs.

$$A_t = 5.65 + .36^* (P \text{ rice}_{t-1}) - .29^{**} (P \text{ corn}_{t-1}) \\ (5.66) \quad \quad \quad P \text{ cotton}_{t-1} \quad (-1.89)$$

$$R^2 = .78$$

$$d = 1.65$$

Price elasticities: rice/cotton -.36
corn -.29

Source: Divisão de Economia Rural for all data.

T A B L E III

RICE SUPPLY, RIO GRANDE DO SUL, 1950-62

A = Area in ha.

P = Price of 50 Kg. em casca in 1953 Cr\$.
Deflator = Conjuntura Econômica's No.2

t = Year - 1950 = 1

All variables in natural logs

$$A_t = 3.9908 + .2608^* (P \text{ rice}_{t-1}) + .5512^* (A_{t-1}) + .2898^{***} (t) \\ (2.8057) \quad \quad \quad (3.6103) \quad \quad \quad (1.7407)$$

$$R^2 = .93^*$$

$$d = 1.44$$

$$= .44$$

Price elasticities: Short-run .26
Long-run .63

Source: Instituto Rio Grandense do Arroz for Area and farm price data

The predictive value of the Center-South rice area function does not appear terribly great. For 1965, it underestimated area planted in rice by 9.7%.

3. Peanuts - São Paulo

São Paulo has produced 92-93% of Brazil's total output of peanuts in recent years according to SEP. Supply response functions were estimated only for São Paulo State, using data from the Divisão de Economia Rural. The greater accuracy of these data more than compensates the exclusion of the 7-8% of output produced by other states in the Center-South.

Cotton was included as a substitute product. Log and linear functions were estimated for production and area planted. The log production response function gave substantially higher significance to the coefficient of cotton price, and only it is presented here.

Prod = Production in 1,000 sacks of 25 Kg. em casca.

P = Price in 1947 Cr\$

Peanuts : 25 Kg

Cotton : 15 kg

t Year : 1949 = 100

Variables in Natural logs, except for t.

$$\text{Prod}_t = 1.30 + 1.48^* (P \text{ peanuts}_{t-1}) - .64^{***} (P \text{ cotton}_{t-1}) + .49^{***} (\text{Prod}_{t-1}) + .05^{**} (t) \\ (4.37) \quad (-1.58) \quad (2.61) \quad (2.05)$$

$$R^2 = .92$$

$$d = 1.50$$

$$\gamma = .51$$

The correlation between Prod_{t-1} and t was .85, high enough to have created serious problems of significance and instability in the coefficients of these two variables. The correlation between t and cotton price $_{t-1}$ was - .48, which could be slightly troublesome. To check the stability of the coefficients when the multicollinearity is eliminated, they were re-estimated in first differences. This gave:

$$\Delta \ln \text{Prod}_t = .04 + 1.40^* (\Delta \ln P \text{ peanuts}_{t-1}) - .91^{**} (\Delta \ln P \text{ cotton}_{t-1}) + \\ (4.30) \quad (2.37)$$

$$.47^{**} (\Delta \ln \text{Prod}_{t-1}) \\ (2.17)$$

$$R^2 = .69$$

$$d = 2.02$$

$$\gamma = .53$$

Only the coefficient of cotton price changes substantially in first differences, from $-.64$ to $-.94$ while the significance level of both prices rises and that of lagged output falls somewhat. The choice between the two functions depends principally upon their predictive value. Neither function, however, predicted well for 1965 and 1966, in spite of the high R^2 . Errors in both years were greater than 20%, for both functions.

One thing is clear from the functions and previous work: peanut production is highly responsive to relative prices both in the short and long-runs. Short-run elasticities appear to be 1.40-1.50 with respect to peanut price and perhaps $-.64$ to $-.94$ with respect to cotton price. The adjustment coefficient is significant and about .5, meaning that about 50% of adjustment to long-run equilibrium occurs in the first year. Long-run elasticities are very high (see Table IV)*

TABLE IV
LONG-RUN PRODUCTION RESPONSE ELASTICITIES

	<u>PEANUTS</u>	
	<u>Log Form</u>	<u>Log Form in First Differences</u>
Peanut Price	3.02	2.98
Cotton Price	-1.31	-2.00

Production of the crop is among the most unstable, in part as a result of this price responsiveness. Minimum prices in peanuts could bring substantial benefits. But given the vigorous responsiveness of cotton and peanuts producers to prices, special care must be exercised in fixing minimum prices near their equilibrium levels to avoid serious distortions in resource allocation.

4. Mandioca

Some slight and still unsatisfactory evidence was found of the response to mandioca and corn prices of area planted in mandioca. Production response functions, on the other hand, gave considerably

* See section IV below for the limitations of such estimates and their possible biases.

poorer results than area functions, with lower R^2 and non-significant price coefficients. The function estimated is as follows:

A = Area in ha

P = Price in 1947 Cr\$/ton

t = year, 1948 = 0

All variables in natural logs, except t.

$$A_t = - .67 + .13^{**} (P \text{ mandioca}_{t-1}) - .073^{***} (P \text{ corn}_{t-1}) \\ (1.79) \quad (-1.50) \\ + 1.04^{*} (A_{t-1}) + .003 (t) \\ (6.30) \quad (.46)$$

$$R^2 = .99$$

$$d = 2.75^{**}$$

Clearly the coefficient of A_{t-1} has been determined mainly by its high correlation with time (.98). Its magnitude implies a δ of -.04, an absurd figure. Multicollinearity also caused the non-significance of the trend, and A_{t-1} apparently represents supply shifts over time better than the exponential growth function included in the equation.

Some weak response to price may be indicated -- estimated short-run elasticities were .13 with respect to mandioca price and -.07 with respect to corn price. In any case the equation should be re-estimated in first differences or by eliminating A_{t-1} altogether. This work has not yet been completed.

5. Corn

All attempts to estimate supply functions for corn area and output failed. In part this is a result of the fact that over 50% of corn produced appears to be consumed by the producing unit as animal feed, mainly for hogs and poultry. We attempted to take account of this factor by including the price of hogs in the previous year as an independent variable. In the realization, however, no price coefficients were significant at the 10% level -- not corn, hogs, rice nor beans. Time and lagged area (output) were the only significant coefficients ,

hardly a satisfactory state of affairs.

Sérgio Brandt's attempt to estimate a corn supply function for São Paulo using Divisão de Economia Rural data showed corn price significant only at 5-10% levels (one-tail test), and R^2 on the order of only .78 to .83 in functions with strong trends and including the lagged dependent variable.* In the absence of better data on livestock feeding, the possibility for improvement in corn supply functions seems slight.

6. Conclusions -- Center-South

We have found, then, significant response to prices in the Center-South for cotton, peanuts, rice and perhaps mandioca. The corn functions were unsatisfactory. Beans and soybeans were not attempted. Except perhaps for the peanut function for São Paulo, the other functions must be treated with reservations, particularly as we shall see because of the quality of the basic data. The functions may be sufficient to show what close observers really knew all the time, that farmers in Brazil do respond to prices, but their use for quantitative prediction within the minimum price program is necessarily restricted.

With cotton and mandioca, we saw the pitfalls which may arise because of high correlation between the lagged dependent variable and the time trend, although with peanuts and rice, multicollinearity does not appear to have disturbed the estimates.

B. North-Northeast

Supply functions were estimated for cotton, corn, rice and mandioca. The cotton function was poor, mainly because it was impossible to incorporate satisfactorily the behavior of tree cotton, perhaps 58% of total cotton production in the Northeast.** Tree cotton takes

* See Sérgio Brandt et al, Estrutura da Oferta de Milho no Estado de São Paulo, (mimeo) São Paulo, 1964. Multicollinearity between time and lagged production led to a negative time trend, a clearly erroneous result for this product.

** See Mercado e Comercialização do Algodão do Nordeste, BNB/SUDENE, Fortaleza, 1964, p.55.

two years as rule to produce in quantity, and then it lasts for 4-5 years. Corn showed no significant price coefficients. For rice and mandioca, however, there is some indication that farmers in the Northeast do respond to price stimuli.

In all estimates the years 1951-52 and 1958-59 were eliminated because of the serious droughts in 1951 and 1958. Given the strong influence of weather on yields in the region, only area functions were estimated.*

1. Rice

The correlation between A_{t-1} and time was .98. Thus estimates were made eliminating A_{t-1} as an independent variable. The results of the linear function are: **

A = ~~Area~~ Area in ha

P = Prices in Cr\$ 1947

$$t = \text{Year, } 1948 = 0$$

$$A_t = 28599.29 + 163.79^{**} (P \text{ rice}_{t-1}) - 25.76 (P \text{ beans}_{t-1}) + 32442.47^{**} (t)$$

(2.69)

(-1.36)

(18.81)

R² = 98

$d = 1.50$

The coefficient of corn price was smaller than its standard error and was not included in the final form presented here.

Producers of rice, then, appear to respond to price changes, as the price elasticities calculated at the means, 1962-64, show.

TABLE V
PRICE ELASTICITIES OF RICE AREA,
NORTH-NORTHEAST

Rice price	.30
Bean price	-.10

* It does not appear that the distortions introduced by serious droughts in the Northeast were taken into account in the ANPES study cited above. This may help explain the non-significant price coefficients it found for the region.

** Although the R^2 's were the same, .98 the linear function had higher significance in the price coefficient.

And the response in the North-Northeast is of the same order as in the Center-South, .30 in the short-run as compared with .21 in the Center-South.

In addition to the function presented here, however, the adjustment model should be re-estimated in first differences. This work has not been completed.

2. Mandioca

Serious multicollinearity between prices* made it impossible to distinguish the effects of mandioca price from those of potential substitute products, corn and beans, and these prices had to be excluded.

$$\begin{aligned} R_{\text{corn mandioca}} &= .91 \\ R_{\text{beans mandioca}} &= .79 \end{aligned}$$

The results of the linear model for area are:

A = Area in ha

P = Price in 1947 Cr\$/ton

t = Year, 1948 = 0

$$A_t = 186,983.16 + 301.39^{**} (P_{\text{mandioca}_{t-1}}) + .54^{**} (A_{t-1}) + 9572.85^{**} (t) \\ (2.31) \quad (3.40) \quad (2.85)$$

$$R^2 = .99^{**}$$

$$d = 2.44$$

$$\delta = .46$$

The correlation between lagged area and time was .96. Thus the coefficients for the trend and A_{t-1} may be highly unstable and perhaps erroneously significant. This possibility should be tested using first differences. The estimation of the function excluding lagged area gave coefficients for the price and mandioca less than their standard errors.

* There is a strong upward trend in deflated prices for most agricultural products in the Northeast, perhaps arising from the inability of the regions agriculture to meet the rising demands of the area.

Price elasticities calculated at the means, 1962-64, were .09 in the short-run and .18 in the long-run. As in the Center-South the response of mandioca to price appears very small, although statistically significant.

3. Conclusions

Thus there is some evidence, at least, that farmers in the Northeast respond to price stimuli for rice and mandioca much as their colleagues in the Center-South. Cotton, a completely commercial crop, should show the greatest response. Since 58% of the crop in the Northeast appears to be tree cotton, it was not possible to set up a satisfactory model for total cotton supply. It would be necessary to estimate two separate cotton supply functions, but there is no breakdown of the two types in available time series data. The inadequacy of our data and model, however, does not lead to the conclusion that cotton producers in the Northeast do not respond to price stimuli. It appears then that ANPES' conclusions that farmers in the Northeast do not respond to price are exaggerated.

IV. CRITIQUE OF THE RESULTS

We turn now to an analysis of the limitations of the estimates presented. Two problems will be treated. (1) probable bias in the coefficient of adjustment and long-run price elasticities and (2) the poor quality of the data employed. The points here are applicable not only to the present estimates, but to all previous and future work in Brazil using Nerlove models and/or SEP data.

A. Probable Bias in the Estimation of the Adjustment Coefficient

The Nerlove adjustment model is subject to probable bias in the estimation of $(1 - \delta)$.^{*} The bias arises from the omission of independent variables correlated with lagged output (area). It can be shown that bias in $(1 - \delta)$ will also bias the estimate of the long-run supply elasticity.^{**}

* All Nerlove's models, including as they do some combination of the lagged dependent variable, are subject to the same bias.

** On this point see G.E. Brandow's critique of Nerlove's model and Nerlove's reply in the Journal of Farm Economics, August, 1958, pp. 719-728.

The problem occurs more or less as follows. Much the greater part of production variations in Brazil are not explained directly by price variations, but by shifts in the supply function which are assumed to leave the shape of the supply function unchanged. The principal sources of these shifts are greater factor supply (land and labor) and changes in production costs arising from technological improvement and/or changes in the prices of inputs. In Brazil data on these shift factors are extremely sparse. Thus they are inevitably excluded from the statistical supply function. The inclusion of a time trend is a rough attempt to capture the effects of these excluded shift variables.

These shift variables will not, in general, change at constant rates over time, but may come in spurts, or they may act only in part of the period considered. Examples of such shifts important for the current analysis are: (1) radically changing real prices of fertilizers of the period examined; (2) the dissemination of hybrid corn seeds beginning in the late 1950's; (3) great improvement in cotton seeds in the middle and late 1950's; (4) a much more liberal official credit policy to agriculture in the late 1950's and 1960's; (5) the rise and subsequent decline of coffee prices in the 1950's and 1960's.

Such irregular shifts will not be captured adequately by the time trend alone. However lagged output (area) will probably reflect well these shifts in the supply curve, and will probably be correlated with the residual variations arising from these excluded variables and not explained by the trend alone. In this case, the coefficient of lagged output (area) reflects not only the coefficient of adjustment but also the effects of excluded variables correlated with lagged output and not fully explained by the trend. This biases both the coefficient of adjustment and the long-run elasticity. In general, the correlation will be positive and the bias upward.

The example of hybrid corn is illuminating. Suppose that the availability of land and labor for corn grows at a constant exponential

rate, shifting the supply function at that same rate. Suppose further that after 1958 hybrid corn is introduced and disseminated at a rapid rate. Production rises more rapidly as costs of production are reduced. This will be well represented by lagged output but not by the typical time trend. Thus the use of hybrid corn will be positively correlated with lagged output, as will be the residual not explained by prices and the trend.

We would expect, then, that the coefficient of lagged output would be more significant than that of the trend (it reflects all shift variables) and to be biased.* In first differences, the change in output will be positively correlated with the residual effects of the excluded shift variables after subtracting the time trend effect. Thus in both cases the coefficient of lagged output (area) will be biased upward as will be the long-run elasticity.** The extent of this bias for each product estimated here is impossible to evaluate without additional information. All that is required for bias, however, is a misspecification of the form of the trend in the shift variables.

This source of bias -- the exclusion of independent variables correlated with some of those included -- is not unique to the adjustment model. Indeed, if the model represents reality, then the exclusion of lagged output (area)*** would give an upward bias to the time trend coefficient and perhaps a downward bias to the price coefficient (if output and price in t are negatively correlated). It does mean, however, that we must be cautious in interpreting the adjustment coefficients and long-run elasticities, especially when the latter

* This in fact occurred in an estimation of corn supply in São Paulo made by Sérgio Brandt et al. The coefficient of lagged output was highly significant, while that of time was negative and not significant at the 5% level. See Estrutura da Oferta de Milho no Estado de São Paulo, (mimeo), Divisão de Economia Rural, S. Paulo, 1964.

** For a proof of the biases arising from excluding variables correlated with lagged output, see Marc Nerlove in Journal of Farm Economics, August, 1958, pp. 723-28.

*** that is, a traditional supply function.

appear very high.³ While biases in the adjustment coefficient are not of great importance if our main interest is to show only that farmers respond to prices, they assume considerable weight if we are trying to show the impact of stable agricultural prices over a period of, say, three years, arising from a more effective price support program.

B. Data Problems

The output and area data supplied by SEP are of doubtful precision, and the low quality of the data seriously vitiates statistical estimation of supply response. Recently it has been argued that since some supply functions show a significant response of farmers to relative prices in the Center-South, SEP's data cannot be as bad as many people have thought. This argument is inconclusive. Even poor data may show some response to price, but the magnitude of this response and especially of the response to substitute products may be considerably misstated.

In evaluating SEP data we need not be limited to such speculations. For the State of São Paulo we have since 1954 output and area data estimated by a scientific sampling of approximately 2,000 farms. These estimates are quite close to independent estimates for coffee and cotton available through IBC's control mechanism and data on deliveries to cotton mills in the State. SEP's estimates, on the other hand, are made subjectively by IBGE's statistical agents, many of whom are in no position to make accurate estimates.

Comparisons of production data from these two sources for São Paulo show great differences both in year-to-year fluctuations and in

³ An example of this occurs with Sérgio Brandt's estimate of rice supply in São Paulo. The coefficient of adjustment was only .15, giving a long-run supply elasticity with respect to rice price of 4.10. The low adjustment coefficient has no economic foundation for upland rice and appears to be the result of the omission of other variables correlated with lagged output, and/or a misspecified trend. See Estimativas de Oferta de Amendoim, Arroz e Mamona no Estado de São Paulo (mimeo), Divisão de Economia Rural, São Paulo, 1965.

trends, as is shown in Table I. The divergences are especially acute for corn, rice and beans. Year-to-year fluctuations are much smaller, as a rule, in the SEP data. If such fluctuations are primarily in response to prices, we might expect considerable downward bias in supply elasticities based on SEP figures. The smaller annual fluctuations in output may arise in part from IBGE's rule that reported declines in output greater than 3% and increases greater than 10% must be justified in writing by the statistical agent.

TABLE I
COMPARISONS OF PRODUCTIONS ESTIMATES FOR SÃO PAULO,
1954/64 FROM SEP AND THE DIVISÃO DE ECONOMIA RURAL

		<u>Average Annual Percentage Change in Output, Absolute value</u>	<u>Percentage Change in Output, 1954/56 to 1962/64</u>
Corn	DER	20.3	72.6
	SEP	9.1	25.5
Rice	DER	18.4	43.6
	SEP	11.2	-1.0
Beans	DER	28.2	48.1
	SEP	10.7	-24.6
Cotton	DER	16.3	3.8
	SEP	12.6	16.2
Peanuts	DER	28.3	162.5
	SEP	21.3	250.1

Sources: Anuário Estatístico and Divisão de Economia Rural da Secretaria da Agricultura do Estado de São Paulo.

Comparisons between the levels and growth of output 1940-50 between the Agricultural Census and SEP have also shown enormous differences,* although in this case the Census may also be of questionable quality.

* See Conjuntura Econômica, Agosto, 1955, pp. 61-68.

The indications are, then, that SEP's data are of dubious quality, as is to be expected given the methods by which they are gathered. No amount of econometric manipulation can improve the quality of the data, and hence all results obtained using SEP's figures must be taken with a grain of salt.

Still more sophisticated supply models could be estimated. In my opinion, because of the data problem, further refinement of the supply functions is not warranted, and their results would be of dubious practical use. The most efficient way to improve supply and demand functions for policy use is to revamp completely the whole data system for agriculture. Progress on this front has lagged far behind the eagerness and sophistication of econometricians in Brazil.