CA

WP172

AGRICULTURAL DEVELOPMENT SYSTEMS EGYPT PROJECT

UNIVERSITY OF CALIFORNIA, DAVIS

REGRESSION ANALYSIS OF POTATO AND TOMATO PRICES AT ROD EL FARAG MARKET, CAIRO, WEEKLY DATA, 1979-1981

by

Jerry Foytik
University of California, Davis
Nabil Habashy
Ministry of Agriculture, Egypt

Economics
WORKING PAPER

USAID/EGYPT

REGRESSION ANALYSIS OF POTATO AND TOMATO PRICES AT ROD EL FARAG MARKET, CAIRO, WEEKLY DATA, 1979-1981

by

Jerry Foytik
University of California, Davis
Nabil Habashy
Ministry of Agriculture, Egypt

Assistance from the Agricultural Development Systems Project of the University of California, Egyptian Ministry of Agriculture, and USAID, is gratefully acknowledged, but the author is soley responsible for the views expressed in this paper.

Economics
Working Paper Series
No. 172

Note: The Research Reports of the Agricultural Development Systems: Egypt Project, University of California, Davis, are preliminary materials circulated to invite discussion and critical comment. These papers may be freely circulated but to protect their tentative character, they are not to be quoted without the permission of the author(s).

July, 1983

Agricultural Development Systems:
Egypt Project
University of California
Davis, Ca 95616

REGRESSION ANALYSIS OF POTATO AND TOMATO PRICES AT ROD EL FARAG MARKET, CAIRO, WEEKLY DATA, 1979-1981[1]

By Jerry Foytik and Nabil Habashy

This study is concerned primarily with the price-quantity relationships prevailing for potatoes and tomatoes sold at the Rod El Farag market during 1979-1981. Weekly data are used for determining the magnitude of shifts in levels occurring over the course of a year. Price is related to quantity and the shift variables, rather than using quantity as the dependent variable. Several regressional equations are computed to indicate how much results vary when different models are assumed to express how the price mechanism operates in this portion of the distributive system.

The analysis indicates that meaningful relationships can be formulated. In brief, price is definitely related negatively to the quantity sold and positively to a trend factor. Also, the level of the price-quantity function shifts substantially and in a somewhat orderly fashion on a weekly basis.

METHOD

Ordinary least squares methods (OLS) are used for relating the median weekly price (Y) at the Rod El Farag market to the quantity sold (Q), trend (T), and weekly dummy variables (X_i). Several linear functions are fitted. No attempt is made to fit logarithmic or other curvilinear functions. The identical equations are used for relating the minimum weekly price (P) to the independent variables. Both sets of regressions are fitted for potatoes and for tomatoes. For example, one equation for each commodity is expressed

---

[1] An analysis of monthly prices for grapes, potatoes, and tomatoes sold at the El Nozha market, Alexandria, is reported in Working Paper 171 of this series.

$$P = A + B_1Q + B_2T + C_1X_1 + C_2X_2 + . . . + C_nX_n.$$

As will be indicated below, all OLS equations gave low values for the Durbin-Watson statistic, indicating the existence of positive autocorrelation among successive observations. Hence, autoregressive equations of the first order are computed.

DATA USED

The Egyptian Ministry of Agriculture collects various prices. One set gives the weekly range in prices. It provides two price series used for the dependent variable: the low end and median, or middle, of the weekly range. The highs could be used also, but are not here.[2]

Only three independent variables are used because of the limited availability of information. Hence, some factors considered relevant, e.g., consumer income, could not be included. Quantity is represented by weekly sales as reported by the Ministry.

Dummy variables are introduced for determining the weekly changes in the level of the price-quantity function. Having weekly data for three years provides numerous observations--probably more than required to determine meaningful results. Furthermore, shifts between successive weeks are assumed to be relatively small. Hence, a sample of weekly observations is taken for fitting the regression equations.

For potatoes, every fifth week starting with week 4 of each year is included, giving 30 observations. The number is doubled by also including the

---

[2]A third equation could be fitted using the high prices of the monthly range. This was not done merely to avoid increasing the number of equations fitted.

immediately following week--i.e., weeks, 5, 10, etc. Then each pair of adjoining weeks is considered to represent two separate observations for a biweekly period. Thus data for weeks 4 and 5 for each year give two observations for period 1, data for weeks 9 and 10 are observations for period 2, etc. This means that the biweekly periods are separated by three weeks of data omitted in fitting equations. Five weeks separate period 10 of one year and period 1 of the next.

Weekly data furnished for tomatoes cover only the first six months of each year. Hence, the biweekly periods are specified somewhat differently. Pairs are separated by only one week. Thus period 1 includes weeks 1 and 2, period 2 includes weeks 4 and 5, . . ., period 9 includes weeks 25 and 26.

A time trend is introduced as a proxy for the combined net effect of omitted factors. It is measured in biweekly periods, taking into account the exclusion of July-December for tomatoes. Values for potatoes range from T = 1 for period 1 of 1979 to T = 30 for period 10 of 1981. For tomatoes, the range is from T = 1 to T = 43.

Data for the weeks included in the analysis appear in Tables A and B (see end of report). The averages given at the foot of each column refer only to these data. That is, they are not averages for all weeks, including those omitted in fitting the equations.

OLS RESULTS

Table 1 summarizes the values computed for four pairs of equations--two pairs use average price (Y) for the dependent variable and two use minimum price (P). The top equation of each pair includes the dummy variables (X_i) to represent periodic shifts in the price-quantity equation, while the bottom equation omits these shifts.

TABLE 1: Potato Prices at Rod El Farag Market in Cairo
Regression Analysis of Weekly Data, 1979-1981

Eqn.	Dependent Variable & Constant	Regression Coefficient			Supplemental Measures		
		Q	T	X	R^2	SE	DW
1	Y = 95.286	-1.041 (3.1)	0.936 (6.4)	a/	.7456	10.98	0.91
2	Y = 96.475	-1.419 (5.6)	1.146 (5.0)		.4957	15.45	0.85
3	Y = 110.646	-1.113 (2.6)		a/	.6000	13.76	0.63
4	Y = 115.375	-1.516 (5.0)			.2909	18.33	0.62
5	P = 94.528	-1.433 (3.7)	0.893 (4.4)	a/	.6873	12.79	0.98
6	P = 92.686	-1.551 (5.7)	1.101 (4.5)		.4802	16.49	0.88
7	P = 109.180	-1.503 (3.3)		a/	.5694	15.01	0.76
8	P = 110.841	-1.644 (5.3)			.3108	18.98	0.68

Figures in parentheses are t-statistics for regression coefficients. Measures R^2 and SE are adjusted for degrees of freedom.

P - low of the high-low range of weekly prices, LE per ton. $\bar{P} = 91.55$

Y - middle of this range, LE per ton. $\bar{Y} = 97.75$

Q - weekly sales, in 100 tons. $\bar{Q} = 11.7333$

T - time measured in biweekly subperiods, T = 1 for weeks 4 and 5, 1979--
see Table A. $\bar{T} = 15.5$

X - shifts in levels for biweekly periods from average annual level.

a/ - biweekly levels were computed for these equations. The following values for equations (1) and (5) are representative of other equations.

Week	Equation 1	Equation 5	Week	Equation 1	Equation 5
4 & 5	- 9.85	- 8.83	29 & 30	2.77	0.62
9 & 10	-10.43	-10.35	34 & 35	5.57	3.07
14 & 15	8.76	13.57	39 & 40	20.83	18.75
19 & 20	-15.10	-14.90	44 & 45	16.12	13.36
24 & 25	-12.21	-13.52	49 & 50	- 6.46	- 1.77

Examination of the tabulation reveals several results of interest. The net effect of quantity (Q) and of trend (T) is in the expected direction and is highly significant statistically. Introducing the weekly shifts improves the fits considerably, i.e., SE is reduced and hence R^2 is increased. Furthermore, these weekly changes appear to follow a somewhat regular pattern, rather than merely being random shifts. Including these shifts in the formulation serves to reduce the net effect of T and of Q. Computed values for the Durbin-Watson statistic are so low that the existence of positive autocorrelation cannot be rejected.

Relating price to quantity, trend, and the weekly shifts provides the best OLS fits. From the statistical view point equation (1) gives the best explanation of price variations during 1979-81. About 75 percent of the variance in potato prices is accounted for by fluctuations in sales and by changes due to the upward trend and the weekly shifts occurring during the course of a year.

Values for the same eight equations for tomato prices are in Table 2. Generally, they indicate results similar to those discussed for potatoes. However, they do differ in some respects. Introducing the weekly shifts improves the fit to a lesser extent, decreases the net regression coefficient for quantity more than in the case of potatoes, and increases the trend effect rather than reducing it.

Again, the best OLS fits are those relating price to quantity, trend, and the weekly shifts. However, representing the dependent variable by average price or minimum price gives equally good fits. About 80 percent of the variance in tomato prices is explained by both equations (1) and (5).

TABLE 2: Tomato Prices at Rod El Farag Market in Cairo
Regression Analysis of Weekly Data, 1979-1981

Eqn.	Dependent Variable & Constant	Regression Coefficient Q	T	X	Supplemental Measures R^2	SE	DW
1	Y = 138.050	-1.861 (3.9)	1.232 (4.7)	a/	.7964	23.57	0.82
2	Y = 193.219	-3.284 (9.6)	0.883 (2.9)		.7045	28.40	0.81
3	Y = 198.713	-2.889 (5.7)		a/	.6974	28.74	0.52
4	Y = 221.344	-3.584 (10.2)			.6623	30.36	0.73
5	P = 133.558	-2.032 (4.2)	1.202 (4.5)	a/	.7942	24.19	0.83
6	P = 188.036	-3.437 (10.2)	0.809 (2.9)		.7268	27.88	0.82
7	P = 192.795	-3.035 (5.9)		a/	.7042	29.00	0.55
8	P = 215.338	-3.728 (10.9)			.6887	29.75	0.74

Figures in parentheses are t-statistics for regression coefficients. Measures R^2 and SE are adjusted for degrees of freedom.

P - low of the high-low range of weekly prices, LE per ton. $\bar{P}$ = 93.704

Y - middle of this range, LE per ton. $\bar{Y}$ = 104.426

Q - weekly sales, in 100 tons. $\bar{Q}$ = 32.6231

T - time measured in biweekly subperiods, T = 1 for weeks 1 and 2, 1979, see Table B.

X - shifts in levels for biweekly periods from average annual level.

a/ - biweekly levels were computed for these equations. The following values for equations (1) and (5) are representative of other equations:

Week	Equation 1	Equation 5	Week	Equation 1	Equation 5
1 & 2	-12.69	-11.79	16 & 17	55.15	50.89
4 & 5	- 7.36	- 8.91	19 & 20	0.41	- 2.51
7 & 8	- 0.78	- 0.99	22 & 23	-31.98	-30.10
10 & 11	6.72	8.00	25 & 26	-24.90	-23.14
13 & 14	15.43	18.55			

AUTOREGRESSION

Values for the Durbin-Watson statistic are very low for all the above 16 regression equations and support the hypothesis of positive autocorrelation among successive residuals. Hence, a first order autoregressive model was introduced in an attempt to obtain better estimators for the structural parameters. OLS equations (1) and (5) for potatoes and tomatoes were recomputed by the Cochrane-Orcutt iteration method. Results are given in Table 3.

Durbin-Watson values are raised to almost 2.0 for the recomputed equations. These higher values no longer justify the assumption of positive (or negative) autocorrelation. Of course, the net regression coefficients are altered. Those for Q are changed by about 30 percent--decreased for potatoes and increased for tomatoes. Those for T are changed in the opposite direction and by a smaller percentage.

Introducing the autoregression coefficient rho (ρ) into the formulation reduces the standard error by 16 percent for potatoes and 21 percent for tomatoes. In other words, using equations (1A) or (5A) instead of (1) or (5) improves the accuracy of predictions. It will be noted, however, that the tabulated values of R^2 are decreased. This apparent contradiction can be explained by examining the method used for computing the values.

Equations (1) and (1A) for potatoes are used to illustrate the computations for the two methods. In both cases, R^2 is determined by comparing the variance unexplained by the equation with the variance in the dependent variable. The latter variance is computed differently--with reference to the actual values of Y for (1) and with reference to the rho-transformed values of Y for (1A). An alternate R^2 value for (1A) might be determined by relating the variance unexplained by (1A) to the variance in

TABLE 3: OLS and Autoregressive Results for Potatoes and Tomatoes
Rod El Farag Market, Cairo, Weekly Data 1979-1981

Eqn.[a]	Constant	Regression Coefficient[b]			Supplemental Measures		
		Q	T	ρ	R^2	SE	DW
Potatoes							
1	Y = 95.286	-1.041 (3.1)	0.936 (6.4)		.7456	10.98	0.91
1A	Y = 89.932	-0.720 (2.6)	1.058 (3.0)	0.592 (5.6)	.5773	9.11	1.93
5	P = 94.528	-1.433 (3.7)	0.893 (4.4)		.6873	12.79	0.98
5A	P = 88.628	-1.033 (3.1)	0.989 (2.5)	0.561 (5.2)	.4921	10.90	1.82
Tomatoes							
1	Y = 138.050	-1.861 (3.9)	1.232 (4.7)		.7964	23.57	0.82
1A	Y = 157.649	-2.433 (5.7)	1.090 (2.0)	0.665 (6.5)	.6878	18.53	1.85
5	P = 133.558	-2.032 (4.2)	1.202 (4.5)		.7942	24.19	0.82
5A	P = 151.110	-2.573 (5.8)	1.085 (2.1)	0.653 (6.3)	.6848	19.17	1.79

Note: See Tables 1 and 2 for symbols used.

[a]The four OLS equations are those listed in Tables 1 and 2. The companion autoregressive equations are identified by the letter "A".

[b]All equations include dummy variables to indicate biweekly shifts from average levels. Shifts for equations (1) and (1A) are:

Potatoes			Tomatoes		
Week	Equation 1	Equation 1A	Week	Equation 1	Equation 1A
4 & 5	- 9.85	- 9.33	1 & 2	-12.69	-24.65
9 & 10	-10.43	-10.97	4 & 5	- 7.36	- 6.03
14 & 15	8.76	5.52	7 & 8	- 0.78	8.60
19 & 20	-15.10	-13.87	10 & 11	6.72	10.61
24 & 25	-12.21	- 9.89	13 & 14	15.43	6.16
29 & 30	2.77	4.87	16 & 17	55.15	46.28
34 & 35	5.57	8.03	19 & 20	0.41	9.40
39 & 40	20.83	23.51	22 & 23	-31.98	-27.66
44 & 45	16.12	13.49	25 & 26	-24.90	-22.71
49 & 50	- 6.46	-11.36			

original Y values. Specifically, the three R^2 values are computed as follows:

	Equation 1	Equation 1A	Alternative
Original SD^2	472.80		472.80
Transformed SD^2		196.45	
SE^2	120.60	83.03	83.03
Difference	352.20	113.42	389.77
R^2	.7456	.5773	.8244

Which of the latter two values is appropriate depends on how the researcher wishes to interpret and use the results obtained. Hence, R^2 values computed by the alternative method are given in the following tabulation together with the values shown in Table 3.

	Potatoes		Tomatoes	
	Eqn 1 & 1A	5 & 5A	Eqn 1 & 1A	5 & 5A
OLS	.7456	.6873	.7964	.7942
Autoregressive	.5773	.4921	.6878	.6848
Alternative	.8244	.7724	.8737	.8707

In all four cases, the alternative R^2 is greater than R^2 for the OLS equation. Of course, this outcome follows from the procedure used because the standard error is reduced by the autoregression model.

Biweekly shifts in the price-quantity function for equations (1) and (1A) appear in the footnote to Table 3. Introducing the rho coefficient does not change the pattern of these shifts very much, particularly in the case of potatoes.

CONCLUSION

The results clearly indicate an empirical relationship of weekly price (minimum or average) to sales, an upward trend, and parallel shifts in the net price-quantity function during the weeks of a year. An autoregressive model is needed to overcome difficulties due to positive autocorrelation.

Several modifications should be introduced to improve the results. For example, the present formulation excludes income as an independent variable because only annual information was available when the analysis was made. This may be an important omission since income increased sharply during the three-year period considered. Of course, the income effect is included indirectly by the trend factor. It would be much better, however, to introduce income explicitly. If monthly (or even only quarterly) estimates can be obtained, they should be used.

The computed equations reveal definite shifts in the price-quantity function on a weekly basis. Admittedly, the indicated weekly effect may follow from the arbitrary manner in which it is represented in the model. For example, the slope of the price-quantity relation may vary, and possibly to a considerable extent, during different portions of the year. This could be tested by determining equations separately for summer and winter seasons or on any other basis that appears reasonable. Possibly weekly markets are interrelated. For example, the price effect of quantity may be due to both current sales and the quantity sold during the immediately preceding week. In that case, lagged sales should be added as a separate independent variable.

Price is used as the dependent variable and quantity as an independent variable for all equations computed for this study. Some researchers prefer to reverse these factors. This can be tried. In that case, it will be necessary to guard against introducing multicollinearity by using both price and income as separately independent variables if the two are highly correlated.

In the meantime, the regressions computed can be used to say something about how the price mechanism operates on the Rod El Farag market. The results can be used for estimating prices for periods beyond 1981 provided underlying market conditions have not changed substantially.

TABLE A: Weekly Input Data For Analysis of Potato Prices
Rod El Farag Market, Cairo, 1979-1981

Week	1979	1980	1981	1979	1980	1981
	Minimum Price P			Median Price Y		
4	67	65	80	75	78	82
5	77	70	80	82	78	82
9	90	66	73	92	76	79
10	92	70	73	96	78	79
14	80	100	90	85	105	95
15	77	88	90	84	93	95
19	60	80	80	65	85	90
20	60	76	80	65	80	90
24	58	90	90	69	95	95
25	80	90	90	84	95	95
29	100	115	90	100	118	105
30	100	120	90	100	120	105
34	100	100	120	100	110	127
35	100	100	120	100	110	127
39	111	135	128	116	135	138
40	114	135	130	117	135	140
44	95	80	150	108	90	150
45	90	80	152	112	88	153
49	51	80	114	59	84	118
50	50	80	101	58	85	105
Aver.	82.6	91.0	101.0	88.4	96.9	108.0
	Quantity Q			Time T		
4	1322	2251	1218	1	11	21
5	2015	1311	1239	1	11	21
9	707	1813	1491	2	12	22
10	546	1330	1465	2	12	22
14	2205	2030	2520	3	13	23
15	2422	2023	2283	3	13	23
19	903	1785	1393	4	14	24
20	1257	1484	1323	4	14	24
24	483	1377	1106	5	15	25
25	266	833	707	5	15	25
29	385	651	427	6	16	26
30	366	673	399	6	16	26
34	224	633	532	7	17	27
35	224	602	343	7	17	27
39	252	231	133	8	18	28
40	483	371	139	8	18	28
44	1141	1750	210	9	19	29
45	1904	1981	140	9	19	29
49	2870	1346	2889	10	20	30
50	2814	1176	2002	10	20	30
Aver.	1139	1283	1098	5.5	15.5	25.5

Symbols used:
P - low of the high-low range of weekly prices, LE per ton
Y - middle of this range, LE per ton
Q - weekly sales, in tons
T - time measured in subperiods with both weeks of each pair assigned the same value, T = 1 for weeks 4 and 5, 1979.

Source: Data supplied by Ministry of Agriculture, Egypt

TABLE B: Weekly Input Data For Analysis of Tomato Prices
Rod El Farag Market, Cairo, 1979-1981

Week	1979	1980	1981	1979	1980	1981
	Minimum Price P			Median Price Y		
1	40	80	88	48	95	98
2	40	79	90	48	90	100
4	48	62	100	58	78	112
5	40	55	100	55	70	112
7	37	67	120	54	78	130
8	56	78	103	64	89	118
10	89	98	114	94	108	126
11	84	101	146	92	114	152
13	100	110	176	108	120	176
14	100	149	213	108	154	213
16	80	251	240	98	274	240
17	79	200	240	90	220	240
19	59	131	75	68	148	96
20	55	103	58	62	116	80
22	58	48	50	64	54	70
23	69	45	81	76	51	94
25	32	84	79	50	92	87
26	26	84	70	40	91	76
Aver.	60.7	101.4	119.1	70.9	113.4	128.9
	Quantity Q			Time T		
1	4568	4344	2503	1	18	35
2	3895	4275	2236	1	18	35
4	4031	5299	3022	2	19	36
5	4098	4213	2953	2	19	36
7	4322	4375	3583	3	20	37
8	3720	4039	3516	3	20	37
10	2940	3213	3115	4	21	38
11	3206	2794	2880	4	21	38
13	2682	2122	1591	5	22	39
14	2203	1494	895	5	22	39
16	2971	557	470	6	23	40
17	3318	1396	270	6	23	40
19	3409	3296	3649	7	24	41
20	3380	5533	4273	7	24	41
22	2774	5343	3399	8	25	42
23	3146	4647	2837	8	25	42
25	4141	3366	3457	9	26	43
26	4833	4121	3455	9	26	43
Aver.	3535	3579	2672	5.0	22.0	39.0

Symbols used:
P - low of the high-low range of weekly prices, LE per ton
Y - middle of this range, LE per ton
Q - weekly sales, in tons
T - time measured in subperiods with both weeks of each pair assigned the same value, T = 1 for weeks 1 and 2, 1979

Source: Data supplied by Ministry of Agriculture, Egypt