Wheat Prices in Spain, 1857-1890: An Application of the Box-Jenkins Methodology

Daniel Peña - Nicolás Sánchez-Albornoz *
University of Madrid - New York University

The early statistical tools which economic historians used in time series analysis (simple or moving averages, index numbers, linear trends...), although very simple, proved of great importance for the development of economic history as a discipline. Now, however, they seldom satisfy the substantive concerns of contemporary economic historians.

Elaborate methods of univariate and, even better, of multivariate time series analysis are now opening new fields for historical research. Statistical techniques developed particularly during the 1970's, as yet not well known by historians, suggest that we may be entering into a new phase in such research. By means of parametric models, the typical behaviour of the various components of time series may be accurately described allowing for a wider range of hypotheses to be considered. The dynamic analysis of several series helps to identify what direction the relationship between them follows. Causality, a matter of deep concern for historians, can thus be tackled.

In the following discussion, we will explore one application of this approach to historical data. Time series of prices were used in early statistical analyses by economic historians; they may help, once again, to test the modern tools. The particular set of price data on which we will operate, has already been

* Escuela Técnica de Ingenieros Industriales, Universidad Politécnica (Madrid) and Department of History, New York University. The authors would like to thank the Escuela de Organización Industrial for the use of their facilities in the processing of data.

through a variety of statistical analyses. The data proved to be highly reliable. They may now serve for further exploration of their content.

Substantively, our aim is to lay the ground for studying the formation of the national domestic market by looking at the behaviour of prices at different local markets. An increasingly strong relationship among them should prove that integration is in progress. The direction of this relationship may also show how these markets are linked together.

Section one of this paper will deal with the general data set, the variables selected and the specific problem to be discussed. A brief introduction to the method applied here will follow. In section three the univariate models which best describe the three time series chosen will be presented. Later, the relationship between them will be modeled. A concluding section will put together what are the major methodological and substantive contributions of this paper.

2. *The Series*

The historical problem where the method will now be tested concerns *when* and *how* did the market for agricultural products originate in Spain. During the 1800's, Spain suffered from economic underdevelopment which lasted until well within this century. One sign of this backwardness was the imperfect integration of her market agriculture. In the second half of the XIXth century, the development of communications — in particular, the building of the railroads — and an array of circumstances and economic measures, promoted the commercialization of agricultural products. This became a decisive step in the articulation of this market.

In previous papers, the behaviour of agricultural prices during those years was examined in order to find clues to such a transformation. The information on which these studies was based was the price series of wheat, barley, wine and olive oil for each of the fifty provinces into which Spain is divided. Monthly data for the period 1856 to 1890 were available. These were originally published in the government newspaper *La Gaceta de Madrid* (Sánchez-Albornoz, 1975b and 1979). In the last decade of this period, prices underwent a change in level and variability. In general, they tended to decrease, except in the case of wine, due to international pressure. Changes in wheat prices were specially significant since wheat was a staple for most of the population (Sánchez-Albornoz, 1975a).

In these studies an attempt was made to outline, by means of factor analysis, areas of price covariation, in order to set up economic spaces as well as their possible alteration with the passage of time (Sánchez-Albornoz, 1973 and 1974). What direction the relation between price areas took was intuitively found, but it could not be precisely determined. Also, it was not possible to establish how one series would influence another of the same economic space.

Moreover, it was not possible to model the behaviour of the time series either by areas or individually. *Grosso modo*, the series were affected by the general food crises of 1857 and 1868 (Sánchez-Albornoz, 1977, Chapters 1 through 3), although not all to the same extent. Later, some of the provinces still experienced similar local crises. The lowering of the price levels and their variability was also felt differently in each province.

Analyses of the kind which will be done here are too arduous to be performed for all sets of series, or even for the series of only one product. Therefore, in this paper we will start by considering only the monthly wheat prices in Valladolid, Zaragoza and La Coruña between July 1857 and December 1890 — 402 consecutive observations from each place (see Graphs 1-3). It must be kept in mind that these prices are averages of weekly values recorded in local markets scattered around the province. Valladolid, primarily a grain producer, was then perhaps the most important agricultural centre in northern Castille. In 1886, this province was the third largest in shipping grain by railroad and the second largest with regard to flour (Casañas, 1979). Zaragoza played a similar role for Aragon, though less grain and more flour were shipped from there. In 1886, Zaragoza got some 6,000 tons of grain but sent 20,900 tons of flour to the northeastern provinces by rail. La Coruña, a minimal producer of wheat but wide open to commerce on its seaboard, depended on coastal shipping for its supplies. It is estimated that its deficit on flour reached 7,300 tons in 1886. The contrast between Valladolid and Zaragoza, on the one hand, and La Coruña, on the other, is twofold: inland vs. coastline, producer vs. consumer.

The law prohibiting foreign grain imports was in effect until 1868 except in time of crisis and even then special authorization was required. In order to meet their deficits, non-producing areas had then to get grain from other regions in Spain. The prohibition was later replaced by a protectionist policy. Although restricting the volume of grain which could be imported, at least it allowed its entry. This change in tariff law is assumed to be responsible for the lower variability observed in wheat prices from the penultimate decade of the XIXth century on.

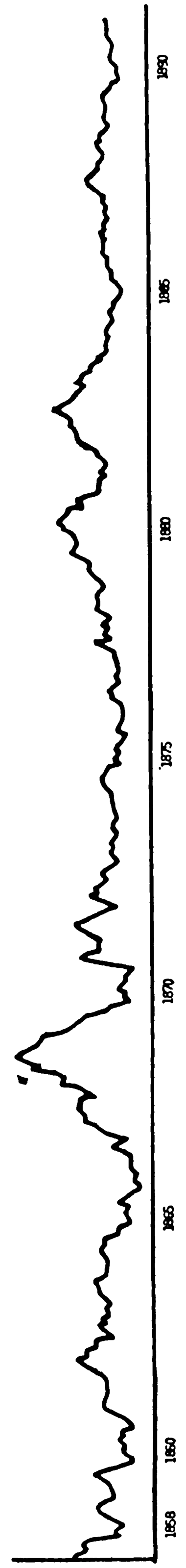
3. *Methodology*

The method applied here in the analysis of the price series was developed by G.E.P. Box and G.M. Jenkins. Their book (Box and Jenkins, 1970) is today the classical reference manual. In the last years, this method has been enriched through the contributions of Box and Tiao (1975 and 1977), Haugh and Box (1977), Granger and Newbold (1977) and Jenkins (1979), among others. A "cookbook" aiming at social scientists, historians among them, was recently published (McLeary and Jay, 1981).

The points which can be dealt with from this approach are many. In the

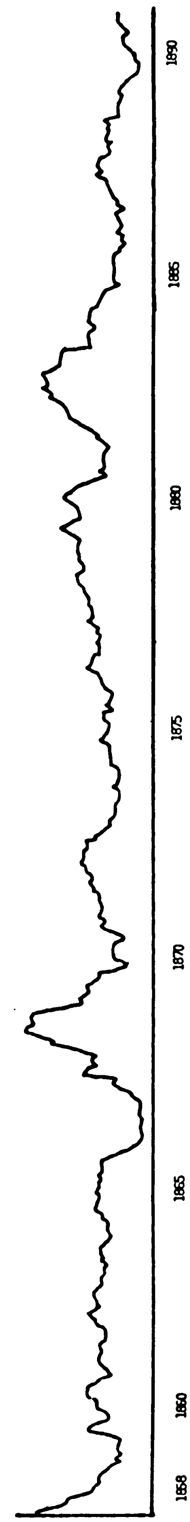
GRAPH 1

WHEAT PRICES. VALLADOLID (1857/1890)



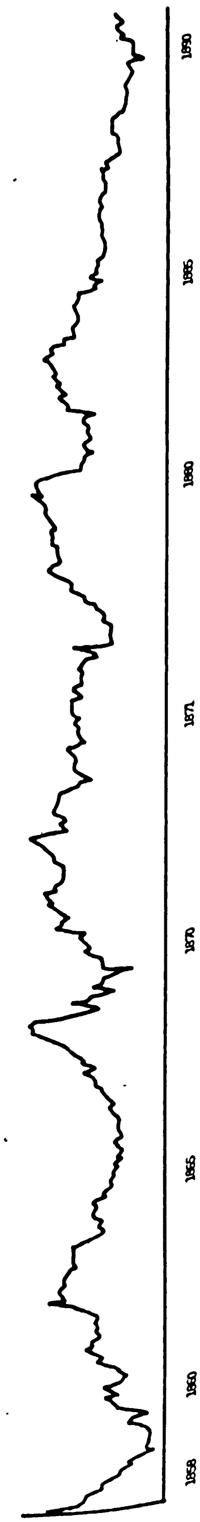
GRAPH 2

WHEAT PRICES. ZARAGOZA (1857/1890)



GRAPH 3

WHEAT PRICES. LA CORUÑA (1857/1890)



first place, let us suppose that we have a time series for which we want to build a parametric model that can describe the evolution of the series, only taking into consideration its past. The Box-Jenkins method provides a very general procedure for the construction of a *univeriate stochastic model*. This model separates that part of the series which can be explained by its history from that which depends on a random component, i.e. the residual or noise. This last factor, which by nature is unpredictable, encompasses the non-systematic effects of all the variables which have influenced the series.[1]

In order to detect the structure of the time dependence between observations of a single series, two statistical tools are used which are based on concepts that should be kept in mind. Those tools are the simple autocorrelation function (which we will denot *saf* from now on) and the partial autocorrelation function (*paf*). The *saf* provides the coefficients of simple autocorrelation (*csa*) for different lags. These coefficients measure the linear relationship which exists between the observations at the instant t and those at the instant t-1, t-2,... The *csa* of order 12, measures, for example, the relationship between the observations made every 12 months, all along the total sampling period, and therefore shows its seasonality.

The partial autocorrelation function is calculated from the coefficients of partial autocorrelation (*cpa*). The *cpa* of lag k measures the relationship between the observation at time t and those at time t-k, where one eliminates from the first the effects due to intermediate observations, t-1, ..., t-k-1. These coefficients allow one to calibrate the *differential* effect of the different lags.

If from a single series, we move on to a group of time series $\{Z_t\}$ which describe the dynamic behaviour of system, the appropiate type of model depends on the causality relation present among the series. Following Granger (1969), we will say that a series X_t has a causal effect on another series Y_t when the knowledge of the history of X_t (that is to say, the history of its values at times t-1, t-2... previous to t) allows us to forecast better values Y_t than would be possible if we did not have such information. For example, the number of vehicles is a cause of gasoline consumption in the sense that each increment in their number affects the consumption of this product. In this case the causality is one-way; i.e. from vehicles to consumption. The inverse does not necessarily have to occur. Increments in the consumption of gasoline in a certain month do not have to be associated with the growth in the number of vehicles. When the relationship between two series shows two-way causality, we can say that there is feedback. Prices of substitute goods

[1] The formula according to which the decomposition may be carried out is:

$$Z_t = f(Z_{t-1}, Z_{t-2}, Z_{t-3} ...) + a_t$$

where $f(Z_{t-1}, Z_{t-2} ...)$ represents the part of the series that can be explained by its past, and a is the residual component or noise that in statical terms is a process of normal interdependent random variables, with a mean of zero and a constant variance.

normally reveal (in a free market) feedback or two-way causality, since the variations of the price of one product influence the price of the other or vice-versa.

The most simple multivariate case of time series appears when the set of available series $\{Z_t\}$ can be divided into two groups. The first consists of only one series Y_t which is the result of the components of another group $\{X_t\}$ without feedback from Y_t to $\{X_t\}$. A model can then be built explaining the evolution of Y_t as a function of the explicative variables $\{X_t\}$. In short, that is the way that classical regression analysis works. The *transfer function* [2] models typical of the Box-Jenkins approach, while embracing the special case of regression techniques, are also formulations general enough to represent any type of dynamic behaviour of the series.

The problem turns out to be more complex when the set $\{Z_t\}$ cannot be divided into two, as in the previous case, due to the presence of feedback among the series; i.e. having interaction among all of them. In this case, we will have to resort to a *multivariate stochastic model* [3] which takes into consideration the observed dynamic interdependence.

The basic instrument to identify interdependent relationships among series is the analysis of the cross correlation functions (*ccf*). There are calculated from the coefficients of cross correlation of order k which measure the existing relationship between the observations of another series at the moments t-k. Thus, the classical Pearson correlation coefficient is a coefficient of cross correlation of order zero. The cross correlation function is calculated from the residuals of the univariate models and not from the original data of the series.[4]

Finally, the effect that specific atypical events such as political changes, new regulations or natural catastrophes, etc., have on a series or group of series, can also be modeled. It is possible to establish this effect by means of an *intervention analysis* [5] which is applicable in any of the cited cases.

---

[2] A model of the transfer function between two series Y_t and X_t (where X_t influences Y_t but not vice-versa) is a mathematical representation of how the values of the series X_t at time t affect Y_t at time t, t+1, t+2, ...

[3] A multivariate stochastic model between two series Z_t and W_t, where feedback exists, describes mathematically how the values of Z_t at the instant t influence W_t at times t, t+1, t+2, ... and how this influence rebounds upon Z_t, modifying its future values at the same time.

[4] Calculating coefficients of correlation among time series, which have not been subjected to a screening operation to eliminate trends, seasonality and other systematic components, is extremely dangerous. For example, it is possible to obtain high correlation coefficients between series that have a growth tendency even if they are independent in reality.

[5] Intervention analysis allows the calibration of what effect an isolated event of duration t has upon the series. For example, the impact which a new law regulation established at time t_0 has on the series of prices Z_t. Intervention analysis is directed

The construction of a model for a series or group of series using the Box-Jenkins method follows an interactive process which we will illustrate in the following pages. The method is gradual, in the sense that the more complex models are constructed from the results obtained from the more simple models. Every investigation of a group of series begins with the construction of a univariate model for each of the series treating them as if they were isolated.

4. *Univariate Models*

In order to illustrate how we arrived at the univariate models that we are proposing, we will explain concisely the steps followed in the analysis of the Valladolid series (average monthly prices for wheat from July 1857 to December 1890). In fact, we will concentrate upon the period between July 1857 and July 1871. For the sake of brevity, we will exclude from this paper the similar analyses done for the Zaragoza and La Coruña series.[6] We will, however, include those results which are useful to characterize them or to compare them against Valladolid. The choice of Valladolid for this analysis was due to the fact that (as will be shown) prices here influenced those in the other two provinces and *not* vice-versa. This unidirectional movement makes Valladolid the logical choice for analysis.

It is convenient to divide a series as long as that of Valladolid (402 cases) into small sections and study each separately. If any changes in the structure of the time dependence of the series exist, they are more easily detected this way; that is, in the way in which the prices of one month are related to the prices of the previous months. The first period ranges from July 1857 to June 1871 (14 years or 168 cases); the second period, from July 1871 to June 1885 (again 14 years or 168 cases), and the third from July 1885 to December 1890 (5 years and 6 months, or 66 cases). The division into periods is arbitrary but it corresponds *grosso modo* to the division into an old regime, a transition period, and a new period of price behaviour, respectively.

To begin with, the diagram of the series (Graph 1, based on original data) suggests that the variability is larger when the level of prices is high. Therefore, the first step which must be taken is to transform the data so that they will have a constant variance.[7]

---

towards estimating the effects of this variation of the external conditions on the values in the series at time t_o, t_o+1, t_o+2, ... The effect each time can be different and its magnitude and direction are identified and estimated empirically.

[6] The interested reader can consult Peña (1979) for the details pertaining to the modeling of these series.

[7] The condition of constant variance is one of the basic hypothesis in the formulation of linear models. When this does not occur, the data will be transformed (for example, by taking logs or square roots, etc.) so that the new transformed data will possess this property.

From studying the range/mean diagram,[8] it is clear that it will be more convenient to work with logarithms. This transformation makes sense, since the changes will then be expressed in a form proportional to the level of the prices instead of being additive.

It is clear that the level of prices varies with time, as can be seen by inspection of Graph 1. In order to eliminate the trend as well as the fluctuations, it will be necessary to differentiate.[9] After doing this a new series emerges in which we have eliminated a large portion of the fluctuations, and since it is the differential of a logarithmic series, it describes the evolution of the rate of monthly increment of the prices. In this new series we find that some values are extreme, i.e. they are either very large or very small.[10] Is this discrepancy perhaps a result of errors in the original data?

In order to verify this point, we compared the Valladolid series with those of Zamora and Palencia, neighbouring provinces to the west and the north respectively, but economically analogous. In general, all three series varied similarly. However, when the extreme values appeared, it was not so. While Zamora and Palencia agreed once again, Valladolid showed untimely values with respect to its neighbours, in addition to being atypical to its own history. Thus, the suspicion that the data are indeed in error is doubly justified.

In order to correct those errors, the most precise statistical procedure is interpolation by means of the *intervention analysis*.[11] If the Valladolid series

---

[8] In order to determine what transformation needs to be applied, one should study how the variability of the series (measured by the range) changes with regard to the trend (measured by the average of the series in each subinterval) for different periods. The form of this relationship indicates what type of transformation is needed so that the variability will be constant during the whole period. For example, if the relationship between the range and the mean is linear, we will take logarithms.

[9] Differentiating the series is a procedure done to eliminate the trend. This allows not only the elimination of polynomial deterministic tendencies (linear, quadratic, etc.), but also random variations at the level of the series due to, for example, an economic crisis, such as was this case. Once these aspects, detectable by inspection, are eliminated, we can carry out a more rigorous analysis of the forms of the time dependence that are more complex.

[10] We will consider values to be exceptional when they lie at more than 3 times the typical deviation from the mean of the series.

[11] Intervention analysis, to which we have referred in note 4, is also useful for obtaining the optimum interpolation when we suspect there are errors in the data. If the apparent error occurs in observation Z_t, we establish a model that reflects a change in the volue of Z_t but has no other effect on the rest of the series by using:

$$z_t = w_o I_t + z_t^{\star}$$

where Z_t is the original series, I_t is a series with a value 1 at time t and zero every-had been the only one available, there would be no other choice. However, since prices for Zamora and Palencia were available, it was possible to find values for Valladolid in line with the behaviour of the series. This rough kind of interpolation makes sense. By comparing the results obtained both by the simpler and the most sophisticated procedures, we found that they were equivalent. In subsequent analyses, such as those done for Zaragoza and La Coruña, we therefore only resorted to the simpler procedure; that is, we interpolated values in agreement to the registered values in similar or neighbouring districts.

The analysis of *saf* and *paf* for different degrees of differentiation (see Graph 4 for $\Delta \ln v_t$) leads to the models shown in Table 1. The model which best represents this span of the series is number 1. This one shows the smallest variance and also has the smallest value of Box-Pierce Statistics[12] (Box and Pierce, 1970). Written in a more developed from, this model is

$$V_t = V_{t-1} \left(\frac{V_{t-1}}{V_{t-2}}\right)^{0.25} u_t$$

where V_t is the price of wheat in Valladolid at time t and u_t is a random variable with mean 1 and a log-normal distribution. Therefore, the price of wheat in Valladolid in any month of this period is the same (excepting random

TABLE 1

SUMMARY OF THE VALLADOLID MODELS

	Model	σ_a^2	Q (39-NP)
1	$(1-.25\,B)\,\Delta\,Ln\,V_t = a_t$ (.07)	0,00387	32,7
2	$(1-.29\,B)\,\Delta\Delta_{12}\,Ln\,V_t = (1-.89\,B^{12})\,a_t$ (.07) (.02)	0,00415	41
3	$(1-.2\,B)\,\Delta^2\,Ln\,V_t = (1-.97\,B)\,a_t$ (.07) (.02)	0,00400	36
4	$(1-.22\,B)\,\Delta^2\Delta_{12}\,Ln\,V_t = (1-.89\,B)\,(1-.83\,B^{22})\,a_t$ (.09) (.05) (.04)	0,00164	44,35

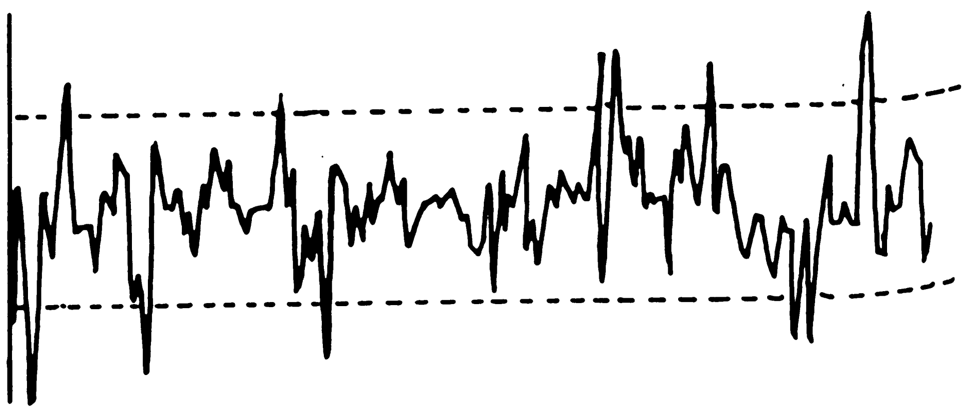
where else, and $Z_t^\star$ is the corrected series. It is clear that Z_t $Z_{t'}$, in all periods except at time t', where

$$z_{t'}^\star = z_{t'} + w_o$$

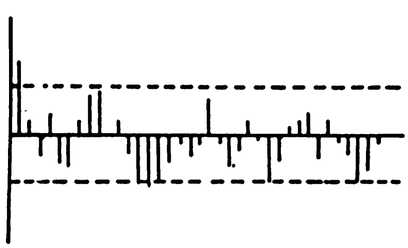
and for that reason the value w_o measures the magnitude of the error at that instant.

12 The Box-Pierce test is a measure of the fitness of the model. It allows one to check the hypothesis that the residuals of the model built are independent from each other and, therefore, no longer contain any information of the history of the series. In all the tables, this test is represented by the symbol Q (.) where the number in parenthesis indicates the number of degrees of freedom.

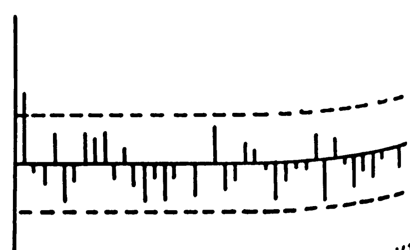
GRAPH 4



Ln V1



autocorrelation function of ∇LnV1



Partial autocorrelation of ∇LnV1

variations) as the price of the previous month multiplied by the rate of change of the previous month raised to 0.25.

We have also constructed univariate models for the other two periods of the Valladolid series as well as a model for the overall period. Table 2 summarizes the best model for each part and the general model finally obtained. In this table, the numbers 1, 2 or 3 attached to V, Z or C indicate the period to which the corresponding model applies.

From the univariate analysis carried out, several conclusions can be reached. Some refer to the general application of the method to historical research, while others refer to the information alone.

The construction of a univariate model for a time series is first of all the most efficient way to detect errors in the data and to refine the series before proceeding with a more complex analysis (especially when it involves interdependence of variables). Inappropiate values were deleted from our series this way.

The univariate model also permits one to locate values that are not erroneous, but are still not typical, giving the historian clues towards punctual events which may have otherwise gone unnoticed. The investigation of the circumstances giving rise to these values would have required extensive investigation of qualitative sources. Because of the lack of time, this research has not been done.

The model rigorously describes the interdependence present among the observations made as a function of time. In short, it provides information not always available about the characteristics of the period studied. For example, the seasonality (or lack of it) in the agricultural prices detected by the model suggests how inefficient (or efficient) the trade circuits may have been.

The univariate model constitutes, on the other hand, only the first step in the process of modeling the causal dependence between time series, such as will be done in the following section.

From the Valladolid series (Table 2), the models studied suggest that the structure of time dependence of the price in one month with respect to the previous months remains constant during the whole period. In all the cases, the dependence is autoregressive, of the first order, and the coefficient is not significantly different from one period to another, with a level of significance of $\alpha = 0,05$.

Secondly, the variance of the residuals in all these models does decrease significantly from one period to another when we evaluate it through an F-test. This implies that the exogenous factors not included in the series history which conditioned the price evolution in Valladolid had a smaller influence towards the end of the XIXth century than at the beginning. The net result is a market progressively more predictable, in relative terms, and less subject to sudden random variations.

In the third place, the Valladolid series does not present any consistent

TABLE 2

SUMMARY OF UNIVARIATE MODELS

	Name		Model	Variance	(Q (39)
Valladolid	$V1_t$	Period 57 - 71	$(1-.25\,B)\,\Delta\,Ln\ V1_t = a_t$ (.07)	0,00387	32,7
	$V2_t$	Period 71 - 85	$(1-.15\,B)\,\Delta\,Ln\ V2_t = a_t$ (.07)	0,00138	27,4
	$V3_t$	Period 85 - 90	$(1-.38\,B)\,\Delta\,Ln\ V3_t = a_t$ (.05)	0,000456	30,7
	V_t	TOTAL 57 - 90	$(1-.24\,B)\,\Delta\,Ln\ V_t = a_t$ (.05)	0,00227	33,1
Zaragoza	$Z1_t$	Period 57 - 71	$(1-.36B)\,\Delta\,Ln\ Z1_t = a_t$ * (.07)	0,00155	38,7
	$Z2_t$	Period 71 - 85	$(1-.21B)\,\Delta\,Ln\ Z2_t = a_t$ (.07)	0,00086	25
	$Z3_t$	Period 85 - 90	$\Delta\,Ln\ Z3_t = a_t$	0,00072	37
	Z_t	TOTAL 57 - 90	$(1-.26\,B)\,\Delta\,Ln\ Z_t = a_t$ (.07)	0,00136	12,4 **
Coruña	$C1_t$	Period 57 - 71	$\Delta\,Ln\ C1_t = a_t$	0,00194	28,6
	$C2_t$	Period 71 - 85	$\Delta\,Ln\ C2_t = a_t$	0,00068	25
	$C3_t$	Period 85 - 90	$(1+.29B)\Delta\,Ln\ C3_t = a_t$ *** (.12)	0,00041	23,6
	C_t	TOTAL 71 - 90	$\Delta\,Ln\ C_t = a_t$	0,00124	36,5

Notes:

* The model of Zaragoza includes three significant interventions needed to get a proper representation.

** The value of Q is estimated with 11 degrees of freedom.

*** The negative AR term disappears when two interventions are applied.

evidence of seasonal factors, such as those we had assumed in a previous analysis (Sánchez-Albornoz, 1975b: 33). Since at no time did the government regulate domestic prices, the group probably responsible for the disappearance of seasonality in the market prices were the merchants who stored the wheat after the harvest, then gradually brought it out for sale. The differences in the prices during the year were not sufficiently attractive to make speculation profitable. Therefore, speculation seems to have been put off until the periods of food crises arrived which happened every ten years.

These conclusions are only partially applicable to Zaragoza. The disappearance of the autoregressive term in the third period, coupled with a significant decrease of the AR coefficient in the first, indicate that the market in Zaragoza evolved in such a direction so as to become increasingly efficient, a weak efficiency however, it is meant, since the only information available is that provided by the history of the series (Table 3). The reduction of the effects produced by random factors and the absence of seasonality are just as evident in Zaragoza as they are in Valladolid.

TABLE 3

Period	Model	Gain	Var. Res.
7/57 - 6/73	$\text{Ln } Z_t = \frac{\underset{(.04)}{.21}}{\underset{(.09)}{1-.67B}} \text{Ln } V_t + \frac{a_t}{\underset{(.07)}{(1-.16B)}\,\Delta}$	.62 (.14)	0,00144
6/73 - 12/90	$\text{Ln } Z_t = \frac{\underset{(.05)}{.25}}{\underset{(.09)}{(1-.67B)}} \text{ Ln } V_t + \frac{a_t}{\Delta}$	.76 (.14)	0,000697

With respect to La Coruña, the random-walk structure shown during the whole period evokes an efficient market, in a weak sense also (Table 4).[13] The price series there is similar in structure to the previous ones.

TABLE 4

Period	Model	Gain	Var. Res.
7/57 a 6/73	$\text{Ln } C_t = (\underset{(.05)}{.90} + \underset{(.05)}{.09}\text{ B}) \text{ Ln } V_t + \frac{a_t}{\Delta}$	.18 (.06)	0,00165
7/73 a 12/90	$\text{Ln } C_t = (\underset{(.06)}{.09} + \underset{(.07)}{.16}\text{ B}) \text{ Ln } V_t + \frac{a_t}{\Delta}$	.25 (.08)	0,000898

[13] The theoretical mathematical representation of an efficient market is precisely the one found in La Coruña, frequently detected also in many stock markets. This implies that there was a free competition and availability of information. Therefore, the best forecast for the value of the prices in a period is obtained from the value observed on the immediately preceding period. The dependence of the Valladolid and Zaragoza series is due to the presence of an autoregressive parameter which, as it is shown in Table 2, turns out to be statistically significant.

In sum, we can conclude that:

a) The dependency structure of wheat prices is stable in Valladolid, a producing area, and in La Coruña, a consuming area, during the second half of the XIX century. It changes slowly in Zaragoza. The dependence there was at the beginning similar to that of Valladolid, but gradually takes up the structure of La Coruña.

b) The variability of the random effect not modeled decreases slowly in all three series.

c) There is no evidence of a seasonal dependence in the whole period.

5. *Modeling of the dependence*

The original data (Graphs 1 through 3) suggest that there is a strong dependence relationship between the three series. In all of them, substantial increases in prices are found in 1857 and in 1868, which were years of general food crises. In contrast, the price of wheat there was exceptionally low in 1866 and relatively stable for the rest of the period. Valladolid records the lowest price of the three series in the period of time analyzed (1857 through 1890) with a mean of 18.63 pesetas per hectolitre, a standard deviation of 3.38, and a coefficient of variation of 0.179. Zaragoza is next lowest (a price between one to five per cent above that in Valladolid) with a mean of 19.45, a standard deviation of 3.09, and a coefficient of variation of 0.159. Finally, La Coruña is highest (a price 30% above that of the first), with a mean of 24.97, a standard deviation of 3.16, and a coefficient of variation of 0,127.

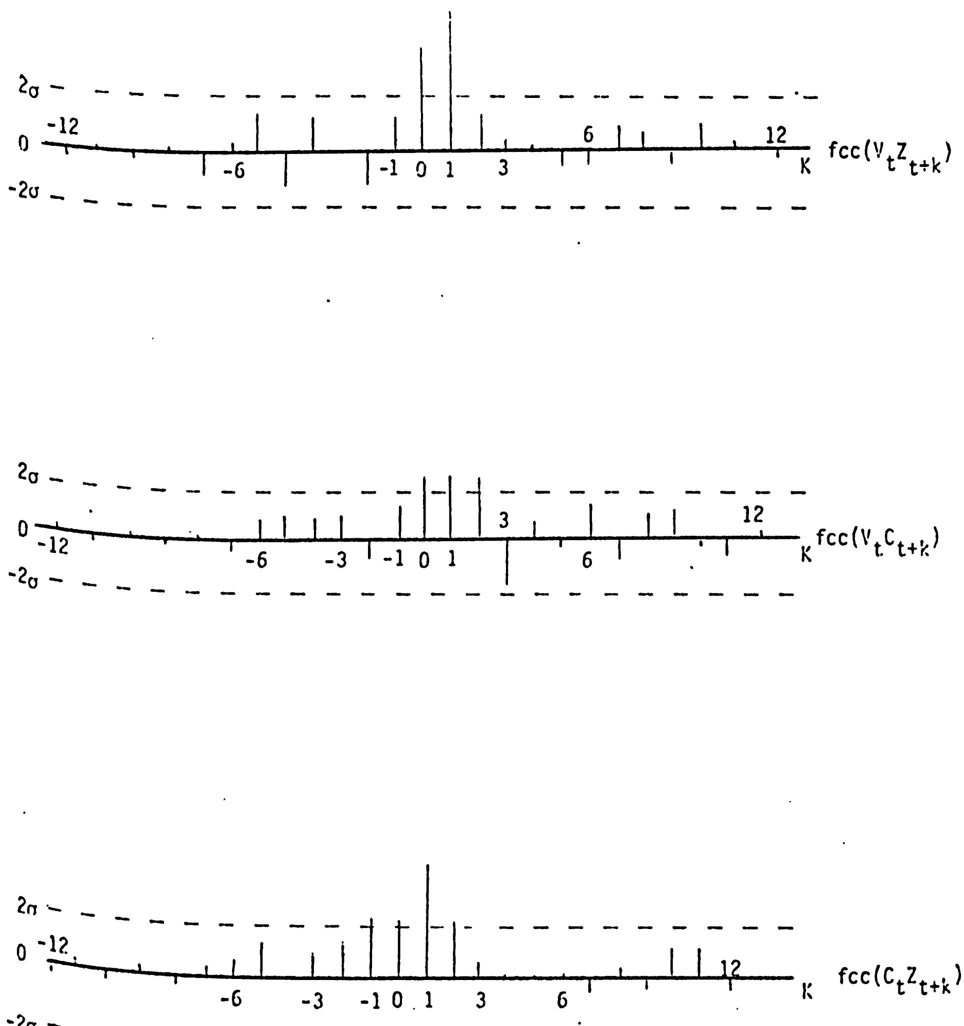
In order to identify what causal relationship there is between the prices of these provinces, we need to estimate the *ccf* between the residuals of the univariate models of the three series. Previous univariate analyses have shown that only the time series of Valladolid and Coruña display a stable correlation structure, while that of Zaragoza exhibits instead a vanishing autoregressive term. The Zaragoza series should thus be split into several segments in order to find periods during which the data exhibit stable correlation functions. A division in two parts almost equal in length met our needs. The first one goes from July 1857 to June 1873 and the second one from July 1873 to December 1890. The autocorrelation appears as it is in the two other provincial series.

The *ccf* between the residuals of the univariate models for the three series during the first period only are displayed in Graph 5. The first function corresponds to the relationship between Valladolid and Zaragoza. The positive value at zero indicates a significant instantaneous relationship and the high value for lag 1 indicates that the monthly wheat price in Valladolid influences the price of wheat in Zaragoza the next month. The coefficients to the right of zero (positive lags) indicate causality directions from Valladolid towards Zaragoza, while the coefficients to the left of zero reflect the influence

GRAPH 5

CROSS CORRELATION FUNCTIONS

(1857 - 1873)



that Zaratoga exerts on Valladolid. Since the latter ones are not significant, we can conclude that in this period the direction of causality is from Valladolid toward Zaragoza, with virtually no feedback. Moreover, the relationship works in a very short term, as would be expected.

The first curve in Graph 6 represent the relationship between the same two series in the second period. The graph reveals once again the one-way causality from Valladolid to Zaragoza. The same features are clearly visible in Graph 7, which contain data for the whole period. The curve suggest that the relation structure is possibly stronger in the second period, an aspect which we will confirm by constructing a model for this dependence.

If we go on to study the relationship between Valladolid and La Coruña, we will find that Valladolid exerts influence over la Coruña, but not vice-versa. However, it has been found that there exists a bivariate relationship between la Coruña and Zaragoza in both of the subintervals analyzed as well as for the entire period. This conclusion is made apparent by the presence of significant coefficients of cross correlation on both sides of the zero lag.

Table 3 shows how stable the price response is in Zaragoza vis-a-vis the price variation in Valladolid. Whenever the price increases one per cent in Valladolid during the period 1857-1873, it moves subsequently up 0.62 per cent in Zaragoza. This response is not transmitted punctually, but spreads over time. During the first three months after the price changed in Valladolid, 76 per cent of the total effect, that is of the 0.62 rise already mentioned, was transmitted to the prices in Zaragoza. Three months later, that is in one semester, up to 91 per dent of the same effect was felt in Zaragoza. Most of the response came therefore quite early. Finally, that part of the price variation in Zaragoza which remains unexplained by the behaviour of the Valladolid prices also exhibits a low autoregressive structure.

During the second time period, the transfer function moves up to 0.76, meaning that a ten per cent increase in the price of Valladolid reaches over the long term a 7.6 per cent rise in Zaragoza. The shape of the response and the lag structure are similar to that of the first period. The autoregressive term is however no longer present in the noise.

Comparing both periods we may now conclude that the relationship was quite stable during the entire 34 year period. The influence of Valladolid over Zaragoza seems to be growing slowly, although such an increment is statistically not significant.

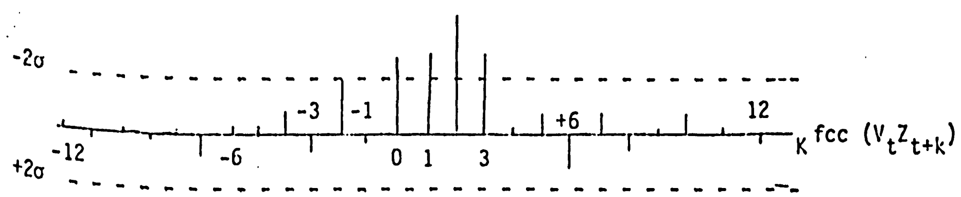
As far as Coruña is concerned, the Table 4 shows that, during the period 1857 to 1873, any 10 per cent increase in the wheat price in Valladolid only leads to a 1.8 per cent rise in the long run. During the second period, it goes up to 2.5 per cent. In both segments, there is a single lag in the response, which implies that the relationship occurred in the very short term.

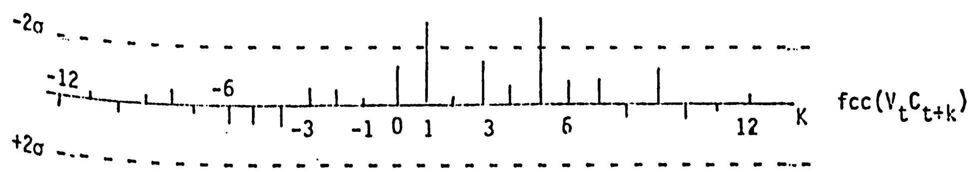
To sum up, the shape of the response to the price variation in Valladolid appears stable in Zaragoza and Coruña during the entire period considered.

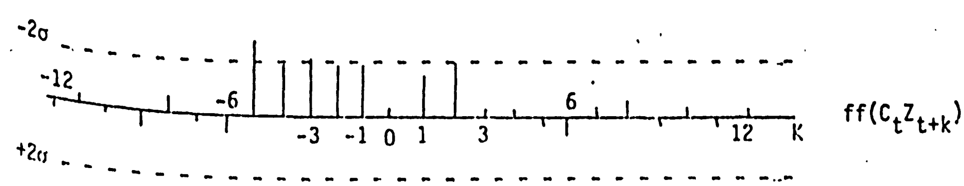
GRAPH 6

CROSS CORRELATION FUNCTIONS

(1873 - 1890)



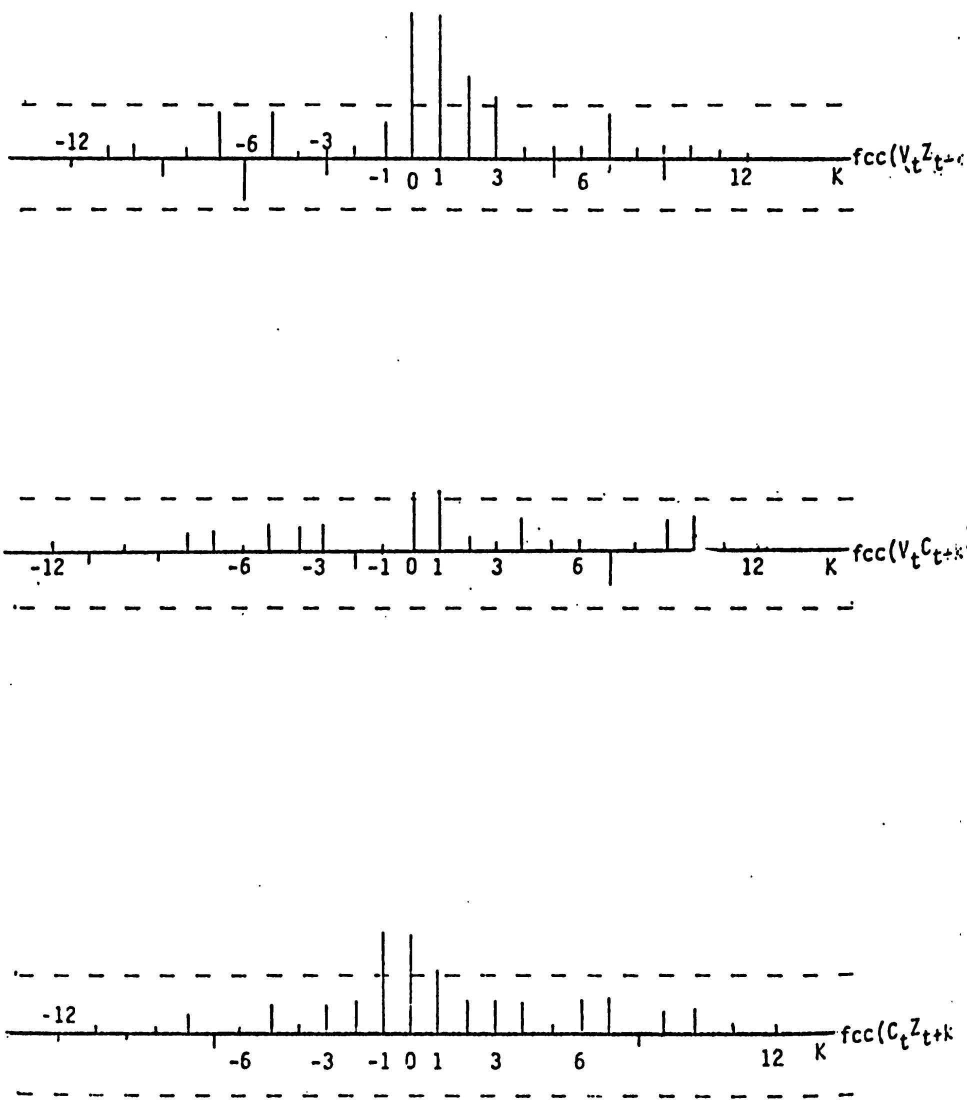




GRAPH 7

CROSS CORRELATION FUNCTIONS

(1857-1890)



The relationship also appears intensifying. The magnitude of the response was certainly stronger and lasted longer in Zaragoza than in Coruña.

6. *Conclusions*

The conclusions can be divided into two groups: the methodological and the substantive ones. First, this approach allowed us to detect and correct errors in the data. Without such a screening, we hardly could have reached consistent results in the analysis of the time series. The construction of a univariate model for each series helped also to identify changes in structure, to appraise the relative importance of the random variables with regard to the systems, to check the presence of seasonality and to test the relative efficiency of the market. The analysis of dependence between series not only allowed us to detect the direction of causality, but also to build a transfer model which reflects the magnitude and the extent of the relationships between series. The application of the Box-Jenkins approach, therefore, offers a general frame for treating time series rigorously.

Concerning results, the series of the three provincial markets present common features: the prices lack seasonality; they experience first sharp fluctuations still related to food crises, and finally reduce their variability. This slowdown indicates, among other things, that the crises disappeared.

All three markets exhibit strong differences. The level of prices in Valladolid and Zaragoza were similar. In La Coruña they were higher. In all three cases, whenever the level was high, there was a lower variability. In the relationship between provinces, Valladolid prevails. This is the source of a unidirectional motion with no feedback. The time lag between series seldom goes over a few months. Price variations in the most inner province are thus quickly transmitted to the other two.

Insofar as these three series allow us to speculate about the national market, it can be said that it appears to lag in development. The sharp regional contrasts and the steep fluctuations evident here no longer occurred then in the agrarian economy of the rest of Europe. This backwardness does not necessarily imply that the old agrarian system remained in Spain unchanged. The lack of seasonality already suggest a capitalistic transformation in the grain trade. Operating on a higher level, this commerce no longer speculated over price differences within the year, but rather on a mid-term basis, aiming at eventual food crises.

The fact that prices were passed along from one province to another does not imply that grains were necessarily moving among them; rather, it only means that information was circulated, which eventually was able to generate flows of goods. The price in La Coruña depended upon the price in Valladolid not because she regularly supplied wheat to the Galician province, but rather because La Coruña could expect to receive it from Valladolid or because her customary sources of supply were also influenced by the Castilian market. The information worked on expectations. In a short time, one market could react to the stimulus coming from the other. The fact that the information was circulated and that it eventually induced the flow of goods denotes a significant degree of association. It is apparent that modern tendencies were intermixing with the old ones.

The tariff forced, in our opinion, the merging of old and new tendencies. The prohibition of grain imports, instituted in 1820, brought about a greater interdependence between regions — a requisite for market integration. However, the process was halted half-way — due in part to the inability to develop transportation fast enough, but also because of the import prohibition itself. Actually, the periodic fluctuations did not disappear but instead increased up to the 1870's. Spain then barely produced enough wheat to feed its inhabitants, and nevertheless was prevented from importing it, except in case of extreme scarcity. Speculation in time of crises continued to be a source of substantial, though irregular, profits. Thus, the increase in the interdependence between provinces did not level off prices or made them more regular.

Stabilization and leveling happened later, during the 80's. This was a consequence of the development of transportation (railroads in particular), and the substitution of the prohibitionist by a protectionist policy. The structure of the dependence between the series examined indicates, however, under what kind of restrictions this change took place. In line with the aims of protectionism, the producing zones continued to set the price of grain. The relationship between provincial markets was seldom bidirectional. In spite of its capitalistic orientation and its exposure to international movements, the national market did not become open to reciprocity.

---

[14] The models for the 73-79 period are:

$$\text{Ln } C_t = \frac{\underset{(0.02)}{0.09}}{1-\underset{(0.06)}{0.86}\,B}\ \text{Ln } V_t + \frac{a_t}{\Delta} \qquad \text{Gain: } \underset{(0.22)}{0.64}$$

$$\text{Ln } Z_t = \frac{\underset{(0.04)}{0.24}}{1-\underset{(0.08)}{0.68}\,B}\ \text{Ln } V_t + \frac{a_t}{\Delta} \qquad \text{Gain: } \underset{(0.15)}{0.74}$$

REFERENCES

BOX, G.E.P. and G.M. JENKINS (1970), *Time Series Analysis. Forecasting and Control*, San Francisco: Holden-Day (2d ed. 1976).

BOX, G.E.P. and G.M. JENKINS (1973), " Somme comments on a Paper by Chatfield and Prothero and on a Review by Kendall ", *Journal of the Royal Statistical Society* A, 135 (3): 337-344.

BOX, G.E. P. and G.C. TIAO (1975), " Intervention Analysis with Applications to Economic and Environmental Problems ", *Journal of the American Statistical Association*, 70 (349): 70-79.

BOX, G.E. P. and G.C. TIAO (1977), " A Canonical Analysis of Multiple Time Series ", *Biometrika*, 64 (2): 355-365.

CASAÑAS, M. (1979), " La comercialització des productes agraris l'any 1886. Reflexions sobre el mercat nacional espanyol ", *Recerques*, 9: 117-132.

HAUGH, L.D. and G.E.P. BOX (1977), " Identification of Dynamic Regression (Distributed Lag) Models Connecting Two Time Series ", *Journal of the American Statistical Association*, 72 (357): 121-130.

GRANGER, C.W.J. (1969), " Investigating Causal Relations by Econometric Models and Cross-Spectral Methods ", *Econometrica*, 37: 424-438.

GRANGER, C.W.J. and P. NEWBOLD (1977), *Forecasting Economic Time Series*, New York: Academic Press.

JENKINS, G. (1979), *Practical Experiences with Forecasting*, G.J.P. Editors.

MC CLEARY R., R.A. HAY, Jr. et al. (1980), *Applied Time Series Analysis for the Social Sciences*, Beverly Hills: Sage Publications.

PEÑA, D. (1979), *Modelización de la dependencia dinámica entre mercados agrícolas en España durante la segunda mitad del siglo XIX*, Documento de Trabajo 10. BJ. 79, Madrid: Escuela de Organización Industrial.

PEÑA, D. and N. SÁNCHEZ-ALBORNOZ (1983), *Dependencia dinámica entre precios agrícolas. El trigo en España, 1857-1890. Un estudio empírico*, Madrid, Banco de España. Servicio de Estudios. Estudios de Historia Económica, nº 8.

SÁNCHEZ-ALBORNOZ, N. (1973), *Determinig Economic Regions from Time Series Data. A Factor Analysis of the 19th Century Wheat Prices*, Occasional Papers num. 1, Ibero American Language and Area Center, New York University.

SÁNCHEZ-ALBORNOZ, N. (1974), " Congruence among Spanish Economic Regions in the Nineteenth Century, *Journal of European Economic History*, 3: 725-745.

SÁNCHEZ-ALBORNOZ, N. (1975 a), " La integración del mercado nacional. España e Italia ", in *Jalones en la modernización de España*, Barcelona; Ariel: 101-146.

SÁNCHEZ-ALBORNOZ, N. (1975 b), *Los precios agrícolas durante la segunda mitad del siglo XIX*, Vol. I, *Trigo, y cebada*, Madrid: Banco de España.

SÁNCHEZ-ALBORNOZ, N. (1977), *España hace un siglo: una economía dual*, 2d ed., Madrid: Alianza Editorial.

SÁNCHEZ-ALBORNOZ, N. and T. CARNERO (1981), *Los precios agrícolas durante la segunda mitad del siglo XIX*, Vol. III, *Vino y aceite*, Madrid: Banco de España.