Resilience and Dynamic Adjustments in Agroecosystems:
A Quantile Approach

by
Jean-Paul Chavas
and
Salvatore Di Falco

Abstract: The paper presents an investigation of agroecosystem dynamics with an application to yield data from England over the period 1885-2012. The analysis relies on a Threshold Quantile Autoregressive (TQAR) model. The model allows for lag effects to vary across quantiles of the distribution as well as with the values taken by the lagged variables. While it includes as special case the quantile autoregressive (QAR) model, the model specification provides a flexible representation of dynamics and asymmetric adjustments. Applied to wheat yield, the analysis documents the dynamics and persistence of yield adjustments to shocks. It finds statistical evidence against QAR models. The estimates indicate the presence of instability in the lower quantile of the distribution. This is consistent with presence of resilience in wheat yield.

Keywords: agroecosystem dynamics, resilience, yield, quantile regression, threshold

JEL: Q1, C14, C22, C51

1 Jean-Paul Chavas is Professor, Department of Agricultural and Applied Economics, University of Wisconsin, Madison, WI 53706, USA; Tel: (608)261-1944; Email: jchavas@wisc.edu. Salvatore Di Falco is Professor, Department of Economics, University of Geneva, Geneva, Switzerland; Email: salvatore.difalco@unige.ch. This research was supported in part by a Graduate School research grant from the University of Wisconsin, Madison.
Resilience and Dynamic Adjustments in Agroecosystems: 
A Quantile Approach

1. Introduction

Resilience is an ecological concept. It indicates the ability of a natural system to withstand a shock and to recover quickly from adverse perturbations (e.g., Holling; Gunderson). Resilience gives insight to the dynamic of a system (Perrings).\(^1\) In the last twenty years the link between resilience and sustainability has been recognized.\(^2\) In the literature, many connections have been drawn between resilience and sustainable development (e.g. Arrow et al.; Folke et al.; Walker and Salt; Mäler) to the extent that some authors use the notions of resilience and sustainable development as synonymous: "A system may be said to be Holling-sustainable, if and only if it is Holling-resilient" (Common and Perrings, p. 28). As a result, a growing body of literature has been published focusing on the identification of the theoretical conditions under which resilience can be taken as a sustainability criterion (e.g., Common and Perrings; Perrings; Holling and Walker; Mäler et al.; Derissen et al.). The empirical evidence on the role of resilience in dynamic adjustments of agroecosystems is, however, very scarce.\(^3\) This paper aims to fill this gap by providing a case study on agroecosystem production dynamics. The paper develops a method to evaluate the yield distribution and its dynamic adjustments. The analysis relies on a threshold quantile autoregressive (TQAR) model that provides a flexible representation of the dynamics in the yield distribution (as discussed below).

The objective of this paper is to study how yields adjust to shocks. In particular, the analysis examines how agroecosystems production (captured by yield) adjusts to adverse shocks from one growing season to the next. The paper makes two contributions to the literature. The first contribution is methodological. The paper develops a method to address resilience issues in
agroecosystems by evaluating the yield distribution and its dynamic adjustments. The analysis relies on a threshold quantile autoregressive (TQAR) model that provides a flexible representation of the dynamics in the yield distribution (as discussed below). The second contribution is its empirical application to wheat yield in England. We combine historical and more modern data to obtain useful information on both short term and longer term effects of weather shocks. The empirical analysis documents the presence of significant dynamics in wheat yields.

Our methodological contribution can be situated in the context of the econometric analysis of dynamics, which typically involves the specification and estimation of a difference equation expressing current values of a state variable as function of past values (e.g., Hamilton 1994; Enders). When, the equation is linear, this leads to the autoregressive (AR) model, which has been commonly used in applied time series analysis (e.g., Hamilton 1994; Enders). But the AR model is restrictive in two ways: 1/ it focuses on the dynamics of the mean (at the exclusion of the dynamics other moments); and 2/ it neglects the possibility that dynamics can vary in different points of the distribution. Extending the dynamics to higher moments has been done in generalized autoregressive conditional heteroscedastic (GARCH) models (e.g., Bollerslev).

Other extensions have explored nonlinear dynamics (e.g., Pesaran and Potter; Potter), including threshold autoregressive models (Tong; Cao and Tsay), smooth transition autoregressive models (Terasvita and Anderson) and Markov switching models (Hamilton, 1989). Another direction for extension has been exploring dynamics using quantile regression. This includes the quantile autoregressive (QAR) model proposed by Koenker and Xiao. The QAR model is more flexible than the AR model in the sense that it allows dynamics to vary across quantiles, thus providing a
basis to explore asymmetric dynamics across quantiles of the distribution (Koenker and Xiao; Tsong and Lee).

This paper relies on a threshold quantile autoregressive (TQAR) model to support an investigation of asymmetric dynamics in wheat yield. Following Cai and Stander, and Galvao et al., a TQAR model includes as special case the quantile autoregressive (QAR) model (which allows dynamics to vary across quantiles). But as in threshold models (e.g., Tong, Hansen), the TQAR specification is more flexible than the QAR model: it allows dynamics to vary also across regimes. We consider a particularly flexible specification of dynamics and asymmetric adjustments by considering multiple regimes defined according to the value taken by each lagged variable. Note that this is more flexible than the TQAR models typically found in previous literature. Thus, our approach allows for nonlinear dynamics in a parsimonious specification that makes it attractive for empirical analysis. As such, this paper contributes to the literature on applied nonlinear dynamics (e.g., Tong; Pesaran and Potter; Potter; Enders).

Our analysis relies on a reduced form specification of yield dynamics. We argue that a reduced form approach is conceptually valid and empirically tractable: it provides a basis to evaluate how shocks affect yield and its distribution both in the current period and over time. Applied to a reduced form representation of yield dynamics, the TQAR model gives estimates of the evolving distribution function of yield and its dynamic response to shocks. Of special interest are the static and dynamic effects of adverse shocks, i.e. of shocks located in the lower tail of the yield distribution. To the extent that unforeseen weather events are the main source of yield risk, our investigation provides useful insights on yield adjustments to climate change. By allowing dynamics to vary depending on the values taken by lagged variables, the TQAR model provides
a convenient basis to investigate the resilience of yield to adverse shocks. Our empirical investigation finds evidence that wheat yield exhibit resilience to adverse shocks.

This paper is organized as follows. Section 2 presents the conceptual model. It discusses both a structural form approach and reduced form approach to yield dynamics. It shows how a reduced form approach is both conceptually valid and empirically tractable. And it argues that a TQAR model provides a flexible representation of dynamic yield adjustments. Section 3 presents an application to time series data on wheat yield in England over the period 1885-2012. The results are presented in section 4. The econometric analysis finds statistical evidence against both AR and QAR models. The results document the evolution of the distribution of wheat yield and its dynamic adjustments to shocks. They indicate the presence of instability in the lower quantile of the distribution. This is consistent with presence of resilience in the dynamics of wheat yield. Finally, section 5 concludes.

2. The Model

We want to investigate the determinants and evolution of crop yield. At time t in a particular location, denote the yield of given crop by the variable \( y_t \in Y \subset R \). Assume that \( y_t \) evolves over time according the p-th order difference equation

\[
y_t = f_0(y_{t-1}, \ldots y_{t-p}; z_t, z_{t-1}, \ldots, z_{t-n})
\]

where \( z_t \in Z \subset R^r \) is a vector of r variables affecting \( y_t \), and \( f_0 \) is a function mapping \( Y^p \times Z^n \) into \( Y \). The vector \( z_t \) captures management (e.g., cultural practices, crop rotation, fertilizer and pesticide use, genetic selection), soil conditions (e.g., soil structure, depth, organic matter, soil nutrients) as well as environmental effects (e.g., diseases, weather conditions). Equation (1)
allows for complex dynamics involving lag effects of $y_t$ (up to $p$ lags) and $z_t$ (up to $n$ lags). Assume that the vector $z_t$ has its own dynamics given by

$$z_t = g_0(y_{t-1}, \ldots, y_{t-p}; z_{t-1}, z_{t-2}, \ldots, z_{t-n}), \quad (2a)$$

where $g_0$ is a function mapping $Y^p \times Z^n$ into $Z$. Equation (2a) allows for dynamic effects in $z_t$ (with up to $n$ lag effects in $z_t$) and $y_t$ (with up to $p$ lags). When $z_t$ includes soil quality and pest population, then (2a) represents the dynamic evolution of the agro-ecosystem supporting agricultural production. When $z_t$ includes management (e.g., cultural practices, fertilizer use), then equation (2a) represents the farmers’ management decision rules used in crop production. And when $z_t$ includes weather (e.g., rainfall, temperature), then (2a) captures the dynamics of weather conditions relevant in the analysis of climate change issues. As such, equation (1) and (2a) provide a generic representation of dynamics involving both $y_t$ and $z_t$. Importantly, the lag effects of $y$ or $z$ on $y_t$ and $z_t$ in (1) and (2a) allow for feedback effects.

After successive substitutions, note that (2a) can be alternatively written as

$$z_t = g_0(y_{t-1}, \ldots, y_{t-p}; g_0(y_{t-2}, \ldots, y_{t-p-1}; z_{t-2}, \ldots, z_{t-n-1}), z_{t-2}, \ldots, z_{t-n})$$

$$= g_0(y_{t-1}, \ldots, y_{t-p}; g_0(y_{t-2}, \ldots, y_{t-p-1}; g_0(y_{t-3}, \ldots), \ldots, z_{t-n-1}), g_0(y_{t-3}, \ldots), \ldots, z_{t-n})$$

$$= \ldots$$

$$= g(y_{t-1}, y_{t-2}, \ldots; z_0, y_0), \quad (2b)$$

where $(z_0, y_0)$ denote initial conditions for $(z_t, y_t)$. Assume that the lagged effects of $(y_{t-m-1}, y_{t-m-2}, \ldots)$ on $z_t$ or $y_t$ in (1) or (2b) are negligible for some $m \geq p$. Then, substituting equation (2b) into (1), equation (1) can be written as the $m$-th order stochastic difference equation

$$y_t = f(y_{t-1}, \ldots, y_{t-m}, e_t), \quad (3)$$

where $f(y_{t-1}, \ldots, y_{t-m}, e_t) = f_0(y_{t-1}, \ldots, y_{t-p}; g(y_{t-1}, y_{t-2}, \ldots), g(y_{t-2}, \ldots), \ldots)$, and $e_t$ is a random vector assumed to be identically and independently distributed with a given distribution function. The vector $e_t$ represents unobservable effects (e.g., unpredictable weather shocks).
Equations (1) and (3) provide each a valid representation of dynamics. Equation (1) is a “structural equation” that captures the effects of current and lagged values of \( z \) on yield \( y_t \). As noted above, the \( x \) variables can represent the effects of management, agro-ecological conditions and weather on crop yield. Equation (3) is the corresponding “reduced form” equation that represents the net effects of past yields on current yield \( y_t \). While it does not reflect structural information about the determinants of yields, equation (3) provides a valid representation of yield dynamics. And it has the advantage of not requiring direct measurements on the variables in \( x \) and their lagged values. In general, the \( x \) vector includes many variables, some of them somewhat difficult to measure. On that basis, the “reduced form” specification given in (3) will be easier to use in applied work. The dynamic analysis presented in the rest of the paper will focus on equation (3).

Given (3), define the conditional distribution function \( F(c \mid y_{t-1}, \ldots, y_{t-m}) = \text{Prob}[y_t \leq c \mid y_{t-1}, \ldots, y_{t-m}] \). The associated conditional quantile function is defined as the inverse function \( Q(r \mid y_{t-1}, \ldots, y_{t-m}) = \inf \{c : F(c \mid y_{t-1}, \ldots, y_{t-m}) \geq r\} \) for \( r \in (0, 1) \). This includes as special case the conditional median \( Q(0.5 \mid y_{t-1}, \ldots, y_{t-m}) \). Both the distribution function \( F(c \mid y_{t-1}, \ldots, y_{t-m}) \) and the quantile function \( Q(r \mid y_{t-1}, \ldots, y_{t-m}) \) are generic: they provide a complete characterization of the dynamics of \( y_t \) under a general specification of \( g(y_{t-1}, \ldots, y_{t-m}, e_t) \) in equation (3). In the rest of the paper, we will make extensive use of the quantile function \( Q(r \mid y_{t-1}, \ldots, y_{t-m}) \) in the analysis of the dynamics of \( y_t \).

In general, equation (3) can be alternatively written as the first-order difference equation

\[
\begin{bmatrix}
    y_1 \\
    y_{t-1} \\
    \vdots \\
    y_{t-m+1}
\end{bmatrix}
= \begin{bmatrix}
    f(y_{t-1,1}, \ldots, y_{t-m,1}, e_t) \\
    y_{t-1,1} \\
    \vdots \\
    y_{t-m+1,1}
\end{bmatrix}
\equiv h(w_{t-1}, e_t).
\]  

(4)
where \( w_t \in \mathbb{Y}^m \subset \mathbb{R}^m \). Equations (3)-(4) include as a special case the standard autoregressive model AR(m) where \( f(y_{t-1}, \ldots, y_{t-m}, e_t) \) is linear in \( (y_{t-1}, \ldots, y_{t-m}) \) and (3) becomes \( y_t = \alpha_0(e_t) + \alpha_1 y_{t-1} + \ldots + \alpha_m y_{t-m} \). In this context, from (4), denote the derivatives of \( h(\cdot) \) with respect to \( w_{t-1} \) by the \((m \times m)\) matrix \( A \equiv \partial h(w_{t-1}, e_t)/\partial w_{t-1} = \begin{bmatrix} \alpha_1 & \alpha_2 & \ldots & \alpha_m \\ I_{m-1} & 0_{m-1} \end{bmatrix} \) where \( I_{m-1} \) is an identity matrix of dimension \((m-1) \times (m-1)\), \( 0_{m-1} \) is a \( (m-1) \times 1 \) vector of zeros. Let the Eigenvalues of \( A \) be \( (\lambda_1, \ldots, \lambda_m) \), where the roots are ordered according to their modulus, \( |\lambda_1| \geq |\lambda_2| \geq \ldots \geq |\lambda_m| \). The dynamics of \( y_t \) can then be assessed using the Eigenvalues of the matrix \( A \): the dynamics of \( y_t \) is asymptotically stable (in the sense that \( \lim_{t \to \infty} y_t = y^c \) for any initial condition \( y_0 \)) if and only if the dominant root of \( A \) is in the unit circle: \( |\lambda_1| < 1 \); and the speed of convergence to the long run equilibrium \( y^c \) is measured by \( |\lambda_1| \).

Note that, under an AR(m) specification, the quantile function takes the form \( Q(r \mid y_{t-1}, \ldots, y_{t-n}) = a_0(r) + \alpha_1 y_{t-1} + \ldots + \alpha_m y_{t-m} \). Thus, the standard autoregressive model imposes two sets of restrictions on the nature of dynamics: 1/ the conditional quantiles are linear in \( (y_{t-1}, \ldots, y_{t-m}) \); and 2/ the quantile slopes \( (\alpha_1, \ldots, \alpha_m) \) are constant across quantiles \( r \in (0, 1) \). Both are rather strong restrictions that we will relax below.

Below, we explore the dynamics of \( y_t \) allowing for nonlinearity of \( f(y_{t-1}, \ldots, y_{t-m}, e_t) \) in equations (3)-(4). In the case where \( f(y_{t-1}, \ldots, y_{t-m}, e_t) \) is differentiable in \( (y_{t-1}, \ldots, y_{t-m}) \), define the \((m \times m)\) matrix \( A(z_{t-1}, e_t) \equiv \partial h(z_{t-1}, e_t)/\partial z_{t-1} \). Evaluated at point \( (z_{t-1}, e_t) \), denote the Eigenvalues of \( A(z_{t-1}, e_t) \) by \( \lambda_1(z_{t-1}, e_t), \lambda_2(z_{t-1}, e_t), \ldots, \lambda_m(z_{t-1}, e_t) \), where \( \lambda_1(z_{t-1}, e_t) \) is the dominant root of \( A(z_{t-1}, e_t) \). These Eigenvalues can provide useful information on dynamics. Indeed, in the neighborhood of point \( (z_{t-1}, e_t) \), the dynamics of \( y_t \) is locally stable if \( |\lambda_1(z_{t-1}, e_t)| < 1 \); and the magnitude of \( |\lambda_1(z_{t-1}, e_t)| \) provides a measure of local stability of \( y_t \).
Relying on the conditional quantile function \( Q(r \mid y_{t-1}, \ldots y_{t-m}) \), we examine alternative specifications for \( Q(r \mid y_{t-1}, \ldots y_{t-m}) \). We focus our attention on the case where the conditional quantile function takes the form \( Q(r \mid y_{t-1}, \ldots y_{t-m}) = x(y_{t-1}, \ldots y_{t-m}) \beta_r, r \in (0, 1) \), where \( x(\cdot) \) is a \((1 \times k)\) vector and \( \beta_r \) is a \((k \times 1)\) vector of parameters. This restricts the analysis to situations where conditional quantiles are linear in the parameters \( \beta_r \). This assumption is made for the sake of convenience: it will help simplify our econometric analysis. Importantly, this specification allows the parameters \( \beta_r \) to vary across quantiles. In addition, the functions \( x(y_{t-1}, \ldots y_{t-m}) \) can possibly be nonlinear, thus allowing for the presence of nonlinear dynamics.

Consider a sample of \( n \) observations on \((y, x)\). Denote the \( i \)-th observation by \((y_i, x_i), i \in N \equiv \{1, \ldots, n\}\). Throughout the paper, we assume that the \((n \times k)\) matrix \( X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \) has rank \( k \). For a given quantile \( r \in (0, 1) \) and following Koenker, the quantile regression estimate of \( \beta_r \) is

\[
\hat{\beta}_r \in \arg\min_{\beta} \{ \sum_{i \in N} \rho_r(y_i - x_i \beta) \},
\]

where \( \rho_r(w) \equiv w \cdot [r - I(w < 0)] \) and \( I(\cdot) \) is the indicator function. As discussed in Koenker, the quantile estimator \( \hat{\beta}_r \) in (5) is a minimum distance estimator that can be obtained by solving linear programming problems.

We first consider a simple linear specification: \( x(y_{t-1}, \ldots y_{t-m}) = [1, y_{t-1}, \ldots y_{t-m}] \) and \( Q(r \mid y_{t-1}, \ldots y_{t-m}) = \beta_{0r} + y_{t-1} \beta_{1r} + \ldots + y_{t-m} \beta_{mr} \). This corresponds to the Quantile Autoregressive model QAR(m) proposed by Koenker and Xiao. For a given \( r \), this restricts the conditional quantile function to be a linear function of \((y_{t-1}, \ldots y_{t-m})\). However, as emphasized by Koenker and Xiao, the quantile slopes parameters \( \beta_r = (\beta_{0r}, \beta_{1r}, \ldots, \beta_{mr})' \) can vary across quantiles (in contrast to the standard AR model). As such, the QAR model provides a more flexible
representation of dynamics than an AR model. In particular, Koenker and Xiao argue that, when the \( y_t \)’s are covariance stationary and the coefficients \( \beta_r \) are treated as random variables, a QAR process can allow for transient forms of explosive behavior (again in contrast to a stationary AR model which cannot). In this context, Koenker and Xiao also show that the quantile estimator \( \hat{\beta}_r \) in (5) is consistent and satisfies a central limit theorem.

While the QAR model allows for quantile slopes to vary across quantiles, it remains restrictive in the sense that, for a given \( r \), it still constrains the quantile impacts of \( (y_{t-1}, \ldots, y_{t-m}) \) to be linear. Next, following Cai and Stander and Gavao et al., we relax this assumption by introducing thresholds and regime switching in QAR. We consider the case of multiple regimes. This is done by partitioning the set \( Y \) into \( K \) mutually exclusive intervals: \( Y = (S_1, \ldots, S_K) \subset \mathbb{R} \), where \( S_j = (b_{j-1}, b_j), \ b_{j-1} < b_j, \ j \in J \equiv \{1, \ldots, K\} \), the \( b_j \)’s being threshold points, \( b_0 \) being a lower bound of \( Y \) and \( b_K \) being an upper bound. In this context, we consider the following Threshold Quantile Autoregressive (TQAR) specification

\[
Q(r \mid y_{t-1}, \ldots, y_{t-m}) = \beta_{0r} + \sum_{j \in J} y_{t-1} D_j(y_{t-1}) \beta_{1jr} + \ldots + \sum_{j \in J} y_{t-m} D_j(y_{t-m}) \beta_{mjr},
\]

where \( D_j(y) = \begin{cases} 1 & \text{when } y \in S_j, \ j \in J \end{cases} \). The parameter \( \beta_{ijr} \) in (6) measures the marginal effect of \( y_{t-i} \) on the \( r \)-th quantile of \( y_t \) when \( y_{t-i} \) is in the \( j \)-th interval \( S_j \). The specification in (6) allows dynamics to vary across regimes defined according to the values taken by lagged variables. By allowing for multiple regimes, (6) provides a very flexible representation of dynamics and asymmetric adjustments. In particular, as noted in footnote 1, by allowing for multiple regimes across multiple lagged variables, model (6) is more general (and thus less restrictive) than most threshold models found in previous literature (e.g., Tong; Cai and Stander; Galvao et al.).
The specification (6) generates a model that remains linear in the parameters (as in (5)). When the process generating the \{y_t\} is stationary and under some regularity conditions, the quantile estimator \( \beta_r^c \) in (5)-(6) is consistent. As showed by Koenker and Galvao et al., a central limit theorem implies that its asymptotic distribution is

\[
\sqrt{n} [\beta_r^c \cdot \beta_r^c] \rightarrow N[0, r(1-r) \Omega_1^{-1} \Omega_0 \Omega_1^{-1}]
\]

(7a)

where \( \Omega_0 = \lim_{n \to \infty} \left( \sum_{i=1}^{n} (x_i' x_i)/n \right) \) and \( \Omega_1 = \lim_{n \to \infty} \left( \sum_{i=1}^{n} (f_i x_i' x_i)/n \right) \) are positive definite matrices, and \( f_i \in (0, \infty) \) is the conditional density function associated with the i-th observation. And the asymptotic covariance matrix of \( \beta_r^c \) and \( \beta_r^c \) is

\[
\text{ACov}(\sqrt{n} \beta_r^c, \sqrt{n} \beta_r^c) \rightarrow [\min(r, r') - r r'] \Omega_1^{-1} \Omega_0 \Omega_1^{-1}.
\]

(7b)

The asymptotic distribution of \( \beta_r^c \) given in (7a) and (7b) provides a basis for conducting hypothesis testing. Note that (4) reduces to the QAR model when \( \beta_{i1r} = \beta_{i2r} = \ldots = \beta_{iKr} \) for all \( i \in \{1, \ldots, m\} \) and all \( r \in (0, 1) \). Thus, the TQAR specification (4) nests QAR as a special case. As such, it provides an empirical basis for testing the validity of the QAR model. In situations where the condition \( \beta_{i1r} = \beta_{i2r} = \ldots = \beta_{iKr} \) is not satisfied for some \( i \) and \( r \), then (4) allows the marginal effects of \( y_{t-i} \) on the \( r \)-th quantile of \( y_t \) to vary with \( y_{t-i} \) across the intervals \( S_j \)'s. This permits the presence of multiple regimes for dynamics depending on the values taken by the lagged dependent variables \( y_{t-1}, \ldots, y_{t-m} \). In this case, the TQAR model (4) can exhibit two forms of nonlinear dynamics: 1/ it allows for the effects of \( y_{t-i} \) to vary across the quantiles of \( y_t \) (when the parameters \( \beta_{ijr} \)'s vary with \( r \)); and 2/ it allows for the effects of \( y_{t-i} \) to vary depending on the interval of \( Y \) where \( y_{t-i} \) is located (when the parameters \( \beta_{ijr} \)'s vary with \( j \) for each lag \( i \)). Neither effect is present in the standard AR model. And only the first effect is present in the QAR model. It seems important to allow for both effects as each one reflects a different aspect of nonlinear
dynamics. Indeed, the first effect captures nonlinear dynamics related to the “current state” $y_t$. This contrasts with the second effect which captures nonlinear dynamics related to the “previous states” $y_{t-i}$’s. It is largely an empirical matter to evaluate the presence of these two effects. The TQAR specification (4) offers a formal framework for conducting such inquiries. More generally, it gives an empirical basis to investigate the nature and relative importance of these two effects. This is illustrated next in an empirical application.

3. An Application to Wheat Yield

The analysis is applied to annual data on wheat yield in England over the period 1885-2012. The data are presented in Figure 1, where wheat yield is measured in ton per hectare (t/ha). Our objective is to analyze the dynamics of wheat yield production.

The data reported in Figure 1 indicate that average wheat yield was fairly constant before 1940 but that it exhibited sharp increases after 1950. This reflects in large part the positive and large effects of genetic selection and improved management on wheat productivity over the last 60 years. To capture such effects, we introduced trend variables in the analysis. We introduce an overall time trend measured by the variable $yr_0 = year - 1900$. We also define the trend variables $yr_i = \begin{cases} 0 & \text{when year } < T_i \\ \text{year} - T_i & \text{when year } \geq T_i \end{cases}$, where $T_i$ is a threshold point satisfying $T_i < T_{i+1}$, $i = 1, 2, \ldots$, $s$, $s$ being the number of thresholds. By allowing a shift in trend starting in year $T_i$, the variables $yr_i$ will allow us to capture possible changes in trends over time.\(^9\)

Since we want to study dynamics, we also examine the effects of lagged yields. We start with a simple autoregressive model $AR(m)$, where $y_t = \alpha_0 + \alpha_1 y_{t-1} + \ldots + \alpha_m y_{t-m} + e_t$. We introduce trend effects by letting $\alpha_0 = a_0 + b_0 yr_0 + b_1 yr_1 + \ldots + b_s yr_s$. In this context, we use the
Bayesian Information Criterion (BIC) to choose the number of lags \( m \) and the number of threshold points \( s \). The BIC criterion indicated choosing two lags: \( m = 2 \). With respect to choice of \( s \), the BIC criterion suggested two threshold points \( (s = 2) \) with optimal threshold \( T_1 = 1948 \) and \( T_2 = 2009 \). We deemed this choice undesirable: it would have too few observations (only 4 observations) to support the analysis of the regime where \( t \geq T_2 \). On that basis, we restricted our analysis to the case where \( s = 1 \). In this context, the BIC criterion selected the threshold \( T_1 = 1941 \). On that basis, the rest of our analysis will rely on two lags \( (m = 2) \) and one threshold point \( (s = 1 \) and \( T_1 = 1941 \), with \( yr_1 = \begin{cases} 0 \\ \text{year-1941} \end{cases} \) when year \( \begin{cases} < \\ \geq \end{cases} 1941 \).

Given these choices, we first estimated an AR(2) model applied to the yield data. The results are presented in Table 1. With the exception of the coefficient of \( yr_0 \), all parameters are statistically significant at the 1 percent level. This includes the coefficient of \( yr_1 \) reflecting a change in yield trend starting in 1941. This also includes the coefficients of \( \text{yield}_{t-1} \) (0.382) and of \( \text{yield}_{t-2} \) (0.302), indicating the presence of significance dynamics. Note that the estimated AR(2) process is stationary: the roots of the associated A matrix are 0.773 and -0.391 (both located in the unit circle).

To check the validity of the AR(2) model, we tested for possible serial correlation in its residuals. Using a Breusch-Godfrey test, the p-value for testing the absence of serial correlation was 0.462. Thus, we failed to find statistical evidence of serial correlation in the error term. This indicates that using two lags \( (m = 2) \) properly captures yield dynamics. On that basis, our dynamic analysis presented below relies on second-order difference equations \( (m = 2) \). Finally, we tested for possible heteroscedasticity in the AR(2) model. We used a Breusch-Pagan test to examine whether the variance of \( e_t \) is constant over time. We found evidence of heteroscedasticity: the p-value of the test for homoscedasticity was 0.001. The results indicate
that the variance of wheat yield has increased significantly during the sample period. In the process, we obtain an estimate $\sigma_t^2$ of the variance of $e_t$ at time $t$.

Next, we proceed to specify and estimate the TQAR model (4) applied to wheat yield. First, we introduce the trend terms $y_{r0}$ and $y_{r1}$ in (4) with $y_{r0} = (\text{year} - 1900)$ and $y_{r1} = \begin{cases} 0 & \text{when year} < 1941, \\ \text{year-1941} & \text{when year} \geq 1941, \end{cases}$, and let $\beta_{0r} = a_{0r} + b_{0r} y_{r0} + b_{1r} y_{r1}$. Second, we chose four regimes ($K = 4$) with $Y = (S_1, S_2, S_3, S_4)$ and allow the parameters of lagged variables to vary across regimes. Given the presence of trends and heteroscedasticity, we define the regimes ($S_{1t}, S_{2t}, S_{3t}, S_{4t}$) as follows. After obtaining the mean yield $\mu_t$ at time $t$ and the standardized residual $\hat{e}_t/\sigma_t$ from the AR(2) model, we define $S_{1t}, S_{2t}, S_{3t}$ and $S_{4t}$ respectively as the first, second, third and fourth quartile of the distribution of $[(\mu_t + \sigma_t (\hat{e}_1/\sigma_1), \ldots, \mu_t + \sigma_t (\hat{e}_n/\sigma_n)]$. Then, $D_j(y_{t,i})$ in (4) becomes $D_j(y_{t,i}) = \begin{cases} 1 & \text{when } y_{t,i} \in S_{j,i}, \\ 0 & \text{else} \end{cases}$, $j \in (1, 2, 3, 4)$.

4. Results

Under this specification, the TQAR model (4) was applied to wheat yield and estimated. From (3), the quantile regression estimates were obtained. Estimates for selected quantiles are reported in Table 2. The parameter estimates show that, while the overall trend $y_{r0}$ is not statistically significant, the trend $y_{r1}$ has a positive and statistically significant effect on wheat yield for most quantiles. This reflects large productivity gains in wheat technology over the last 70 years. The estimates also show the presence of dynamics as lagged effects can vary across quantiles and across regimes (as further discussed below).
A number of hypotheses were tested. First, we tested the null hypothesis that the regression parameters $\beta_{ijr}$’s are the same across nine quantiles: $r = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8$ and $0.9$. The Walt statistic for this test was 4.612. With 80 degrees of freedom, the p-value is 0.001. Thus, we find strong statistical evidence that the regression parameters vary across quantiles. This result implies a rejection of the AR(2) model (which assumes that the autoregression parameters are constant across quantiles). It also implies the presence of asymmetric dynamics across quantiles. The nature of this asymmetry is further evaluated below.

Second, we investigated whether dynamics is the same across regimes ($S_1, S_2, S_3, S_4$), each regime reflecting where previous states ($y_{t-1}$ and $y_{t-2}$) were located. We tested the null hypothesis that the parameters $\beta_{ijr}$ are the same across regimes $j = 1, 2, 3, 4$, for all $i \in \{1, 2\}$. This hypothesis is tested using a Wald test for selected quantiles $r \in (0, 1)$. The test results are presented in Table 3. They show strong statistical evidence of differences across regimes $S_j$’s in the lower tail of the distribution: the p-value of the test is 0.001 when $r = 0.1$. This result imply a rejection of the QAR model (which assumes that each regression quantile is linear in the lagged variables) in favor of the more flexible TQAR model. There is also some evidence of differences across regimes when $r = 0.6$ (with a corresponding p-value of 0.084). But Table 3 shows no statistical evidence of differences across regimes for other quantiles. Thus, much of the differences in dynamics across regimes are found in the lower tail of the distribution. These effects are further evaluated below.

Next, we use the estimated quantile model to evaluate the evolution of the distribution of wheat yield over time. The analysis presented below is based on quantile regression applied for all quantiles where the distribution function has a jump. This provides a flexible representation of the distribution of yield capturing how uncertain factors affect agro-ecosystem productivity.
(e.g., the effects of weather shocks or unpredictable pest damages on yield). Of special interest will be the implications of our TQAR model estimates for agroecosystem dynamics and resilience.

The predicted distributions are presented in Figure 2 for selected years. Figure 2 shows a large and steady shift in the yield distribution during the last 70 years. It also shows that the distribution of yield exhibits a larger upper tail and lower tail in 2000. This likely reflects the more recent effects of climate change on the variability of agricultural productivity. As such, our econometric approach provides useful information on how climate change is affecting risk exposure in agriculture.

Figure 3 presents the evolution of quantiles of wheat yield, \( Q(r \mid \cdot) \) for selected values of \( r \). Again, this illustrates the large increases in agricultural productivity over the last few decades. We also examine “relative quantiles” defined as \( Q(r \mid \cdot)/Q(0.5 \mid \cdot) \), i.e. as the ratio of the \( r \)-th quantile yield divided by median yield. Figure 4 presents the evolution of relative quantiles of wheat yields for selected values of \( r \). The documented changes in the distribution of relative yield over time reflect yield dynamics. In particular, Figure 4 indicates that, over the last decade, climate change may have contributed to an increase in downside risk (defined as the risk located in the lower tail of the yield distribution).

To examine the nature of dynamics, the Eigenvalues of the A matrix (where \( A(z_{t-1}, e_t) \equiv \partial h(z_{t-1}, e_t)/\partial z_{t-1} \) in equation (2)) are evaluated using the quantile estimates. The dominant roots of A are reported in Figures 5. Figures 5 show the variations in the domain root across quantiles and across regimes. Over all, the dominant root varies between 0.6 and 1.1. The dominant root is the largest (in the range 1.0 to 1.1) in the lower tail of the distribution (when \( r \) is less than 0.1) when \( y_{t-1} \) is in regime \( S_1 \). A dominant root greater than 1 indicates the presence of locally explosive
behavior. It means that wheat yield tends to escape from the lower tail of the distribution. This finding can be linked to the concept of resilience. Resilience has been defined as the ability to recover from or adjust easily to adverse shocks (Holling; Gunderson). Our quantile results provide evidence that wheat production is resilient: after facing adverse shocks, dynamics is such that yield tends to escape over time from the lower tail of the distribution.

Figures 5 show that the dominant roots are found to be less than 1 when \( r > 0.1 \) for all regimes. This suggests that stable dynamics exists everywhere outside the lower tail of the distribution. As such, the analysis does not find any evidence of multiple state states (see Gunderson). Interestingly, the dominant root does not exhibit a monotonic pattern with respect to quantiles \( r \). The dominant root is found to be in the range \( 0.9-1 \) when \( r = 0.95 \). The next highest values for the dominant root are in the range \( 0.8-0.9 \) around the median (with \( r \) between 0.4 and 0.6). Finally, the lowest value of the dominant root is around 0.7 when \( r \) is either between 0.15-0.3 or around 0.8. It means that the persistence in yield dynamics is lowest for these quantiles.

Finally, Figures 5 show modest changes in the distribution of the dominant root across regimes (\( S_1, S_2, S_3, S_4 \)), except in the lower tail of the distribution (where the changes across regimes are larger). This is consistent with the test results presented in Table 3. It illustrates that differences in dynamics across regimes are mostly present in the lower tail of the distribution. It identifies the presence of significant nonlinear dynamic adjustments in the presence of unfavorable events. This stresses the role of nonlinear dynamics, especially related to downside risk, i.e. to the occurrence of risky events located in the lower tail of the distribution.
5. Concluding Remarks

This paper presented the specification and estimation of a Threshold Quantile Autoregressive (TQAR) model, with an application to the dynamics of wheat yield in England. The model allows dynamics to vary across quantiles (a property shared with QAR). It is more flexible than the QAR model as it allows dynamics to vary across regimes defined according to the values taken by lagged variables. Our specification allows for multiple regimes across multiple lagged variables. As such, it provides a very flexible representation of dynamics and asymmetric adjustments. It allows for nonlinear dynamics in a parsimonious way that makes it attractive for empirical analysis. Of special interest are the static and dynamic effects of adverse shocks, i.e. of shocks located in the lower tail of the yield distribution. This is particularly relevant in the analysis of resilience issues.

The econometric analysis is applied to a reduced form specification of yield dynamics. The estimates of the TQAR model allows us to document the presence of significant dynamics and slow adjustments to shocks. They provide statistical evidence against both AR and QAR models. This indicates that our proposed model provides a more flexible representation of wheat yield dynamics. To the extent that unforeseen weather events are the main source of yield risk, our investigation provides useful insights on yield adjustments to climate change. The results document the evolution of the distribution of wheat yield. They show how climate change may have been contributing to increasing risk exposure in agriculture over the last few decades. The results indicate the presence of instability in the lower quantile. This result indicates the presence of resilience: after facing adverse shocks, dynamics is such that wheat yield tends to escape over time from the lower tail of the distribution.
The analysis presented in this paper provides an illustration of how a TQAR model can be used to evaluate yield dynamics. Note that it could be extended in several directions. First, the TQAR model could be applied to the analysis of other dynamic processes (in economics or biology). For example, it could provide useful information on ecological dynamics in the presence of climate change. Second, while our reduced form approach is empirically convenient, it does not provide direct estimates of structural information related to yield dynamics. This would allow addressing management issues. There is a need to explore applying TQAR models to structural equations. Third, our proposed model defines exogenous regimes for lagged variables. It would be useful to evaluate TQAR dynamics under endogenous regimes. These appear to be good topics for future research.
References


Table 1: Parameter Estimates of AR(2) process for Wheat Yield$^a$

<table>
<thead>
<tr>
<th>Variable</th>
<th>Parameter estimate</th>
<th>Standard error</th>
<th>t-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.734**</td>
<td>0.221</td>
<td>3.321</td>
<td>0.001</td>
</tr>
<tr>
<td>$y_{t0}$</td>
<td>0.001</td>
<td>0.002</td>
<td>0.506</td>
<td>0.613</td>
</tr>
<tr>
<td>$y_{t1}$</td>
<td>0.025**</td>
<td>0.008</td>
<td>2.868</td>
<td>0.004</td>
</tr>
<tr>
<td>yield$_{t-1}$</td>
<td>0.382***</td>
<td>0.092</td>
<td>4.157</td>
<td>0.001</td>
</tr>
<tr>
<td>yield$_{t-2}$</td>
<td>0.302**</td>
<td>0.091</td>
<td>3.307</td>
<td>0.001</td>
</tr>
</tbody>
</table>

$^a$ Asterisks indicate the significance level: * at the 10 percent level, ** at the 5 percent level, and *** at the 1 percent level.
Table 2: Parameter Estimates of Wheat Yield for Selected Quantiles$^a$  

<table>
<thead>
<tr>
<th>Variable</th>
<th>$r = 0.1$</th>
<th>$r = 0.2$</th>
<th>$r = 0.5$</th>
<th>$r = 0.8$</th>
<th>$r = 0.9$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.455</td>
<td>0.736*</td>
<td>0.634*</td>
<td>1.126*</td>
<td>1.101*</td>
</tr>
<tr>
<td>$yr_0$ (overall trend)</td>
<td>0.001</td>
<td>0.001</td>
<td>-0.001</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td>$yr_1$ (trend for year &gt; 1941)</td>
<td>0.020</td>
<td>0.029*</td>
<td>0.026*</td>
<td>0.037*</td>
<td>0.032*</td>
</tr>
<tr>
<td>$yield_{t-1} \times D_1(y_{t-1})$</td>
<td>0.373</td>
<td>0.170</td>
<td>0.120</td>
<td>0.072</td>
<td>0.341</td>
</tr>
<tr>
<td>$yield_{t-1} \times D_2(y_{t-1})$</td>
<td>0.276</td>
<td>0.171</td>
<td>0.114</td>
<td>0.112</td>
<td>0.350</td>
</tr>
<tr>
<td>$yield_{t-1} \times D_3(y_{t-1})$</td>
<td>0.315</td>
<td>0.150</td>
<td>0.131</td>
<td>0.108</td>
<td>0.325</td>
</tr>
<tr>
<td>$yield_{t-1} \times D_4(y_{t-1})$</td>
<td>0.371*</td>
<td>0.215</td>
<td>0.154</td>
<td>0.144</td>
<td>0.343</td>
</tr>
<tr>
<td>$yield_{t-2} \times D_1(y_{t-2})$</td>
<td>0.334</td>
<td>0.426*</td>
<td>0.666*</td>
<td>0.516*</td>
<td>0.306*</td>
</tr>
<tr>
<td>$yield_{t-2} \times D_2(y_{t-2})$</td>
<td>0.362</td>
<td>0.418*</td>
<td>0.605*</td>
<td>0.474*</td>
<td>0.291*</td>
</tr>
<tr>
<td>$yield_{t-2} \times D_3(y_{t-2})$</td>
<td>0.324</td>
<td>0.396*</td>
<td>0.617</td>
<td>0.499*</td>
<td>0.335*</td>
</tr>
<tr>
<td>$yield_{t-2} \times D_4(y_{t-2})$</td>
<td>0.351</td>
<td>0.399*</td>
<td>0.562</td>
<td>0.469*</td>
<td>0.294*</td>
</tr>
</tbody>
</table>

$^a$ The dummy variable $D_i(y_{t-k})$ equals 1 when $y_{t-k}$ is in the i-th quartile of the yield distribution, $i = 1, 2, 3, 4$. Asterisks indicate significance at the 5 percent level.
Table 3: Testing the constancy of parameters across regimes $S_1$, $S_2$, $S_3$ and $S_4$.\(^a\)

<table>
<thead>
<tr>
<th>Quantile</th>
<th>Wald test statistic</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r = 0.1$</td>
<td>16.861</td>
<td>0.001</td>
</tr>
<tr>
<td>$r = 0.2$</td>
<td>1.529</td>
<td>0.175</td>
</tr>
<tr>
<td>$r = 0.3$</td>
<td>0.517</td>
<td>0.794</td>
</tr>
<tr>
<td>$r = 0.4$</td>
<td>1.309</td>
<td>0.259</td>
</tr>
<tr>
<td>$r = 0.5$</td>
<td>1.038</td>
<td>0.405</td>
</tr>
<tr>
<td>$r = 0.6$</td>
<td>1.912</td>
<td>0.084</td>
</tr>
<tr>
<td>$r = 0.7$</td>
<td>2.087</td>
<td>0.060</td>
</tr>
<tr>
<td>$r = 0.8$</td>
<td>1.520</td>
<td>0.178</td>
</tr>
<tr>
<td>$r = 0.9$</td>
<td>0.802</td>
<td>0.571</td>
</tr>
</tbody>
</table>

\(^a\)/The variable $y_{t-k}$ is in regime $S_i$ when $y_{t-k}$ is in the $i$-th quartile of the yield distribution, $i = 1, 2, 3, 4$. 
Figure 1: Wheat yield in England (t/ha), 1885-2012.
Figure 2: Evolution of the distribution of wheat yield over time.
Figure 3: Evolution of quantiles for wheat yield over time.
Figure 4: Evolution of relative quantiles of wheat yield over time (relative to median yield).
Figure 5: Dominant root of $A^a$
a/ The variable $y_{t-k}$ is in regime $S_i$ when $y_{t-k}$ is in the $i$-th quartile of the yield distribution, $i = 1, 2, 3, 4$. 
Footnotes

1 For a comprehensive analysis of different drivers to transitional dynamics, see Bretschger and Pittel.

2 According to Levin et al., resilience should be the preferred way to think about sustainability.

3 Note that much research has been conducted on agricultural productivity and its determinants. For example, the effects of weather shocks on crop yields has been investigated by Lobell and Field, Lobell et al., Nelson et al., Robertson et al., Schlenker and Roberts, Tack et al., Willenbockel, and others. But these analyses have typically been conducted in a static context and do not provide information about dynamic adjustments or resilience. To our knowledge, the only exception is the study of Di Falco and Chavas on the resilience of cereals agroecosystem using dynamic panel data.

4 Previous literature on threshold autoregression has often focused on two regimes defined according to a single lagged variable (e.g., Tong; Cai and Stander; Galvao et al.). By allowing for multiple regimes across multiple lagged variables, our specification imposes fewer restrictions on dynamics.

5 This is the assumption made in Markov representations of dynamic processes (e.g., Billingsley; Meyn and Tweedie).

6 Note that there no loss of generality to assume that the $e_t$’s in (3a) are independently distributed over time since any serial dependency can be captured by elements of $z_t$ and their associated dynamics in (2a)-(2b).

7 The case of discrete distribution is a special case. When $w_t$ can take $s$ possible values ($a_1$, ..., $a_s$), the transition probability from $w_{t-1} = a_i$ to $w_t = a_j$ is $P(i, j) = \text{Prob}[w_t = a_j | w_{t-1} = a_i, w_t = h(w_{t-1}, e_t), i, j \in J \equiv \{1, ..., s\}]$. Letting $p_{ji} = \text{Prob}[w_t = a_j], j \in J$, the dynamics is then represented by the Markov chain: $p_{ji} = \sum_{i \in J} p(i, j) p_{i, t-1}$, $j \in J$ (Billingsley; Meyn and Tweedie).

8 In addition, under a deterministic dynamic system, the largest Lyapunov exponent defined as $\text{LE}(z_0) = \lim_{t \to \infty} \frac{1}{t} \sum_{i=1}^{t} \ln(|\lambda_i(z_i)|)$ measures the rate of expansion of the state trajectory.
along a forward path starting at $z_0$. Under some regularity conditions, a deterministic system
is asymptotically stable starting at $z_0$ if $LE(z_0) < 0$; and it is chaotic if $LE(z_0) > 0$.

While these trend effects apply within the sample period, we do not assume that they would
necessarily hold beyond the sample data.

The BIC criterion is to choose the model specification that minimizes $BIC = -2\ln(L) + k \ln(n)$, where $L$ is the likelihood function of the estimated model, $k$ is the number of
parameters and $n$ is the number of observations (Schwarz).

We treat the four regimes ($S_1$, $S_2$, $S_3$, $S_4$) as given. This simplifies the analysis and avoids
identification issues between the parameters and the thresholds used to define regimes. Note
that the case of quantile regression with unknown threshold points is analyzed by Oka and
Qu and Galvao et al.