NON-PARAMETRIC MODELLING OF SPATIAL PRICE RELATIONSHIPS

March 2006

We apply nonparametric methods to a consideration of price transmission processes within U.S. egg markets at the turn of the XIX century. Gordon (2000) labelled this as an era of "Great Inventions" which contributed to the subsequent years of significant productivity growth and noted that the development of mechanical refrigeration and transportation technologies played an important role in this growth. Our models present certain advantages relative to parametric models traditionally employed in price transmission analyses. We compare results derived from local polynomial modelling to those obtained using alternative nonlinear threshold models. Both techniques suggest that U.S. egg markets were interrelated at the turn of the XIX century. However, nonparametric techniques often suggest a higher degree of price transmission than that implied by threshold models. Results also suggest that threshold models may have difficulties in adequately capturing price relationship dynamics, especially when these are of a highly nonlinear nature.

Introduction

Spatial price analysis uses different statistical techniques to evaluate price relationships between spatially distant markets in order to shed light into the patterns and workings of marketing. Specifically, the literature on spatial price transmission provides several empirical methods that evaluate the extent to which price signals are transmitted across space. As Fackler and Goodwin (2001) note, most tests of spatial price relationships can be economically justified through the point-location model. This well-known, simple and highly stylized economic framework, first discussed by Enke (1951) and Samuelson (1952), represents the equilibrium conditions inherent in the Law of One Price (LOP). Though the terminology has been loosely applied, the (weak version) of the LOP is an equilibrium concept involving that, in a well-functioning market, arbitrage activities will preclude prices of a homogeneous good at two separate locations to differ by an amount greater than transfer costs. This is equivalent to the spatial arbitrage condition that implies that the actions of arbitrageurs will move the spatial price spread toward the costs of transferring goods from one market to another.[1] Moreover, it is important to note that, in order for regional price adjustments to take place, explicit trade between a pair of markets may not be necessary. For example, one may find a situation in which two agents located in two spatially different markets sell into a third common market. One would expect that the actions of buyers would result in equilibrating pressures that should equalize prices without any direct flow of commodities existing between the pair of markets. To the extent that these equilibrating pressures take place, improved information on prices should contribute to spatial price convergence. Of course, such

[1] It should be noted here that other concepts such as market integration or efficiency have also been utilized to characterize situations in which price shocks are transmitted across spatially separate markets (see Fackler and Goodwin, 2001, for further detail).

price-sensitive mechanism should be based upon the existence of adequate transportation and handling facilities that, even if trade flows are limited, create the potential for relatively relevant commodity movements.

The objective of this article is to assess the spatial arbitrage condition by applying nonparametric techniques. We apply these techniques to a sample of U.S. egg prices at the turn of the nineteenth century (October 1881 to October 1911). From a methodological point of view, the paper explicitly considers the nonlinear nature of spatial price behaviour. Nonlinear modelling of price adjustments assumes that there exist different states of nature or regimes and that the regime occurring at a certain point in time determines the dynamic price behaviour. While several econometric procedures have been devised to capture nonlinear price relationships, threshold models are one of the most widely used. Most of these models view price behaviour as a regime-switching process with a band of inaction that guarantees that price differentials will revert to a certain range. Within each regime, price adjustments are assumed to be linear and a discrete jump at a threshold value is supposed to lead regime-switching. In this line of research, Chavas and Metha (2004) have proposed an extended error correction model that allows price dynamics to differ across regimes. While these authors treat regime-switching as exogenous, more general models of asymmetry incorporate this issue as endogenous. These models include threshold autoregressive (TAR) models (Obstfeld and Taylor, 1997; Ghoshray, 2002), or threshold vector error correction (TVECM) models (Lo and Zivot, 2001; Goodwin and Piggott, 2001). All these approaches have in common their parametric nature. Parametric approaches to model price relationships require the formulation of assumptions about the true nature of price behaviour that may prove to be too restrictive or unrealistic.

Contrary to parametric models, nonparametric techniques such as local polynomial modelling (see Fan and Gijbels, 1996, chapter 3) do not require any assumption about the functional form characterizing price behaviour. In being data driven methods, it is the data that informs and determines the shape of the relationships among the variables studied. Local estimators offer yet another advantage over parametric methods. While robust estimation of threshold models requires stationary thresholds, local estimators do not assume constant transactions costs bands. Hence, the use of nonparametric techniques should be preferred when changes in thresholds over the period of analysis are suspected. Up to date, the use of nonparametric techniques to study nonlinear aspects of price transmission has been very limited. Mancuso, Goodwin and Grennes (2003) assessed capital market integration by using local linear regression (LLR) models. Though price transmission between spatially separate food markets has been an important research topic, no analysis has attempted to address spatial food price relationships using nonparametric techniques.[2] The use of nonparametric methods is the main novelty of the paper and represents a contribution to the literature.

Our empirical study focuses on the U.S. egg market at the turn of the nineteenth century and is considered of interest for two main reasons. First, because technological developments during the second industrial revolution benefited the exchange of a wide variety of goods and contributed to increased economic interdependence and spatial price convergence (Holmes, 1913; Zapoleon, 1931). Perishable commodities were amongst the most favoured.[3] Specifically, the adoption of mechanical refrigeration by

[2] An exception is the paper by Barrett and Li (2002) that includes a semi-parametric test for spatial market equilibrium.

[3] As it is well known, the second industrial revolution, which can be roughly dated from 1860 to 1900 comprises a series of inventions such as the electricity, the internal combustion engine, communication innovations, and so on (Gordon, 2000). All these inventions involved substantial changes in the patterns

the 1890s represented a vital element to the development of food production and distribution industries. Though previous to the 1890s the regional exchange of perishable foods was, to a certain extent, promoted by the investment in ice-refrigerated cars by some U.S. meatpackers, refrigeration based on ice played a relatively small role in food production and distribution. As a result, prior to the 1890s there was little cold storage of perishables (Goodwin, Grennes and Craig 2002). Mechanical refrigeration had a relevant impact on both the geography and economics of the egg industry (Anderson, 1953). Hence, the analysis of the economic impacts of food preservation mechanisms provides an interesting contribution to the literature. However, econometric estimations that reflect price linkages in food markets during the period of relevant changes described above are very scarce. Goodwin, Grennes and Craig (2002) and Serra and Goodwin (2004) constitute two exceptions.

Second, in spite of the technical developments already mentioned, the egg manufacturing industry suffered from many shortcomings during the period of analysis. A first problem concerned egg handling (see the next section for more detail). Poor handling techniques would result in a high proportion of eggs not reaching city markets and in important economic losses (Philips, 1909). A second problem was the slow and difficult development of the egg manufacturing industry resulting in small scale operations during the period of analysis. These shortcomings are likely to have limited spatial arbitrage opportunities relative to other commodities. Since trade in manufactured products such as frozen or dried eggs can bring the entire industry into a more competitive environment, it is interesting to see whether a slow industrial development limited regional price convergence for eggs during the period of study.

---

of trade. Exchanges in small cities began to disappear in favour of centralized warehouses close to large centres of population that would concentrate and then distribute commodities in their areas of operation.

Our article is organized as follows. We devote the next section to a description of the technical change and the U.S. egg markets at the turn of the nineteenth century. The econometric methods section discusses the econometric methodology employed. The empirical application is presented in the fourth section. We close the article with a brief summary of the analysis and the derived conclusions.

Technical Change and the U.S. Egg Markets at the Turn of the Nineteenth Century

The last decades of the nineteenth century have been identified as an era of increased development and economic integration (see O'Rourke and Williamson, 1999). Improvements in transportation, refrigeration, and communication mechanisms made it easier for buyers and sellers to contact each other, yielding a higher level of market integration and reduced regional price spreads (Zapoleon, 1931). Economic integration resulted in an increase in trade, a more efficient use of resources and growth in productivity and overall production (Gordon, 2000; Zapoleon, 1931). Economic integration was not limited to capital (Sylla, 1969) or labour markets (Williamson, 1995). It also characterized the evolution of agricultural commodity markets such as grain markets (O'Rourke, 1997; Zapoleon, 1931). Among the improvements mentioned above, the adoption of refrigeration technologies constituted a key factor in the development of food production and distribution and allowed food trade to occupy prominent positions in international trade. Though a considerable amount of food was traded in the form of dried, canned, or preserved foodstuffs, the relevance of fresh produce trade, aided by progress in cold storage and refrigerated shipments, should not be underestimated (the U.S., for example, imported substantial quantities of shell eggs from China at that time) (Zapoleon, 1931; Koudele and Heinsohn, 1960).

Between 1830 and the Civil War, the American urban diet improved as a result of an increase in the use of fresh foods to the detriment of certain traditional items such as salted meats (Cummings, 1940). The growth of cities and urban markets in the Northwest concentrated the demand for these foods. Farmers began producing marketable surpluses first in the immediacies of the town and later further out in the hinterlands. Commercialization of perishables such as eggs or dairy products grew aided by the development of stockyards and the rail lines proliferation in the Midwest (Craig and Weiss, 1993). Widespread use of fresh meats, vegetables, fruits, and other perishables substantially expanded the demand for refrigeration as a means to preserve food. It was between 1830 and 1860 that the first attempts to refrigerate perishables by using natural ice were made. Between 1860 and 1890 there was an increased use of ice to store food which led to the growth of the cold-storage industry and the development of refrigerator cars mainly in the hands of the meat industry. The ice-based refrigeration became relevant to the meat-packing industry as a means of keeping meats in the packing plant, preserving the product during transportation, and producing cured-meat during hot seasons. Although the availability of railroads to the East had allowed a substantial trade in live animals from producing areas to the city centres, the scarce use of refrigeration had compelled slaughter facilities to locate in city centres to avoid spoilage. Advances in ice-based refrigeration involved progress in the coordination of the shipping, wholesaling and retailing of meat, as well as other perishables. They allowed the meat-packing industry to expand and reallocate towards the livestock-producing areas (Anderson, 1953). In spite of the advantages of shipping dressed-meat relative to live animals, railroads, with a vested interest in livestock trade, resisted the development of the refrigerated meat commerce since it reduced tonnage to haul and jeopardized the investments made in livestock loading and feeding facilities. This resulted in the so-called “big four” packers (Armour, Hammond, Morris and Swift) investing in their own cars and becoming to dominate the dressed beef market (Aduddell and Cain, 1981).

Contrary to the meat packing industry, the egg industry did not experience much change with the ice-based refrigeration. By the 1880s ice-cooled rooms displaced old preservation methods such as water glass or lime water. However, eggs were cold stored only as a last resort and they were usually the inferior grade eggs that would be kept refrigerated, thus giving this product a bad reputation (Pennington, 1941).

In spite of the progress made in natural ice-based refrigeration mechanisms, the potentialities of natural ice were limited and posed many problems to food preservation: apart from the fact that large quantities of ice were needed for refrigeration, which required very expensive structures, ice would generate moist, it was inherently unsanitary and limited the range of temperatures that could be produced. As a result, ice played a relatively small role in the production and distribution of food (Anderson, 1953; Goodwin, Grennes, and Craig, 2002). The storage of perishables remained a problem until the widespread adoption of mechanical refrigeration by the 1890s.

The adoption of mechanical refrigeration represented a vital element to food production and distribution since the early 1890s. It increased arbitrage opportunities in short shelf life commodity markets (Anderson, 1953; Goodwin, Grennes, and Craig, 2002; O’Rourke and Williamson, 1999) and it soon became to dominate meat preservation, allowing to intensify and broaden the industry changes that had begun with the introduction of natural ice. Mechanical refrigeration was also key to the production and distribution of other perishables such as milk and milk products or fruits. It became essential for egg distribution as well, given the fact that egg deterioration varies directly with temperature. In spite of this it should be noted that, during the 1890-1913 period, almost no refrigeration was employed in the early stages of egg marketing. By using poor handling methods, eggs were transferred from the farm to the village merchant. The product would receive better attention after reaching the shipper who would send them to the cities using refrigerator cars. Once in the cities, eggs would either be stored in refrigerated warehouses or be forwarded to retail stores where they would receive no better treatment than in farms (Anderson 1953). It was not until after 1913 that egg handling started to improve under the auspices of the United States Department of Agriculture.

A considerable proportion of eggs such as cracked, dirty or small eggs of inferior quality could not be shipped in the shell to the city markets. By 1900 it began the practice of salvaging these eggs by freezing them. However, the egg-freezing industry failed to grow rapidly mainly due to poor sanitation measures that yielded an inferior product (Koudele and Heinsoh, 1960). By the end of the nineteenth century, the egg-freezing industry was still at a very early stage of its development. It was mainly operating on a small scale, due to the fact that eggs were manually separated until 1912, when the hand separator was invented. Additionally, the egg industry was struggling to solve relevant sanitary and refrigeration problems affecting egg products. Another drawback precluding the expansion of the egg industry was the lack of demand for its products, mainly due to the low quality of the final outcome. The use of dried eggs, prior to the freezing of eggs, was scarce and principally limited to army camps. The consumption of frozen eggs, also limited, mainly came from bakers and other food manufacturers. It was not until World War II that the U.S. egg manufacturing industry, especially the egg drying industry, considerably expanded as a result of both public programs to encourage increased production of eggs through government purchases at supported prices and an increased demand from the Armed Services (Koudele and Heinsohn, 1960).

In spite of the limitations affecting the development of the egg-manufacturing enterprise, refrigeration had a relevant impact on the economics of the egg industry. The introduction of refrigerated cars allowed Corn Belt states to develop their production potential to the maximum (Anderson, 1953). Additionally, refrigerated warehouses altered the seasonal production structure for eggs allowing a smoother supply flow throughout the year. With the exception of the analyses of Goodwin, Grennes, and Craig (2002) and Serra and Goodwin (2004), there are no econometric estimations on price transmission processes within agricultural commodity markets during this period of rapid and relevant changes. In our empirical application, we concentrate on the U.S. egg markets.

Though the adoption of technical improvements at the end of the nineteenth century, especially the introduction of mechanical refrigeration, contributed to a significant increase in the regional price convergence of perishable commodities (O'Rourke and Williamson, 1999), it is a priori unclear how these changes affected egg markets. On the one hand, the refrigerator car and cold-storage warehouses expanded arbitrage possibilities. On the other, the combination of the high perishability of fresh shell eggs, the lack of a high scale egg-manufacturing industry, and the still deficient handling methods used on shell eggs, may have limited spatial arbitrage operations of egg products during the period of analysis relative to other commodities, thus limiting price convergence for eggs. Our paper aims at determining how all these issues affected spatial price relationships within the U.S. egg markets.

Econometric Methods

As noted above, spatial price transmission analyses aim at testing predictions derived from the economic theory and show how price shocks are transmitted across space. Specifically, several studies of the spatial arbitrage condition have focused on the analysis of the regional price spread (see Obstfeld and Taylor, 1997; or Goodwin and Piggott, 2001). Early analyses used price correlations to shed light on spatial price linkages (see Blyn, 1973). This approach, however, has been criticized since it may suffer from inferential biases. Given the fact that prices are typically non-stationary, more recent research has paid particular attention to the time-series properties of price data. Cointegration-based tests have been applied (see Sanjuán and Gil, 2001; Alexander and Wyeth, 1994; Zanias, 1993; or Goodwin, 1992). However, these tests have also been criticized since they do not account for transactions costs (see McNew and Fackler, 1997; or Barrett, 1996).

Current studies have recognized the relevance of allowing for transactions costs in spatial price analyses. These costs include all costs related to spatial arbitrage and trade such as transportation and freight charges, risk premium, refrigeration costs, spoilage, etc. The influence of these costs may result in nonlinear price adjustments (see Obstfeld and Taylor, 1997; Sercu, Uppal and Van Hulle, 1995; Hecksher, 1916). Nonlinearities should be present whenever transactions costs create a band of price differentials within which arbitrage activities do not take place because the marginal costs of trading exceed the marginal benefits. As noted, threshold models are one of the most widely used procedures to capture nonlinearities in price relationships. These models represent price linkages by a combination of different regimes corresponding to the trade/ no trade conditions.

Tong (1978) originally introduced nonlinear threshold time series models. Tsay (1989) developed a method to model threshold autoregressive processes and to test for threshold effects in autoregressive models. Balke and Fomby (1997) extended the threshold autoregressive models to a cointegration framework by considering a model where there is a discontinuous adjustment to a long-run equilibrium. Specifically, the equilibrium error follows a mean-reverting threshold autoregression outside a given range, while having a unit root inside the range. Balke and Fomby (1997) suggest to estimate this model through a two-step approach whereby threshold parameters are chosen through a grid search that minimizes a sum of squared errors (SSE) criterion. In this paper we follow the proposal by Balke and Fomby (1997).

The literature review presented above provides evidence that analyses of price transmission have typically been based on parametric approaches. As noted, these approaches may imply too strong and inadequate assumptions on the functional form that characterizes price relationships. In threshold models, for example, each threshold separates two linear segments representing price adjustment under different regimes. Hence, the transition from one regime to another is assumed to be sharp and discontinuous, involving that the price differentials that motivate individuals to undertake arbitrage activities and/or adjust prices, are common across economic agents. The sudden range reversion assumption implicit in threshold models might be too restrictive in a variety of situations. Teräsvirta (1994) contends that time aggregation and non-synchronous trade are likely to favour smoother transition processes. Mancuso, Goodwin and Grennes (2003) argue that discontinuous range reversion may be adequate if transactions costs and uncertainties were homogeneous across different individuals, but might be too restrictive otherwise. The smoothed TAR models introduced by Teräsvirta (1994) partially overcome this limitation by allowing for gradual adjustments between regimes. However, in being parametric, these models still carry the potential for specification biases as a result of an inappropriate parametric assumption. The risk of specification biases grows when one studies periods of relevant changes that are likely to involve structural breaks and alter transactions costs. On the contrary, nonparametric techniques do not require any preliminary assumption about the shape of the functional form characterizing price linkages. Instead, the data completely inform us how relationships look like, which constitutes a clear advantage over parametric methods. Because nonparametric techniques do not require any preliminary guess on price behaviour, we are interested in applying these techniques to a characterization of spatial price relationships, and in comparing the results with those arising from alternative parametric TAR models. For ease of exposition, we first describe the parametric techniques to then offer details on the nonparametric methods employed.

Threshold Autoregressive Models

Obstfeld and Taylor (1997) devise a method to assess price transmission across spatially separate markets in the presence of transactions costs or uncertainty. Specifically, through the use of a threshold autoregressive model (TAR) of price differentials, they model price relationships as a regime-switching process with a band of inaction. This model introduces an important concept: that of commodity points - thresholds that delimit a region where price differentials show no central tendency due to the lack of arbitrage activities. These points may reflect the influence of the aforementioned transactions costs. Threshold models are useful in situations where the economic behaviour cannot be captured by a single regime. This occurs when some forcing-variable leads a switching, that can occur back and forth, among different regimes. These regimes are represented by different parameter estimates of the underlying model.

Usually, analyses of spatial price behaviour take the magnitude of regional price differentials as the variable that determines regime-switching (Serra and Goodwin, 2004; Mancuso, Goodwin, and Grennes, 2003).

A simple autoregressive model (AR) of price differentials can be represented as: $Y_t = \beta X_{t-1} + e_t$, where $Y_t = \Delta X_t = (P_{it} - P_{jt}) - (P_{it-1} - P_{jt-1})$ represents the adjustment in regional price differentials in period t,[4] being P_{it} and P_{jt} the prices of a certain commodity in two spatially separate markets (i and j). Regional price differentials in the previous period $t-1$ are represented by $X_{t-1} = (P_{it-1} - P_{jt-1})$, while e_t is a white noise error term. Under a TAR model, lagged price differentials $\left(X_{t-1}\right)$ allow to distinguish among different regimes that represent different price behaviour. These different regimes are represented by different values of the parameter β. A three regime TAR can be expressed as follows:

$$Y_t = \begin{cases} \beta^{(1)} X_{t-1} + e_t^{(1)} & if \quad -\infty < X_{t-1} \le c_1 \\ \beta^{(2)} X_{t-1} + e_t^{(2)} & if \quad c_1 < X_{t-1} \le c_2 \\ \beta^{(3)} X_{t-1} + e_t^{(3)} & if \quad c_2 < X_{t-1} \le +\infty \end{cases} \qquad (1)$$

where c_1 and c_2 are the threshold parameters or, in other words, the commodity points. Parameters $\beta^{(j)}$, $j = 1,2,3$, represent the speed at which price differentials are corrected. If arbitrage operates between two markets, then price relationships are expected to exhibit the following pattern. When $c_1 < X_{t-1} \le c_2$ the price gap should have no central

[4] Although AR models involve a dependent variable which corresponds to the first-differenced explanatory variable $\left(\Delta X_t\right)$, we use the notation Y_t since it is more convenient in the exposition of the nonparametric techniques which can be applied to more general situations.

tendency, i.e., there should be no error correction. However, for price differentials outside the central band, arbitrage activities are expected to revert price gaps towards this band. Hence, we expect $\beta^{(1)}$ and $\beta^{(3)}$, i.e. price convergence outside the neutral band, to be negative, while $\beta^{(2)}$ should be (equal or) close to zero.

Since we do not impose the restriction that $\beta^{(1)} = \beta^{(3)}$, a three regime TAR is estimated allowing for asymmetries in price transmission. A significant literature that was recently surveyed by Meyer and von Cramon-Taubadel (2004) has examined the extent to which responses to price shocks are asymmetric. While most such studies have considered transmission of shocks among different levels of the market rather than different locations, the same general issues underlie both forms of asymmetries. In this line of research, Ghoshray (2002) evaluated the asymmetry of adjustments in international wheat prices and its implications for market competitiveness, product differentiation, and government intervention. The literature has identified several causes of asymmetries that may be relevant to the U.S. egg markets at the turn of the XIX century. First, as Abdulai (2000) notes, middlemen in the marketing chain with market power, are likely to respond more quickly to price shocks that involve a reduction in their marketing margins, relative to the speed of response when these margins are increased. This may lead to asymmetries in regional price transmission. As explained above, one of the impacts of the introduction of mechanical refrigeration was to concentrate food production and distribution in the hands of a few dealers. Hence, market power could have affected the egg industry during the period of study. A second explanation for asymmetries relates to unequal development of transportation and handling facilities. As Goodwin and Piggott (2001) explain, trade flows tend to primarily occur in one direction, which leads to these inequalities and favours shipments in one direction relative to the other. Though trade data are unavailable, bibliographical references indicate that egg trade would mainly move from Corn Belt states to the highly populated eastern cities such as New York (Anderson, 1953).[5]

The TAR model can be estimated using sequential iterated least squares regression in two steps. The aim of the first step is to estimate threshold parameters. For this purpose a grid search is conducted. The first or lower threshold is searched over the minimum and the median of the lagged price differentials, while the upper threshold is searched over the range that goes from the median to the maximum value of the lagged price differentials. This search is restricted in order to ensure an adequate number of observations in each regime. For a given pair (c_1, c_2), $\beta^{(1)}$, $\beta^{(2)}$, and $\beta^{(3)}$ can be determined through the OLS regressions of Y_t on X_{t-1} for each sub-sample. From this estimation, the residual sum of squares is derived giving $\hat{\sigma}(c_1, c_2) = \sum_{t=1}^{n} \hat{e}_t(c_1, c_2)^2$. The aim of the grid search is to maximize a standard F test for a linear AR against the alternative of a TAR: $F = \frac{\tilde{\sigma}^2 - \hat{\sigma}^2(c_1, c_2)}{\hat{\sigma}^2(c_1, c_2)} n$, where n represents the number of observations, $\hat{\sigma}^2(c_1, c_2)$ stands for the error variance of the TAR model, being $\tilde{\sigma}^2$ the error variance of the AR model. Hence, in the second step of the process, the estimates of c_1 and c_2 are obtained as: $(c_1, c_2) = \underset{c_1, c_2}{\arg\min}\ \hat{\sigma}^2(c_1, c_2)$, which is equivalent to maximizing *F*. The *F* test for the significance of the differences in parameters across regimes does not have a standard distribution, its p-value is determined following the method provided by Hansen (1997).

Prior to the estimation of the TAR models, we evaluate the time series properties of the data by using unit root and cointegration tests. Previous research has shown that

[5] See Blinder (1982) or Wohlgenant (1985) for other factors causing asymmetries.

unit root tests can be seriously distorted by structural changes in time series. Several test statistics have been proposed to correct this problem (see Rappoport and Reichlin, 1989; Banerjee, Lumsdaine and Stock, 1992; Zivot and Andrews, 1992; or Perron, 1997). In order to determine whether price series are non-stationary or whether the apparent non-stationarity is due to a structural break, we use Perron's (1997) sequential test.[6] Second, cointegration among prices is tested using the Johansen (1988) test.[7]

Local polynomial fitting

Locally weighted regression techniques were introduced in the statistical literature in the late 1970s (see Cleveland, 1979; Stone, 1977; or Katkovnik, 1979) and have become a method of choice in the estimation of regression functions due to their advantages that include easy interpretation, good behaviour near the boundary, capacity to adapt to various designs, as well as the existence of fast algorithms for computing them (see Cleveland, 1979; Cleveland, Devlin and Grosse, 1999; Fan, 1992; or Fan and Gijbels, 1995). We use these techniques to estimate a nonparametric version of a threshold autoregressive model of spatial price differentials. As we explain in the introduction section, we hypothesize that these techniques will suggest a smoother price behaviour than that implied by TAR models.

Consider a series of independently and identically distributed observations (X_{t-1}, Y_t) for $t=1,\ldots,n$, from a population (X_{-1}, Y). As noted above, Y_t represents the adjustment in regional price differentials in period t and X_{t-1} is the value of the regional

[6] The truncation lag parameter is selected using the general to specific recursive method proposed by Perron (1997).

[7] The lag length of the vector autoregressive model is selected to ensure non-autocorrelation of the residuals and to minimize the AIC criterion.

price differentials in the previous period $t-1$. Responses of Y are related to the covariate X_{-1} through a regression function m which is unknown. The unknown function can be approximated using a Taylor series expansion, i.e., by modelling m locally around a certain point x_k, using a simple polynomial of order p. If we represent the local regression function by $m(x_k) = E\left(Y \middle| X_{-1} = x_k\right)$, the Taylor series expansion can be expressed as $m(x) \approx \sum_{j=0}^{p} \frac{m^{(j)}(x_k)}{j!}(x-x_k)^j \equiv \sum_{j=0}^{p} \beta_j (x-x_k)^j$, where $\beta_j = \frac{m^{(j)}(x_k)}{j!}$, being $m^{(j)}(x_k)$ the jth derivative of the regression function $m(x_k)$. This is precisely the basic idea behind local polynomial fitting, i.e., to locally approximate a regression function around a certain point x_k, by fitting a regression surface to the data points within a certain neighbourhood of x_k. From the Taylor series expansion, it is evident that the estimator for the function and its derivatives can be derived from: $\hat{m}^j(x_k) = j!\hat{\beta}_j$, where $\hat{\beta}_j$ denotes the solution to the following weighted least squares problem:

$$\min_{\beta_j} \sum_{t=1}^{n} \left(Y_t - \sum_{j=0}^{p} \beta_j (X_{t-1} - x_k)^j \right)^2 K_t \left(\frac{X_{t-1} - x_k}{h_k} \right) \tag{2}$$

where h_k is the bandwidth that controls for the amount of local averaging or, in other words, the size of the neighbourhood of x_k. K is a kernel function whose role is to smooth data points in the given local neighbourhood. More clearly, K is a weighting scheme to the local least squares problem that down-weights the contribution of those observations away from x_k. Local polynomial fitting techniques require the adoption of several important decisions. A first one concerns the selection of the order of the local polynomial (p). As Fan and Gijbels (1996, chapter 3) explain, since the modelling bias is mainly controlled by the bandwidth, the selection of p is less critical. Although a large value for p would presumably reduce the modelling bias, it can yield noisy estimates. These authors explicitly recommend to use the lowest polynomial order according to the formula: $p = j+1$, which involves $p=1$[8] in our case or, more intuitively, a local linear regression estimation. By using a local linear regression, the least squares problem presented above can be reduced to:

$$\min_{a,b} \sum_{t=1}^{n} \left(Y_t - a - b(X_{t-1} - x_k)\right)^2 K_t\left(\frac{X_{t-1} - x_k}{h_k}\right) \tag{3}$$

where $a = \beta_0$ and $b = \beta_1$ are parameters. Another relevant decision that one needs to take in order to apply local polynomial techniques relates to the bandwidth parameter. The literature has emphasized the relevance of bandwidth selection, based on the influences that h_k can have on the final results. Choosing to select an excessively small bandwidth can result in noisy estimates, while a large bandwidth can yield important modelling biases. The bandwidth can be chosen either subjectively or objectively by the data. A frequently used objective bandwidth selection technique is the cross-validation method, which we adopt to choose an optimum constant bandwidth ($h_k = h$). This method chooses h to minimize the squared prediction error: $\sum_{t=1}^{n}\left(Y_t - \hat{Y}_t\right)^2$, where $\hat{Y}_t$ is any estimate of the regression function involving the smoothing parameter h. For each observation t, the estimate $\hat{Y}_t$ is obtained by computing the regression function without

[8] We are only interested in the regression function $m(x_k)$, but not in its derivatives. Hence, $j = 0$.

the tth observation and by predicting Y_t. In our application, the predicted value for Y_t is derived using the Nadaraya-Watson nonparametric regression estimator: $\hat{Y}_t = \left(\sum_{t=1}^{n^t} K_t \left(\frac{X_{t-1} - x_k}{h} \right) \right)^{-1} \left(\sum_{t=1}^{n^t} K_t \left(\frac{X_{t-1} - x_k}{h} \right) Y_t \right)$, where $\sum_{t=1}^{n^t}$ represents the exclusion of the tth observation. The minimization process requires the computation of the squared prediction error at different bandwidth grid points. The bandwidth h is searched over a region defined by 0.1 and 2 standard deviations of the independent variable X_{t-1}.[9]

The second decision relates to the kernel function. In this paper the Epanechnikov kernel is selected as Fan and Gijbels (1996, chapter 3) have shown it to be an optimal weight function: $K(g) = \frac{3}{4}\left(1 - \left|g^2\right|\right) I_{[-1,1]}(g)$. The solution to the problem in expression (3) is given by:

$$\hat{m}(x_k) = \frac{S_{n,2}(x_k) T_{n,0}(x_k) - S_{n,1}(x_k) T_{n,1}(x_k)}{S_{n,2}(x_k) S_{n,0}(x_k) - S_{n,1}(x_k)^2} \qquad (4)$$

where:

$$T_{n,l}(x_k) = \sum_{t=1}^{n} K_t \left(\frac{X_{t-1} - x_k}{h} \right) (X_{t-1} - x_k)^l Y_t$$

$$S_{n,j}(x_k) = \sum_{t=1}^{n} K_t \left(\frac{X_{t-1} - x_k}{h} \right) (X_{t-1} - x_k)^j .$$

[9] It should be noted here that the corrected Akaike information criteria (Hurvich and Simonoff, 1998) was also used as an alternative method for bandwidth selection and the results derived were very similar.

It is thus clear that the derivation of the local linear regression estimator requires to calculate five different quantities ($S_{n,0}$, $S_{n,1}$, $S_{n,2}$, $T_{n,0}$, and $T_{n,1}$). Though we evaluate the estimators at each point in our sample, it is useful to note that, in order to speed up computations and since estimated curves are usually presented in graphical form, the estimators can be also evaluated at grid points only, which reduces the number of kernel evaluations to be done (see Fan and Gijbels 1996, chapter 3).

An intuitive summary of the local linear regression method is presented in the following lines. This method proceeds by first choosing a point x_k and using the local linear regression to derive an estimate of the regression function at that point. The local regression is estimated using data points around x_k. The size of the neighbourhood is controlled by the bandwidth. Once the bandwidth is selected, a weighted least squares regression is estimated, where sample points away from x_k are down weighted relative to the points closer to x_k. The weighting of points is done by using a kernel function.

We would like to note here that, although we implement local linear regression techniques to estimate the nonparametric counterpart of an AR(1) model, nonparametric regressions can be extended to other multivariate settings in a straightforward manner. However, these extensions are not very useful in practice as a result of the "curse of dimensionality" (see Fan, 2000 for further detail). As this author notes, though multivariate nonparametric modelling is an area which is rapidly evolving, it still merits further research.

The very nature of the LLR involves that no parameters are produced on which to base tests of the usual form. Thus, we rely on a visual graphical inspection of the plotted results to infer price behaviour. To assure inferences, asymptotic confidence intervals are computed following Härdle (1990) in the following way $\hat{Y}_t \pm c_a s$, where c_a

is the (100-a)-quantile of the normal distribution, and s is the sample estimate of the variance of $\hat{Y}$. Specifically, s can be computed as: $s = \frac{c_k \hat{\sigma}}{nh\hat{f}}$, where $c_k = \int_{-1}^{1} K(g)dg$, $\hat{f} = \frac{1}{nh}\sum_{t=1}^{n} K_t \left(\frac{X_{t-1} - x_k}{h} \right)$, and $\sigma^2 = \frac{1}{n\hat{f}}\sum_{t=1}^{n} K_t \left(\frac{X_{t-1} - x_k}{h} \right)\left(Y_t - \hat{Y}_t \right)^2$ (see Härdle 1990, p. 98-101 for further detail).

Since the mathematics of nonparametric techniques are easily written in matrix form, we implemented the technique through a user-defined program executed in Matlab version 6.1. As Fan (2000) notes, applications of nonparametric techniques have been hampered by a lack of software, since they are not available in the most common statistical packages. In the next section we apply the techniques described to a consideration of spatial price relationships in the U.S. egg markets at the turn of the nineteenth century.

Data and Empirical application

The basis for our empirical analysis lies in a study that was conducted in 1913 by the U.S. Department of Agriculture (Holmes, 1913). This study was directed at evaluating the extent to which the advent of cold storage had affected markets for certain important commodities, including eggs. To assess price transmission processes within U.S. egg markets at the turn of the nineteenth century, our empirical analysis uses monthly U.S. egg prices taken from Holmes (1913), observed from October 1881 to October 1911 and quoted at different relevant wholesale markets.[10] Zapoleon (1931) stresses the relevance of the development of centralized warehouses located near to the largest

[10] A four page annex in Holmes (1913) offers detailed information on the name of the publications from which prices were obtained, as well as a detailed statement of the grades of the commodities for which prices were compiled. A copy of this annex is available from the authors upon request.

population centres for regional price convergence. These centralized warehouses would collect and then forward commodities within their areas of operation. New York, Boston and Baltimore are selected as three markets representing these new trade structures. Since they are three spatially close markets, equilibrating price adjustments are likely to have occurred between them. As noted before, the introduction of refrigerated cars allowed Corn Belt states to develop their production potential to the maximum, hence it is also interesting to see how prices in producing areas related to prices in markets close to population centres. Since, by the beginning of the XIXth century, Iowa was the first state with regards to the value of egg and poultry production (Philips, 1909) we have also chosen to analyze the Dubuque market. Hence, our price analysis comprises four markets, of which three are located in the East coast, are relatively close to each other (the distance separating Boston and Baltimore from New York is about 200 miles) and represent large centres of population and a Corn Belt market representing a distant supply area (Dubuque is located at more than a thousand miles from New York).

As explained above, in light of its nonparametric nature, the LLR is more easily interpreted by graphical representation which recommends against specifying too complex models. In this regard, we carry out a pair-wise analysis. Pair-wise analyses are very common in the price transmission literature and represent a natural avenue for studying price relationships, since arbitrage conditions should hold for any pair of prices. Goodwin and Piggott (2001) propose to define pairs of prices composed by a central market price $\left(P_{it-1}\right)$ and another market price $\left(P_{jt-1}\right)$, being the central the largest market in terms of volume. Statistics on trade volumes in the period studied are unavailable. As an alternative, census data for the period studied indicate New York as the most populated U.S. city followed, among other cities and at a considerable distance, by Boston and Baltimore. Hence, New York is likely to have been the largest market in terms of volume. As a result, it is also likely to have lead price formation, a point which can be confirmed by the weak exogeneity tests carried out in the framework of the Johansen's cointegration model. Of interest is to note that other studies have also provided evidence of the leading role often played by big markets located in consumption areas in price formation (see Serra and Goodwin, 2004). As a result, we choose New York as the central market price.[11] Hence, the explanatory variable in our model corresponds to the lagged price differential between New York and another market price ($X_{t-1} = P_{it-1} - P_{jt-1}$), while, as explained above, the dependent variable ($Y_t = X_t - X_{t-1}$) shows the change experienced by this price differential in period t. Prices are measured in levels.

Graphs of the price series are presented in figure 1, where one can see that they follow very similar patterns, suggesting important price transmission processes within U.S. markets. Tests evaluating time series properties of the data are offered in table 1. By allowing both a change in the intercept and in the slope, the Perron (1997) test suggests that the null of a unit root cannot be rejected for any price series.[12] Hence non-stationarity is genuine and should not be attributed to a structural break. Johansen tests provide evidence in favour of stationary long run relationships among the pairs of prices. Weak exogeneity tests indicate that the null of weak exogeneity of the New York market cannot be rejected. This is compatible with the relevance of the New York

[11] At the request of an anonymous referee, however, other possible combinations of markets were also investigated, such as the Baltimore-Boston alternative. Results show that arbitrage activities also hold for these pairs of prices.

[12] Results of the other modalities of the test (which are available upon request) do not differ from the ones presented here.

market in the study by Holmes (1913, p. 8), who quoting Mr. F. G. Urner, noted that "…the greatest development of cold storage as a public utility began with the introduction of mechanical refrigeration shortly before 1890." He also noted that New York's egg consumption increased 52.5% from 1900 to 1910 (Holmes, 1913, p. 11). Underlying the interest in how cold storage may have affected New York egg markets was a strong concern that warehousemen were using cold storage to assert market power and raise prices (Holmes 1913, p.7). Warehouse interests countered that, to the contrary, the advent of cold storage had actually reduced prices for consumers of eggs.

Results derived from the application of the TAR technique are presented in table 2. As noted above, we estimate a three-regime TAR in order to allow for asymmetries in the process of price adjustment. The F-test suggests that threshold effects are statistically significant for all pairs of prices. Results are compatible with the existence of arbitrage activities that correct spatial price gaps as suggested by Heckscher (1916). Hence, in spite of the problems that afflicted the egg industry during the period studied, arbitrage activities aided by the aforementioned technical developments were relevant enough to guarantee price transmission across space. Parameters c_1 and c_2 are an estimate of threshold points and the difference between the two is an estimation of transactions costs bands. We find transactions costs bands to be the largest for the New York-Dubuque model indicating large price differentials between these two markets. As explained before, Dubuque is close to a major producing area and thus is likely to have been a net exporter of eggs. Conversely, New York, as well as Boston and Baltimore, represent major consumption centres. Prices in such centres probably include a significant transactions costs charge. As a result, transactions costs are expected to be higher between an importer and a net exporter market than between two consumption centres, and this is what our results reflect. TAR parameter estimates provide evidence of asymmetric price adjustments that confer a certain advantage to New York over the other markets. While negative (or small positive) price differentials ($-\infty < X_{t-1} \leq c_1$) are quickly corrected, positive price gaps ($c_2 < X_{t-1} \leq +\infty$) are arbitraged away at a slower path, i.e., $\left|\beta^{(1)}\right| > \left|\beta^{(3)}\right|$. The first situation whereby $-\infty < X_{t-1} \leq c_1$ corresponds to New York prices being below the other market prices, while $c_2 < X_{t-1} \leq +\infty$ involves New York wholesalers enjoying the highest price. In band parameter estimates are not statistically significant or take low positive values. Consistently with the existence of transactions costs, this implies that either price differentials within the band follow a random walk process, or are even allowed to experience a very small growth.

Results derived from the local polynomial fitting are graphed in figures 2 to 7. Figures 2, 4 and 6 represent the LLR predicted values and, for comparison purposes, the TAR model predicted values. In figures 3, 5 and 7 we present the nonparametric regressions and their confidence intervals. Except for extreme price differentials, confidence bands are tight, which increases the reliability of inferences from graphical inspection. It can be seen that the nonparametric regressions resemble the lines that represent the values predicted by the TAR model. Consistently with the parametric models, nonparametric techniques suggest that deviations from the long-run equilibrium are corrected in a nonlinear fashion. The slope of the LLR is higher for price differentials outside a certain band and smaller within the band. Hence, and as suggested by the TAR, there is a range of price differentials where equilibrating price adjustments may be less intensive, which is compatible with the existence of transactions costs. Additionally and in agreement with TAR models, nonparametric results suggest that out of band adjustments are not symmetric, with negative price differentials being corrected more quickly relative to positive spreads.

In spite of the similarities between the two methodologies, several differences between the models arise. First, as anticipated, in that it does not assume homogeneous transactions costs across economic agents, LLR allows the transition from one regime to another to be smoother relative to TAR models. Second, where TAR models suggest a still market, local polynomial fitting shows that a price adjustment still takes place. In addition, this adjustment can be relatively quick as is the case with the New York – Boston model. It is also important to note that the area of reduced price adjustment in the LLR is shorter than its parametric counterpart (this is especially evident in figures 4 and 6). Hence, nonparametric techniques imply that markets are more strongly interconnected either though information transmission, or through arbitrage, than what one would conclude from simple observation of the parametric model. In the third place, nonparametric techniques suggest that TAR models, since they are estimated with a limited number of regimes, may have difficulties in capturing the true nature of price relationships. According to LLR, big price differentials cause a change in price behaviour that is not adequately captured by the straight lines derived from the parametric method, and that could suggest necessity to allow for more price regimes in the TAR model. It is also possible that this behaviour is only due to the presence of outliers in the sample.

In short, nonparametric techniques offer certain advantages over parametric methods that are mainly due to the fact that they do not require any assumption about the functional form characterizing price linkages and do not impose transactions costs to be homogeneous and constant. This makes these techniques especially well suited to study price transmission processes during periods of relevant changes that are likely to alter commodity points.

Concluding Remarks

Recent studies of regional price transmission have focused on the nonlinear nature of market price linkages which may be caused by transactions costs. Nonlinear modelling of price adjustments assumes that there exist different states of nature or regimes and that the regime occurring at a certain point in time determines the dynamic price behaviour. Though parametric threshold models have been widely used to capture nonlinearities in price relationships, these techniques may be too restrictive to satisfactorily represent price linkages. As an alternative, we apply nonparametric methods to a consideration of the degree of price transmission between U.S. egg markets at the turn of the nineteenth century. Gordon (2000) labelled this as an era of "Great Inventions" which contributed to the subsequent years of significant productivity growth and noted that the development of mechanical refrigeration and transportation technologies played an important role in this expansion. Contrary to parametric models, nonparametric techniques do not require any assumption about the functional form characterizing price behaviour. Instead, the data completely inform us how the relationship looks like, which constitutes a clear advantage over parametric methods. Nonparametric techniques offer further advantages over parametric methods that are mainly due to the fact that they do not require transactions costs to be homogeneous and constant. This makes these techniques especially well suited to study price transmission processes during periods of relevant changes that are likely to alter commodity points. We compare results derived from local polynomial modelling to those obtained using alternative nonlinear threshold models.

Both techniques suggest that U.S. egg markets were closely interrelated at the turn of the XIX century, thus showing the relevance of technological developments in promoting price transmission across space. However, local linear regressions often suggest a higher degree of price transmission than that implied by threshold autoregressive models. More specifically, and contrary to TAR models, LLRs suggest that even small price differentials tend to be arbitraged away. While TAR models support the existence of a band of price differentials within which no adjustment takes place, nonparametric regressions imply price adjustments even within thresholds, albeit at different rates. Hence, and according to nonparametric techniques, markets are more strongly interconnected either through trade flows or information. Results also suggest that TAR models may have difficulties in adequately capturing price relationship dynamics, especially for extreme price differentials.

References

Abdulai, A. (2000). Spatial Price Transmission and Asymmetry in the Ghanaian Maize Market, *Journal of Development Economics* 63: 327-349.

Aduddell, R. M. and Cain, L. P. (1981). Public Policy Toward "The Greatest Trust in the World", *Business History Review* 55: 217-242.

Alexander, C. and Wyeth, J. (1994). Cointegration and Market Integration: An Application to the Indonesian Rice Market, *Journal of Development Studies,* 30: 303-328.

Anderson, O. E. (1953). *Refrigeration in America: A History of a New Technology and Its Impact*. Princeton, NJ: Princeton Univ. Press.

Balke, N.S. and Fomby, T.B. (1997). Threshold cointegration, *International Economic Review* 38: 627-645.

Banerjee, A., Lumsdaine, R. L., and Stock, J. H. (1992). Recursive and Sequential Tests of the Unit Root and Trend Break Hypothesis: Theory and International Evidence, *Journal of Business and Economic Statistics* 10: 271-287.

Barrett, C.B. (1996). Market Analysis Methods: Are Our Enriched Toolkits Well-Suited to Enlivened Markets?, *American Journal of Agricultural Economics* 78: 825-829.

Barrett, C.B. and Li, J.R. (2002). Distinguishing Between Equilibrium and Integration in Spatial Price Analysis, *American Journal of Agricultural Economics* 84: 292-307.

Blinder, A. S. (1982). Inventories and Sticky Prices: More on Microfoundations of Macroeconomics. *American Economic Review* 72: 365-400.

Blyn, G. (1973). Price Series Correlation as a Measure of Market Integration, *Indian Journal of Agricultural Economics* 28: 56-59.

Chavas, J. P. and Metha, A. (2004). Price Dynamics in a Vertical Sector: the Case of Butter, *American Journal of Agricultural Economics* 86: 1078-1093.

Cleveland, W. S. (1979). Robust locally-weighted regression and smoothing scatterplots', *Journal of the American Statistical Association* 74: 829-836.

Cleveland, W. S., Devlin, S. J. and Grosse, E. (1999). Regression by local fitting, *Journal of Econometrics* 37: 87-114.

Craig, L. A., and Weiss, T. (1993). Agricultural Productivity Growth During the Decade of the Civil War, *The Journal of Economic History* 53: 527-548.

Cummings, R. O. (1940). *The American and His Food: A History of Food Habits in the United States*. Chicago: University of Chicago Press.

Enke, S. A. (1951). Equilibrium Among Spatially Separated Markets: Solution by Electrical Analogue, *Econometrica* 19: 40-47.

Epanechnikov, V. A. (1969). Nonparametric Estimates of a Multivariate Probability Density', *Theory of Probability and its Applications* 14: 153-158.

Fackler, P. L., and Goodwin, B. K. (2001). *Spatial Market Integration. Handbook of Agricultural Economics*. G. Rausser and B. Gardner, eds. Amsterdam: Elsevier Publishing.

Fan, J. (1992). Design-Adaptive Nonparametric Regression, *Journal of the American Statistical Association* 87: 998-1004.

Fan, J. (2000). Prospects of Nonparametric Modeling, *Journal of the American Statistical Association* 95: 1296-1300.

Fan, J. and Gijbels, I. (1995). Data-Driven Bandwidth Selection in Local Polynomial Fitting: Variable Bandwidth and Spatial Adaptation, *Journal of the Royal Statistical Society. Series B(Methodological)* 57:. 371-394.

Fan, I. and Gijbels, I. (1996). *Local Polynomial Modeling and Its Applications*. London: Chapman & Hall.

Goodwin, B.K. (1992). Multivariate Cointegration Tests and the Law of One Price in International Wheat Markets, *Review of Agricultural Economics*: 117-124.

Goodwin, B. K., Grennes, T. and Craig, L. (2002). Mechanical Refrigeration and the Integration of Perishable Commodity Markets, *Explorations in Economic History* 39: 154-182.

Goodwin, B. K. and Piggott, N. (2001). Spatial Market Integration in the Presence of Threshold Effects, *American Journal of Agricultural Economics* 83: 302-317.

Gordon, R. J. (2000). Does the New Economy Measure up to the Great Inventions of the Past? National Bureau of Economic Research Working Paper 7833, August.

Ghoshray, A. (2002). Asymmetric Price Adjustment and the World Wheat Market. *Journal of Agricultural Economics* 53:299-317.

Hansen, B.E. (1997).Inference in TAR Models, *Studies in Nonlinear Dynamics Econometrics* 2: 1-14.

Härdle, W. (1990). *Applied nonparametric regression.* Econometric Society Monographs, vol. 19. New York: Cambridge University Press.

Heckscher, E. F. (1916). Växelkursens Grundval vid Pappersmyntfot, *Ekonomisk Tidskrift* 18: 309–312.

Holmes, G. K. (1913). Cold Storage and Prices, *Bureau of Statistics Bulletin* 101. Washington D.C.: United States Department of Agriculture.

Hurvich, C. M. and Simonoff, J. S. (1998). Smoothing Parameter Selection in Nonparametric Regression Using an Improved Akaike Information Criterion. *Journal of the Royal Statistical Society*, Ser. B, 60: 271-293.

Johansen, S. (1988). Statistical Analysis of Cointegrating Vectors. *Journal of Economic*

Dynamics and Control 12: 231-254.

Katkovnik, V.Y. (1979). Linear and nonlinear methods of nonparametric regression analysis, *Soviet Automatic Control* 5: 25-34.

Koudele, J.W. and Heinsohn, E.C. (1960). The Egg Products Industry of the United States. Part I Historical Highlights, 1900-59, Publications of the Agricultural Experiment Station. Historical Bulletin No 423, Kansas State University.

Lo, M.C. and Zivot, E. (2001). Threshold Cointegration and Nonlinear Adjustment to the Law of One Price, *Macroeconomic Dynamics* 5: 533-576.

Nadaraya, E.A. (1964). On Estimating Regression, *Theory of Probability and its Applications* 9: 141-142.

Mancuso, A. J., Goodwin, B.K. and Grennes, T.J. (2003). Nonlinear Aspects of Capital Market Integration and Real Interest Rate Equalization, *International Review of Economics and Finance* 12: 283-303.

McNew, K. P. and Fackler, P. L. (1997). Testing Market Equilibrium: Is Cointegration Informative?, *Journal of Agricultural and Resource Economics* 22: 191-207.

Meyer, J. and Stephan von Cramon-Taubadel (2004). Asymmetric Price Transmission: A Survey. *Journal of Agricultural Economics* 55:581-611

Obstfeld, M. and Taylor, A.M. (1997). Nonlinear aspects of Goods-Market Arbitrage and Adjustment: Hecksher's Commodity Points Revisited, *Journal of the Japanese and International Economies* 11: 441-479.

O'Rourke, K.H. (1997). The European Grain Invasion, 1870-1913, *Journal of Economic History* 57: 775-801.

O'Rourke, K. H., and Williamson, J. G. (1999). *Globalization and History*. Cambridge, MA: MIT Press.

Pennington, M . E. (1941). Fifty Years of Refrigeration in the Egg and Poultry Industry. *Ice and Refrigeration*, CI.

Perron, P. (1997). Further Evidence on Breaking Trend Functions in Macroeconomic Variables, *Journal of Econometrics* 80: 355-385.

Philips, A. G. (1909). The Marketing of Eggs. Kansas State Agricultural College. Experiment Station, Farm Bulletin 162, Manhattan, Kansas.

Rappoport, P. and Reichlin, L. (1989). Segmented Trends and Non-Stationary Time Series, *The Economic Journal* 99: 168-177.

Samuelson, P. A. (1952). Spatial Price Equilibrium and Linear Programming, *American Economic Review* 48: 283-303.

Sanjuán, A.I., and J.M. Gil. (2001). Price Transmission Analysis: a Flexible Methodological Approach Applied to European Pork and Lamb Markets, Applied Economics 33: 123-131.

Sercu, P., Uppal, R. and Van Hulle, C. (1995). The Exchange Rate in the Presence of Transactions Costs: Implications for Tests of Purchasing Power Parity, *Journal of Finance* 10: 1309-1319.

Serra, T. and Goodwin, B.K. (2004). Regional Integration of Nineteenth Century U.S. Egg Markets, *Journal of Agricultural Economics* 55: 593-608.

Stone, C.J. (1977). Consistent nonparametric regression (with discussion), *The Annals of Statistics* 5:595-645.

Sylla, R. (1969). Federal Policy, Banking Market Structure, and Capital Mobilization in the United States, *The Journal of Economic History* 29: 657-686.

Teräsvirta, T. (1994). Specification, Estimation, and Evaluation of Smooth Transition Autoregressive Models, *Journal of the American Statistical Association* 89 (1994): 208-18.

Tong, H. (1978). *On a Threshold Model in Pattern Recognition and Signal Processing*. C.Chen, ed. Amsterdam: Sijhoff and Noonhoff.

Tsay, R. (1989). Testing and Modeling Threshold Autoregressive Processes, *Journal of the American Statistical Association* 84: 231-240.

Watson, G.S. (1964). Smooth Regression Analysis, *Sankhya: The Indian Journal of Statistics, Series A* 26: 359-372.

Wohlgenant, M. K. (1985). Competitive Storage, Rational Expectations, and Short-Run Food Price Determination. *American Journal of Agricultural Economics* 67: 739-748.

Williamson, J.G. (1995). The Evolution of Global Labor Markets Since 1830: Background Evidence and Hypotheses, *Explorations in Economic History* 32: 141-196.

Zanias, G. P. (1993). Testing for Integration in European Community Agricultural Product Markets. *Journal of Agricultural Economics* 44: 418-427.

Zapoleon, L. B. (1931). International and Domestic Commodities and the Theory of Prices, *Quarterly Journal of Economics* 45: 405–459.

Zivot, E. and Andrews, D.W.K. (1992). Further Evidence on the Great Crash, the Oil Price Shock and the Unit Root Hypothesis, *Journal of Business and Economic Statistics* 10: 251-270.

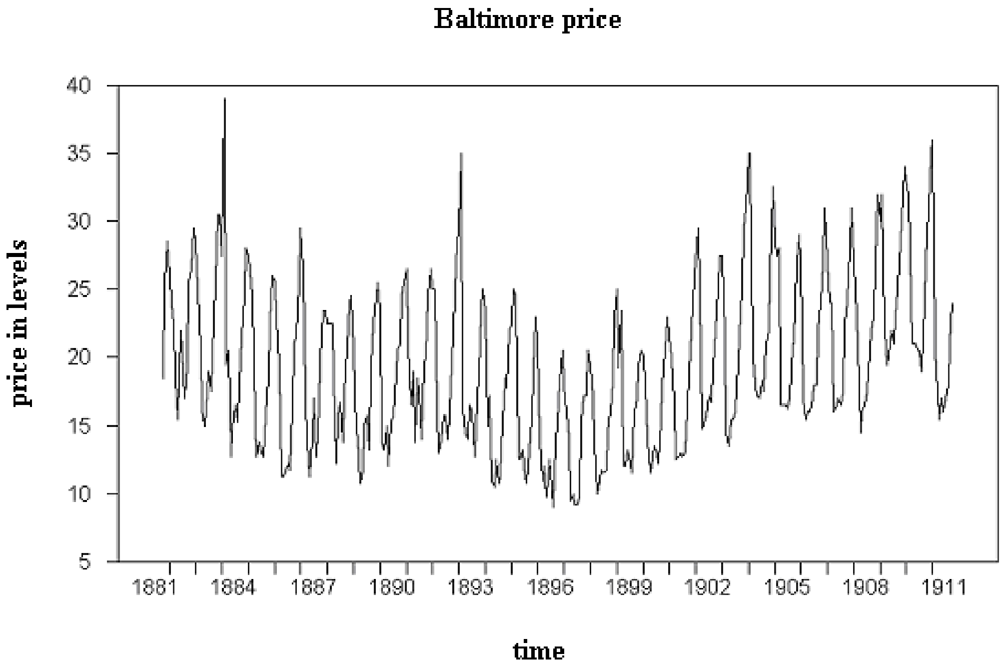
TABLE 1. Time series properties of the data

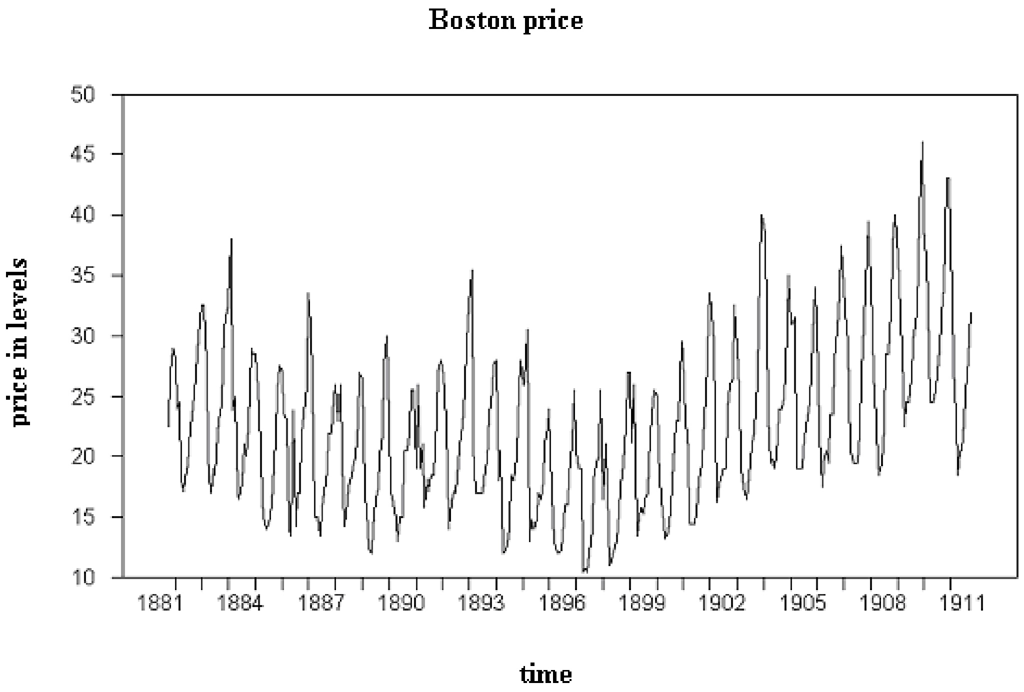
Unit Root Tests			
Price	Perron (critical value at 10%)		Break date
New York	-3.64 (-4.82)		1893:09
Baltimore	-3.39 (-4.82)		1893:01
Boston	-3.79 (-4.82)		1895:02
Dubuque	-3.39 (-4.82)		1893:01
Johansen Tests			
Model	λmax, r=0 (critical value at 10%)	λmax, r=1 (critical value at 10%)	Weak exogeneity test (p-value) Ho: New York is weakly exogenous
New York - Baltimore	11.84 (10.29)	7.4 (7.50)	1.21 (0.27)
New York - Boston	21.72 (10.29)	0.75 (7.50)	1.29 (0.26)
New York - Dubuque	29.46 (10.29)	6.93 (7.50)	0.51 (0.48)

TABLE 2. TAR model parameter estimates

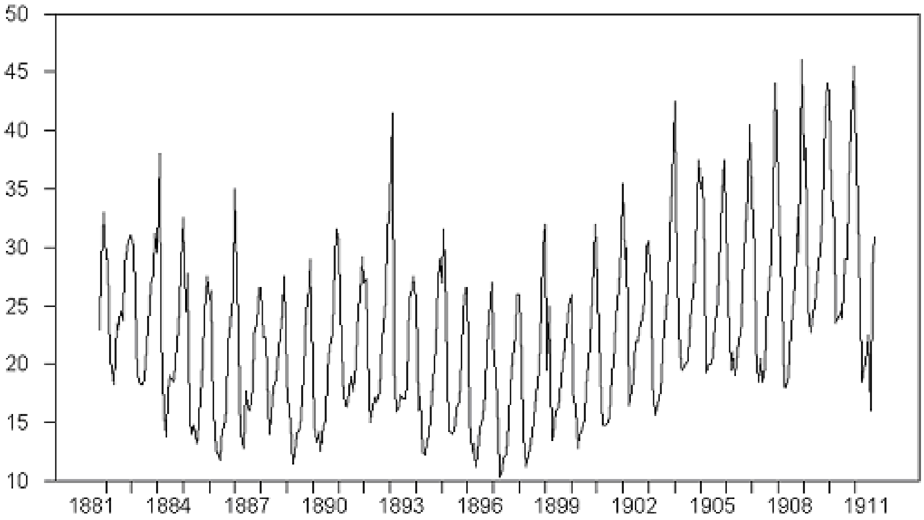
Markets	Thresholds and F-test			TAR parameters		
	c_1	c_2	**F-test (p-value)**	$\beta^{(1)}$ **(standard error)**	$\beta^{(2)}$ **(standard error)**	$\beta^{(3)}$ **(standard error)**
New York-Baltimore	0.13	5.00	53.24 (0.00)	-1.77 (0.26)	0.01 (0.05)	-0.25 (0.05)
New York-Boston	-1.50	1.05	16.89 (0.01)	-0.99 (0.1)	0.09 (0.25)	-0.77 (0.08)
New York-Dubuque	-0.25	7.5	41.65 (0.00)	-1.26 (0.36)	0.11 (0.04)	-0.17 (0.03)

Figure 1. Price Series





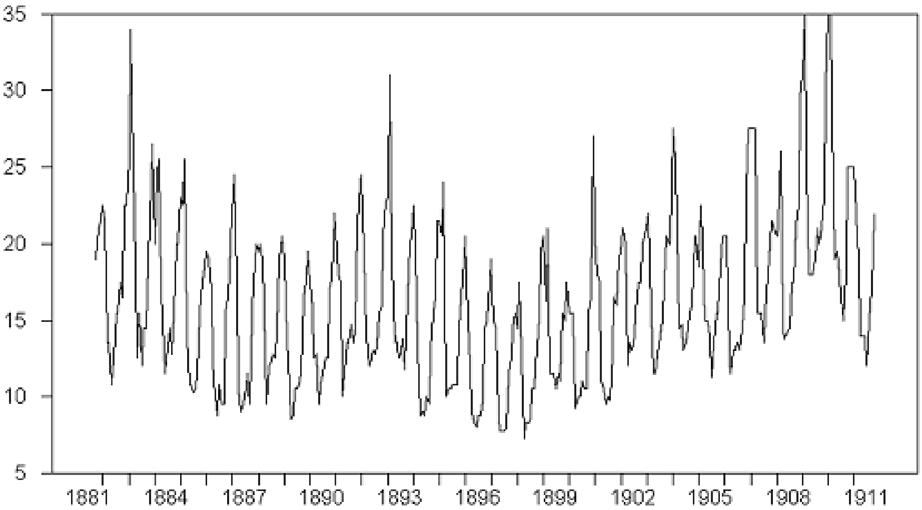
New York price



price in levels

time

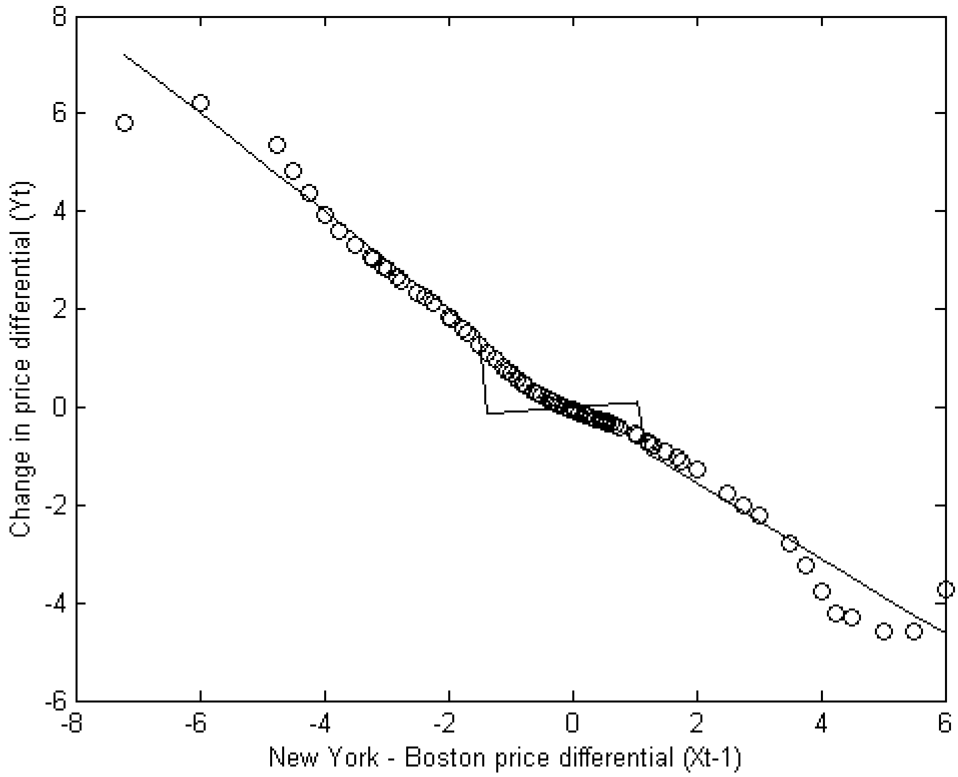
New York price



price in levels

time

Figure 2. LLR and TAR models: Boston-New York

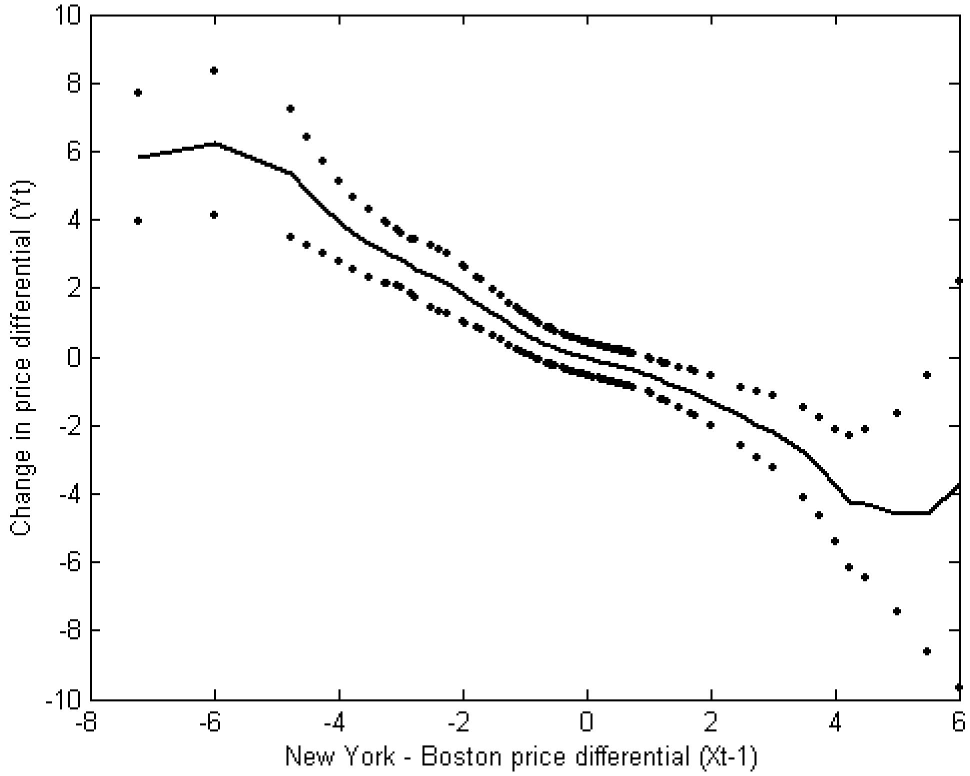


where:

__ represents the TAR model

oo represents the LLR model

Figure 3. LLR model and confidence intervals: Boston-New York

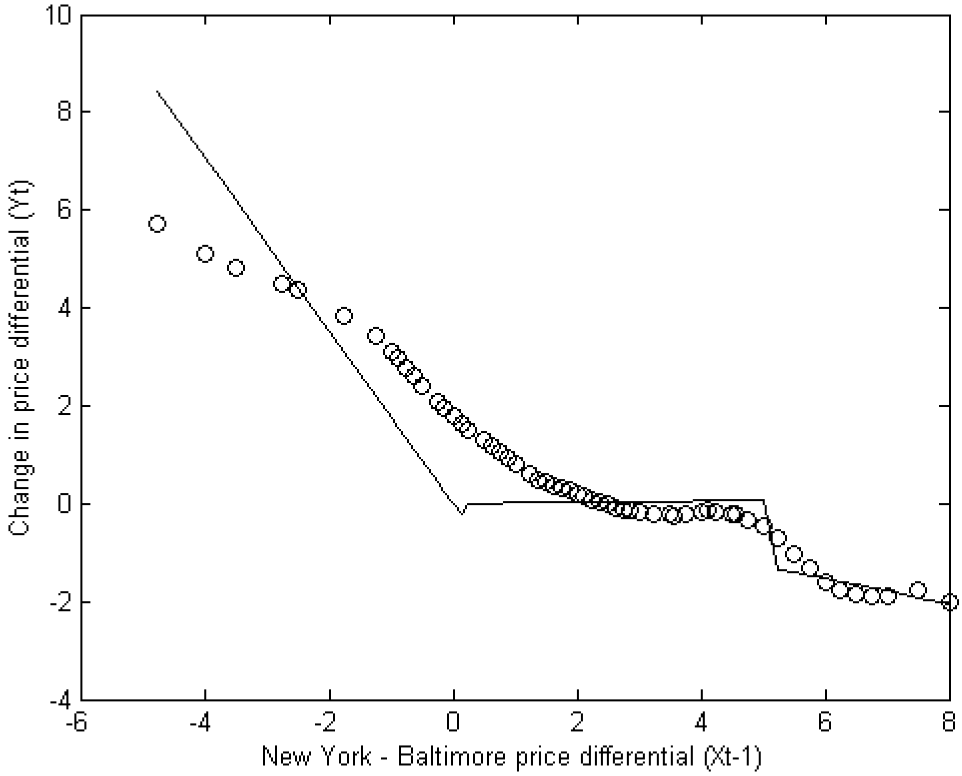


where:

__ represents the LLR model

·· represents confidence bands

Figure 4. LLR and TAR models: Baltimore-New York

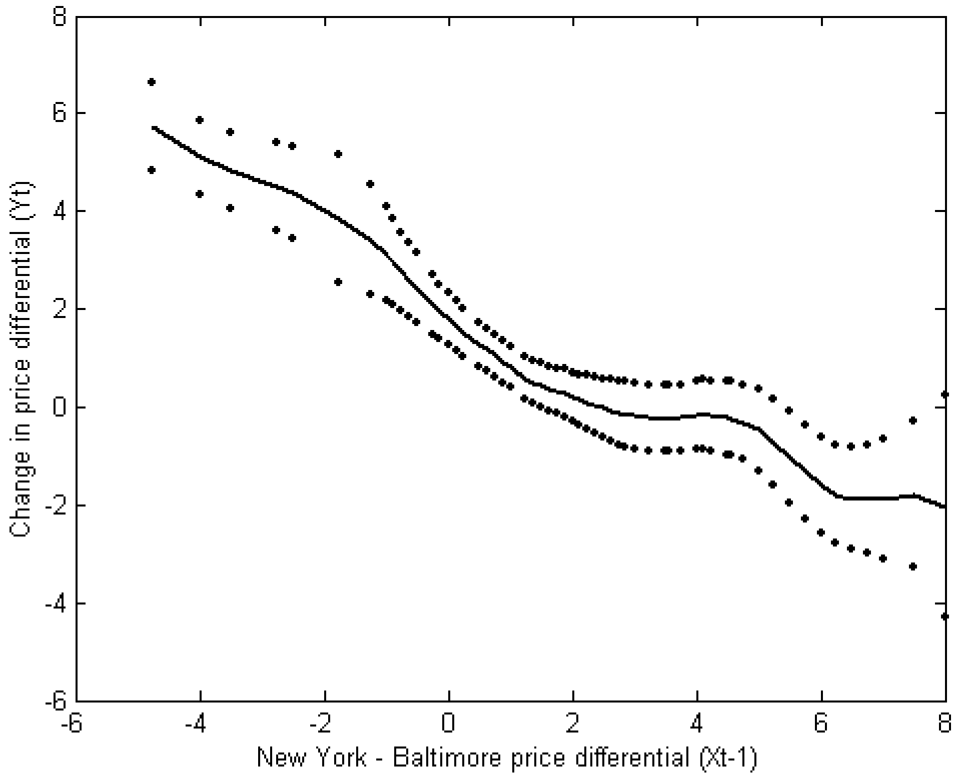


where:

- represents the TAR model

o represents the LLR model

Figure 5. LLR model and confidence intervals: Baltimore-New York

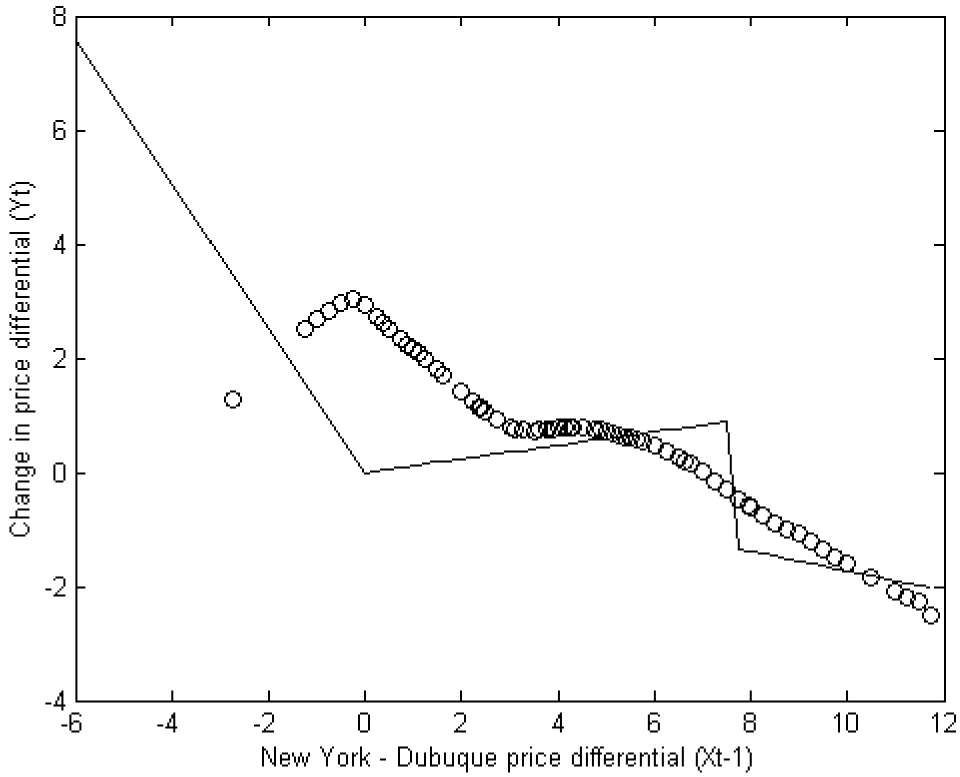


where:

__ represents the LLR model

·· represents confidence bands

Figure 6. LLR and TAR models: Dubuque-New York

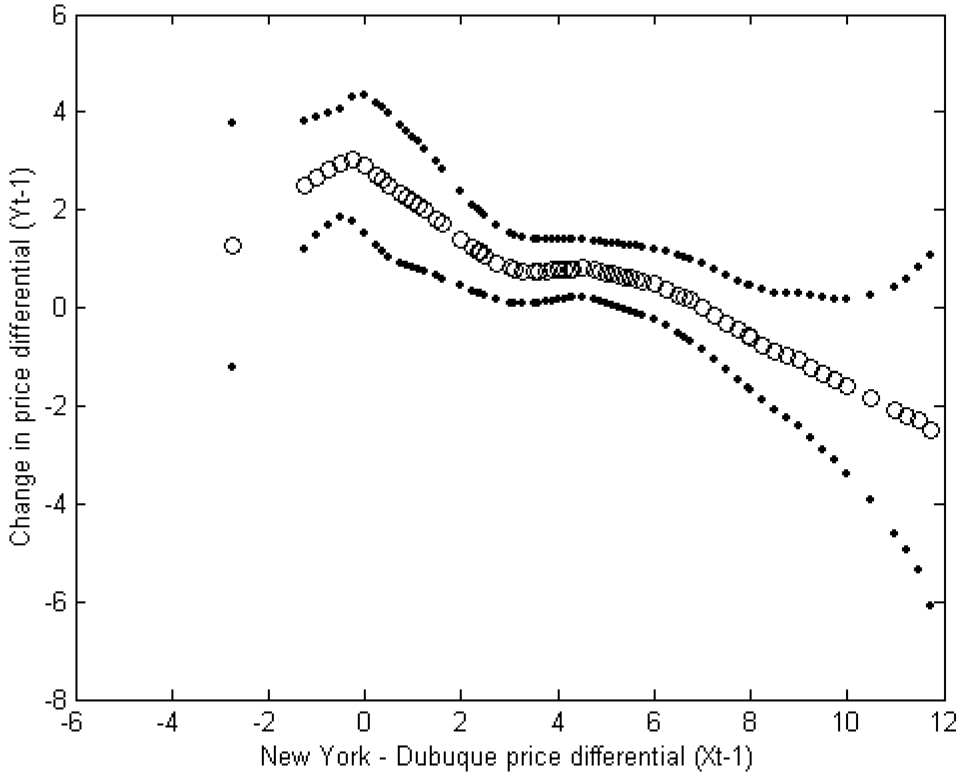


where:

- represents the TAR model

o represents the LLR model

Figure 7. LLR model and confidence intervals: Dubuque-New York



where:

__ represents the LLR model

·· represents confidence bands