Commodity market analysis under rational expectations: the case of the world cocoa market

Andrea Bonfatti

October 4, 2012

Abstract

In this work we specify a model for the global market of a perennial crop, namely cocoa, under linear rational expectations. One feature common to primary commodity markets is the lagged supply response by producers to market signals. This is most true for perennial crops where the biological gestation lag further accentuates such characteristics. As a result, given the low demand elasticity, stocks fluctuate widely in response to supply shocks and a fairly stable relationship between stock levels and world price can be observed in the market.

Such facts lead us to place emphasis on the derivation of the supply equation and to focus on the role of stocks, by including a speculative stock-holding equation in the structural model and deriving a dual solved or reduced form in price and stocks. The solution method follows Gilbert (1995) in reducing the state variables to a restricted set including, along with world income, interest rate and exchange rate, two constructed variables representing short and long run market imbalances. The basic intuition is that agents will be willing to hold stocks if excess supply is positive in the short term but not in the long term too.

The rational expectations hypothesis (REH) generates a set of nonlinear cross equation restrictions, linking structural and reduced form parameters, which provide indirect tests of the REH and in general of model specification. We use such restrictions to recover the structure of the model and carry out a comparative statics analysis in order to assess the model predictions about the effects of changes in the state variables on price and stocks.

The reduced form in price and stocks from a short run version of the REH model, motived by a multivariate dynamic analysis of the series, is estimated using yearly data on the world cocoa market. The estimated coefficients present theoretically consistent signs, as derived in the qualitative analysis, and meaningful magnitudes, apparently confirming the hypothesized effect of market fundamentals on current price and stocks.
Conversely, the derived estimates of the stock elasticities are not satisfactory suggesting further investigation in this direction. The single and the overall set of restrictions are not rejected, suggesting an acceptable model specification and the validity of the model consistent expectations. At this regard, we remark that the failure at rejecting the REH restrictions using estimates from a short-run model seems to imply that the information set used by market participants in forming expectations about future market developments is essentially limited in time. Our initial guess that past events, mostly supply related, would be incorporated in expectations does not seem to be supported by empirical evidence.

*Keywords:* Commodity markets; Structural models; Rational expectations.
1 Introduction

The analysis of primary commodity markets is relevant in many respects as: a) most agricultural commodities constitute basic staple food for populations worldwide; b) they still represent basic inputs in industrial production processes; c) primary commodities are the major source of export revenues of many developing countries and on their production depends the livelihoods of millions of smallholders; d) they have become increasingly important as a financial asset as commodity prices display a countercyclical pattern that make them effective in a portfolio strategy.

One feature displayed by commodity markets is the lagged response by economic agents to market signals even in time series with annual periodicity, as observed by Lord (1991). This is most true for perennial crops, such as cocoa, coffee, tea or rubber, cultivated in developing countries where, in addition to the biological gestation lag, market imperfections and domestic policies generate adjustment costs which cause suppliers to fully react to price changes only after several years. The delayed adjustments may generate temporary, even though not necessarily short-lived, market disequilibria from long term equilibrium relationships, resulting in quite long price cycles.\footnote{Labys (2006) identifies price cycles, their amplitude and duration dependence, for some 21 primary commodities using monthly data over the period 1960-1995. For cocoa and coffee he finds 15 and 16 price cycles, with a mean overall duration of about 25 and 27 months, respectively. In the case of cocoa, the maximum duration in contraction phases is 67 months, 134 for coffee, the longest among all investigated commodities.}

In fact, because of the low price elasticity of world demand, the dynamics of commodity prices is mainly driven by supply shocks, with stocks fluctuating widely with a cyclical pattern, the cycles coinciding with those of production. Such a close correspondence between stocks and price movements can be observed in Figure 4, which shows the pattern of world cocoa price and the stocks-to-use ratio, an indicator of cocoa availability monitored by industry analysts. Stocks of course may play a role of price stabilizers, absorbing excess supply, and that was the goal of the buffer stocks operating under several International Cocoa Agreements, even though they are not effective in case of price spikes due to stockouts.

Modelling of commodity markets has a long history. Traditional world commodity models (Akiyama and Duncan, 1984; Ghosh et al., 1987; Trivedi, 1990) typically comprise: a) blocks of domestic supply and demand equations, geographically disaggregated; b) price transmission equations linking local prices to a world reference price; c) a price equation relating price to world stocks or a proxy for world supply market-balance. Such models are typically estimated equation by equation. The tea model developed by Trivedi (1990) is a relevant variant as the world price is determined within a separate rational expectations (RE) model including a stockholding equation. The RE solution is used in an estimable price equation obtained by the inversion of the stock equation.

Some models focused on speculative stockholding and set price as a function of a single state variable, total availability, equal to production plus carry-
overs, trying to address through dynamic programming the non linearity arising from the non-negativity constraints related to possible stockouts (Deaton and Laroque, 1992; Wright and Williams, 1989; Peterson and Tomek, 2005).

Other authors stressed the importance of market imbalances and the role of quantity variables for modelling expectations (Hwa, 1985; Ghosh et al., 1987; Gilbert and Palaskas, 1990). In particular, Gilbert and Palaskas (1990), unlike traditional RE models, regress price changes on expected quantity variables. The rationale is that, as argued by Ghosh et al. (1987), expected future price changes are in general uninformative, since arbitrated to equal storage cost less convenience yield; rather, expected future supply-demand balances may provide an indication of the direction in which the price must adjust if the market clears over time. Though, as adjustments will eliminate such imbalances, the relevant expected future imbalances are those calculated at a reference price.

Gilbert (1995), along this stream of literature, reduces the information set used in a model for the aluminium market to a limited number of state variables, called market fundamentals, capturing excess supply (or demand) in the short and long-term. The basic intuition is that agents will be willing to hold stocks if excess supply is positive in the short term but not in the long term too. The model allows to obtain predicted values for stocks and prices and to test the rational expectations hypothesis in stockholding behaviour, conditional on the validity of the model.

In this work we build on this latter contribution, specifying an aggregate global model for a perennial crop, namely cocoa, accounting for speculative stockholding behaviour, given the historical importance of physical stocks in the cocoa market. In the model specification we proceed in two steps. First, we set up a rather general ARDL model, which allows to properly deal with the lagged response and is meant to represent an encompassing specification, containing several features useful to modelling perennial crop markets in different frameworks. In fact, depending on the specific modelling purposes and data availability, sensible economic restrictions can then be imposed. In the second step, we derive a stripped-down version of the model and adopt a solution method which, in a rational expectations framework, makes use of the concept of market imbalance (Gilbert, 1995). In deriving a solved or reduced form in price and stocks, we generalize the definition of market fundamental accounting for lags in supply response characterizing perennial crops. A meaningful research question is in fact whether market participants anticipate future market imbalances by including past (typically supply) shocks into their information set, as we would expect in a rational expectations framework.

Moreover, by manipulating the slope coefficients of the solved form, a set of cross-equation restrictions stemming from the rational expectations hypothesis are derived. The restrictions allow us in turn to recover the structure of the model (Appendix B) and to carry out a comparative statics analysis (Appendix C), aimed to predict the effect of changes in the state variables on price and

\footnote{For instance, the ARDL specification proves convenient as it can be readily written in ECM form, particularly suitable to describe long-term equilibrium relationships and the related adjustment processes.}
stocks. The qualitative analysis of the model includes also, in a deterministic setting, a simulation of price and stock response to shocks hitting the system, using the price rational expectation solution and different combinations of values for the structural parameters.

As regards the empirical part, we first provide a consistent estimate of the solved form in price and stocks previously derived, making use of the short and long term market fundamentals, and test the restrictions stemming from the rational expectations hypothesis. Second, we compare these estimates with those obtained from time series models, namely vector autoregressions (VAR), using the same dependent variables.

As to the first goal, the reduced form is estimated using the generalized method of moments (GMM) because of the endogeneity of the market fundamentals. Since we are estimating a short-run model, the REH restrictions provide a test of whether market participants (specifically stockholders) form their expectations rationally, or consistently with the postulated model, using an information set which is in fact limited in time, apparently neglecting events far in the past. We then estimate another version of the model where a time dummy is included to account for a structural break detected investigating the statistical properties of the series.

As to the second objective, one of the most important critics to the traditional approach to econometric modelling, as exemplified by the Cowles Commission works, which motivated the introduction and widespread adoption of VAR modelling within the London School of Economics methodology, concerned the insufficient attention paid to the statistical model, resulting in an inaccurate specification of the system dynamics and unjustified exclusion restrictions (Sims et al., 1990). Therefore, we present different VAR models, moving from simple a-theoretical to more structured specifications where exogenous variables from the solved form are included. A comparison between the GMM estimates of the reduced form and those from a restricted VAR model with a matching specification is then offered.

This exercise must be considered an attempt towards a comparison of full information versus reduced form models which make use of only price data or a very limited number of state variables (Wright and Williams, 1989; Deaton and Laroque, 1992).

The document is organized as follows. Section 2 presents the general specification of the model, where a particular emphasis is placed on the supply side given the peculiar features characterizing perennial crop production. Section 3 illustrates the analytical derivation of the solved form, through the construction of the market fundamentals, and of the set of restrictions stemming from the rational expectations hypothesis. Section 4 provides, in a deterministic setting and using a short run model, an overview of the conditions determining the dynamic stability of the system and the simulation of price and stock response to shocks hitting the system. Section 5 provides a brief description of the world

---

An alternative interpretation is that, for instance, the current or lagged price already incorporates past events having future impact on market developments, as we would expect in a rational expectations framework.
cocoa market. The dynamic characteristics of the series used in the model for a proper dynamic specification are investigated in Section 6. Section 7 presents the GMM estimates of the reduced form, the results from different VAR models and a comparison between the GMM estimates and those from the restricted VAR with a matching specification. Section 8 reports the concluding remarks.

2 Modelling perennial crop markets

2.1 Production

Perennial crops, such as cocoa, display typical features which make the traditional neoclassical approach not suitable to modeling supply response, namely: a) the existence of a biological gestation lag between planting and obtaining yield during which supply conditions may change; b) the bearing of adjustment costs related to the removal and planting of trees; c) the fact that the productivity of trees varies systematically with age; d) the heterogeneous nature of the tree stock, since the age-yield profile and productive life depend on technical change and hence are not invariant with respect to the date (vintage) of the investment.

The heterogeneous nature of capital stock and the time-varying productivity have been traditionally dealt with by making use of the concept of potential output \( q^p_t \), given by the sum, over all mature vintages \( v \), of the tree stock \( K(t, v) \) still in production multiplied by yields \( \delta(t, v) \) associated to each specific vintage:

\[
q^p_t = \sum_v \delta(t, v) K(t, v), \quad \forall v.
\] (1)

Potential output represents a vintage production function under the assumption of a constant coefficients technology, where variable inputs are combined in fixed proportions to capital stock of different vintages.

Data on stock composition and age-yield profiles are often lacking and therefore simplifying assumptions are usually adopted. For instance, if yields are assumed to depend only on age and are positive from age \( k \) through age \( m \), the second equality in equation (2) holds

\[
q^p_t = \sum_v \delta_{t-v} K_{t-v} \approx \sum_{i=k}^m \delta_i K_{t-i} \approx \delta \sum_{i=k}^m K_{t-i}
\] (2)

and if we further assume that yields are uniform over time, we get the third equality. These are of course strong assumptions as we are in fact ruling out technical change and the possibility of accounting for capital heterogeneity.

In order to improve production capacity the farmer has the option of either planting trees on land previously uncultivated or allocated to different crops, or replacing existing aging stands with new trees. As argued by (Ruf et al., 2004), new planting and replanting decisions are qualitatively different, as the

\footnote{We assume a constant density of planting \( d \) so that area planted \( A(t, v) = dK(t, v) \) represents a good proxy for the tree stock.}
slash-and-burn system does not imply the additional opportunity cost of investment of foregoing a present, though declining, output. These adjustment costs, associated to the partial irreversibility of investment, liquidity constraints and uncertainty about future economic conditions may result in inaction bands, where farmers do not respond to price changes, making the standard neoclassical theory of investment not suitable for explaining investment decisions typical of a mature industry (Hill, 1996).

These considerations motivated the attempts of modelling the two investment decisions separately (French and Matthews, 1971; Hartley et al., 1987; Akiyama and Trivedi, 1987), even though data availability has severely limited the widespread use of such modeling strategy.\(^5\)

Therefore, if we are forced to neglect the age composition of the tree stock, the evolution of total productive land \(A_t\) can be written as

\[
A_t = A_{t-1} + M_t - U_t + \xi_t
\]  

(3)

where \(M_t\) denotes new areas entered into production, \(U_t\) areas temporarily out of production because of uprootings of aging trees aimed at replantings or permanently lost to that specific crop production, while \(\xi_t\) represents any stochastic shock negatively affecting the tree stock such as frosts, fires and deseases.

Since data on uprootings, plantings or replantings are hardly available at aggregate level, coming usually from ad hoc household surveys, we follow Mehta and Chavas (2008) specifying a simplified law of motion

\[
A_t = \vartheta A_{t-1} + N_t
\]  

(4)

where \(N_t\) denote net additions to productive areas and \((1 - \vartheta)\) is the annual depreciation rate.

In deriving the supply equation of the global model we proceed by specifying a structural model of supply along the lines of the seminal work by Wickens and Greenfield (1973) on the coffee market in Brazil. Their supply model comprises a vintage production function, given by potential output as defined above, an investment equation for new plantings and a harvesting equation relating actual to potential output, aimed at capturing short-run supply response.

The optimal level of investment \(I^*_t\) in their model is derived by maximizing the discounted flow of expected net revenues from the investment subject to the production function (1)

\[
V_t = \sum_{t=0}^{\infty} \beta^t \left[ (p^e_t - s^e_t)q^p_t - F_t - g(I_t) \right]
\]  

(5)

where \(s^e_t\) is the expected unit cost of harvesting, \(F_t\) are fixed costs, \(g(I_t)\) is a nonlinear function representing planting costs and \(\beta\) is the discount factor. By assuming a quadratic function for \(g(I_t)\), the time path solution for \(I_t\) is a linear

\(^5\)The use of the Kalman filter in models cast in state-space form tries to overcome the shortcomings related to lack of statistical data (Kalaitzandonakes and Shonkwiler, 1992).
function of discounted expected net revenues $R_t^c$

$$R_t^c = \sum_{i=0}^{\infty} \beta^i \delta_i (p_{t+i}^c - s_{t+i}^c)$$

(6)

where $\delta_i$ are age-specific yields and the unobservable variables are proxied by distributed lags of prices and variable input costs, if any are available. Since cocoa production is labour intensive, a suitable proxy for unit costs would be the wage rate in the agricultural sector.\(^6\) More in general, the availability of labour force has been historically an important factor in the early development of the cocoa sector in most producing countries, characterized by the slash-and-burn system.

The substitution of expected price with distributed lags of course introduces a measurement error and makes such naïve expectations not coherent with the rational expectations framework. Nevertheless, in major cocoa producing countries were operating, until the liberalization process took place in the 1990s, marketing boards and caisse systems where a guaranteed farm gate price were announced at the beginning of each crop year, eliminating uncertainty at least for the nearby harvest.\(^7\)

At this regard, we remark that the specification of an aggregated global model unfortunately prevent us from incorporating institutional features specific to single countries. This can be problematic in pre-liberalization periods where domestic policies drive a substantial wedge between producer and border prices, insulating producers from international price variability. While the price transmission issue could be better addressed in a disaggregated model via price linkage equation relating domestic producer prices to the world reference price, we may try to loosely control for domestic policies through the use of time dummies referring to the steps in the liberalization processes occurring in the main producing countries, such as Côte d’Ivoire. Furthermore, dummy variables may also allow us to control in a simple way for asymmetric supply response and inaction bands arising from the above mentioned adjustment costs and uncertain future market conditions in a liberalized framework.

The relevance of such effects is an empirical matter, and the use of an aggregated supply function might blur and hide locally relevant effects. The use of annual data and the related necessity of preserving degrees of freedom suggest parsimony in the use of such dummies, and the same rationale should apply to the choice of the appropriate lag length.\(^8\)

\(^6\) Other outlays are related to the purchase of seedlings, fertilizers and pesticides.

\(^7\) This price was calculated by subtracting to a cif price all marketing costs along the supply chain, including levies, margins and stabilization funds used to guarantee minimum producer prices in period of declining world prices. The cif price was calculated as a weighted average of the prevailing spot price and the price obtained by selling forward part of the future harvest. For this reason the futures price for harvest time delivery could be considered a good proxy of the reference price faced by producers at planting time and, in general, of the extent of the investment by the state in the cocoa sector for the incoming crop year, in terms of subsidized inputs and infrastructure.

\(^8\) While information criteria provide guidance, a relevant lag length for planting decisions could be 3 to 5 years, as this is the gestation lag before the plant starts bearing fruits, even
Therefore, a rather general distributed-lag linear specification for the investment equation, describing areas entered into production in period \( t \), can be written as

\[
N_t = \mu^N_t + \alpha_c(L)p_t + \alpha_f(L)f_t + \alpha_{cf}(L)p^f_t + \alpha_w(L)w_t \tag{7}
\]

where \( f_t \) denotes the futures price of cocoa, \( p_t \) the spot cocoa price, \( p^f_t \) the spot price of substitute crops, such as coffee, and \( w_t \) is a vector including the unit cost of inputs or in general any other monetary factor affecting investment decisions. The formulation of the intercept, \( \mu^N_t = \mu^N_0 + \mu^N_1 \delta \), is general enough to include time dummies or deterministic terms. The coefficients \( \alpha_i(L) \) are polynomials in the lag operator \( L \) that in a very general formulation are given by

\[
\alpha_i(L) = \sum_{j=0}^{n} (\alpha_{ij} + \alpha_{ij}^+ D^+_i + \alpha_{ij}^B B_i + \alpha_{ij}^{B+} B_i D^+_i) L^j \tag{8}
\]

where \( D^+_i \) and \( B_i \) are dummies denoting positive price changes and policy changes, such as liberalizations, respectively.

The harvesting equation aims to explain discrepancies between potential output \( q^p_t \), as a result of past investments, and actual production due to short run supply response to current and lagged cocoa prices, resulting in more intense applications of variable inputs such as pesticides and fertilizers, and possibly to other exogenous factors \( w_t \), and is given by

\[
q_t = \mu^h_t + \gamma p^p_t + \gamma_c(L)p_t + \gamma_w(L)w_t + \gamma_q q_{t-1} + \xi^h_t \tag{9}
\]

where one period lagged production \( q_{t-1} \), while balancing the equation, may capture possible autocorrelation arising from shocks affecting the tree stock due to pest and deseases. The idea is that the permanent component of the composite shock is captured by lagged production so that the disturbance term \( \xi^h_t \) can be assumed to be an iid process, reflecting weather events.

Apart from using Kalman-filtering techniques, structural estimation of the supply system, given by equations (9), (7) and (2), usually turns out to be problematic because of data constraints. For these reasons, we prefer to derive the reduced form of the system, as in Wickens and Greenfield (1973), which represents the model supply equation. Thus, assuming that \( q^p_t = \delta A_t \), with constant average yields, writing (4) as

\[
A_t = (1 - \theta L)^{-1} N_t \tag{10}
\]

and substituting \( q^p_t \) into (9) we get

\[
q_t = \mu^q_t + \beta_q(L)q_t + \beta_c(L)p_t + \beta_{cf}(L)p^f_t + \beta_f(L)f_t + \beta_w(L)w_t + u^q_t \tag{11}
\]

where

\[
\begin{align*}
\beta_{cf}(L) &= \alpha_{cf}(L)\gamma^p \delta \kappa, \quad \beta_f(L) = \alpha_f(L)\gamma_p \delta \kappa, \quad \kappa = (1 - \theta L)^{-1} \\
\beta_c(L) &= (\alpha_c(L)\gamma_p \delta + (1 - \theta L)\gamma_c(L))\kappa, \quad \beta_w(L) = (\alpha_w(L)\gamma_p \delta + (1 - \theta L)\gamma_w(L))\kappa \\
\beta_q(L) &= \gamma_q L, \quad \mu^q_t = (\mu^N_0 \gamma_p \delta + (1 - \theta L)\mu^h_t)\kappa, \quad u^q_t = \xi^h_t.
\end{align*}
\]

though 8 to 11 years pass before the plant reaches full production.
In what follows, because of data constraints, the vector $w_t$ will be substituted by the scalar $x_{2t}$ denoting the exchange rate of major producing countries.

One problem with this approach is the difficulty of identifying the structural parameters from the reduced form. Furthermore, the inclusion of a large number of explanatory variables in the investment equation implies the addition of a corresponding number of lag structures, reducing drastically the degrees of freedom.\footnote{The identification issue, along with the inadequate representation of the vintage technology, have led some authors to reject the econometric approach altogether and adopt a dynamic programming approach, gauged more appropriate to the complexity of the problem (Bellman and Hartley 1985, Trivedi 1986, Weaver 1989).}

\section*{2.2 Consumption}

Consumption demand $c_t$, measured by total world grindings, is specified as a function of a distributed lag of cocoa price $p_t$, the price of cocoa substitutes $p_v^t$, such as vegetable fats and oils, and world income $x_1t$, as measured by the weighted GDP of major consuming countries. In the general specification, lags in consumption have also been introduced to possibly account for long-term equilibrium relationships between consumption and GDP

$$
c_t = \mu_c t + \varphi_c(L)c_t + \varphi_p(L)p_t + \varphi_v(L)p_v^t + \varphi_x(L)x_{1t} + u_c^c. \tag{12}
$$

\section*{2.3 Stockholding}

Stock demand may include transaction, precautionary and speculative components. In Muth’s famous speculative inventory model the incentive to carry over an additional unit of stock is given by

$$
\frac{P^p_{t+1} - \left(1 + R_t + \delta\right)P_t + c(S_t)}{1 + R_t} \quad \text{s.t.} \quad S_t \geq 0 \tag{14}
$$

where arbitrage leads to a situation where either stocks are zero and the incentive to carry additional stocks is non-positive, or stocks are positive and the incentive is zero.

The non negativity restriction and the possibility of stockout may result in non-linear price response as an increase in production can be partially absorbed by increased stockholding, moderating the price fall, but the reverse is not true, unless sufficient carryovers from previous period are available. The resulting
non-linearities can be dealt with only by numerical approximation or simulation techniques even in fairly simple models (Lowry et al., 1987; Deaton and Laroque, 1996). Note that we ignore the non negativity constraint, assuming a positive convenience yield high enough to ensure that positive stocks are always held. Hence, our log-linear stockholding equation is

\[ s_t = \mu_s + \eta_s (L) s_t + \eta_s (p_{t+1} - p_t - r_t) + u_s^t \]  

(15)

where \( \eta_e \) is taken as a constant parameter and lower-case letters denote as before logged variables.

### 2.4 Expectations

Agents form their expectations about future market developments following the rational expectations hypothesis, so that

\[ p_{t+1} = E[p_{t+1} | \Omega_t] + \epsilon_{t+1} \]  

(16)

where \( \epsilon_{t+1} \) is an unforecastable innovation such that \( E_t \epsilon_{t+1} = 0 \).

### 2.5 Market clearing

The model is closed by the market clearing condition\(^{10} \) where total availability, given by production plus lagged carryovers \( s_{t-1} \), equals demand for final consumption and for stocks

\[ q_t + s_{t-1} = c_t + s_t. \]  

(17)

### 2.6 Exogenous variables dynamics

In a rational expectations framework, it is important to find a suitable representation of the stochastic process governing the exogenous variables dynamics as they represent the forcing variables of the system. As regards income \( x_{1t} \), the ADF test presented in Section 6 suggests a trend stationary process, while the other tests possibly indicate a random walk specification with a drift term\(^{11} \). We decided for the latter option, so that we have

\[ x_{1t} = \gamma x_{1,t-1} + w_{1t}. \]  

\(^{10} \) A few studies develop and estimate global commodity models where interestingly a futures market is introduced. Kawai (1983) derives the rational expectations equilibrium solution of both spot and futures market in a model with stockholding very close to the one developed in this work. Palm and Vogelvang (1986) estimate a short-run model for the world coffee market in which, together with futures, other prices along the supply chain are modelled. In all these cases, a clearing condition for the futures market is added to the model.

\(^{11} \) A debate is still alive in macroeconomics as to whether GDP is better represented by a random walk or a trend stationary process.
As to the exchange rate $x_{2t}$, the unit root hypothesis is accepted by all tests, though only weakly, even accounting for the structural break in 1994 corresponding to the devaluation of the Franc CFA. Therefore, we chose a random walk with drift specification given by

$$x_{2t} = \gamma_{20} + x_{2,t-1} + w_{2t}.$$  \hfill (19)

As concerns the interest rate $r_t$, the CMR test rejects the unit root while the ADF and KPSS only marginally fail to do it. Thus, we specify an ARMA(1,1) process with drift term

$$r_t = \gamma_{30} + \rho_3 r_{t-1} + w_{3t} \quad \hfill (20)$$

where the disturbance $w_{3t} = \upsilon \epsilon_{t-1} + \epsilon_t$ includes an autoregressive component. As we can see, a certain degree of arbitrariness is present and hence room is left for further experimentation.

3 The estimable model specification

We now present the model specification used for estimation, obtained from the general formulation by reducing the order of the lag polynomials of the quantity variables, eliminating prices other than spot cocoa prices and the dummies related to asymmetric price response. Such specification proves convenient as the model can be quite easily solved in terms of the market fundamentals and, since the aim is to provide an acceptable representation of price and stocks behaviour at global level, it allows to focus on the main driving variables. Conversely, we are admittedly overlooking cross price effects and institutional features, likely relevant in explaining supply response, included asymmetric price transmission along the supply chain due to domestic policies or market power issues. That would probably require the specification of a more disaggregated model.

The model consists of three behavioural equations for supply, demand and stockholding and a market clearing identity. Consumption demand $c_t$, measured by total world grindings, is postulated as a function of real current cocoa price $p_t$ and real income $x_{1t}$, given by the weighted GDP of major consuming countries

$$c_t = \mu_c + \varphi_p p_t + \varphi_x x_{1t} + u_c^c. \quad \hfill (21)$$

In the short run model supply is specified as a function of previous period production, the lagged cocoa price $p_{t-1}$ (we rule out contemporaneous price response) and the lagged exchange rate of major producing countries $x_{2,t-1}$

$$q_t = \mu_q + \beta_q q_{t-1} + \beta_{1p} p_{t-1} + \beta_{2x} x_{2,t-1} + u_q^q. \quad \hfill (22)$$

Stock demand linearly depends on the expected gain from carrying stocks into the next period, given by the difference between expected $p_{t+1|t}$ and current price, net of storage costs proxied by the interest rate $r_t$

$$s_t = \mu_s + \eta_s s_{t-1} + \eta_c (p_{t+1|t} - p_t - r_t) + u_s^s. \quad \hfill (23)$$
where \( \eta_e \) is taken as a constant parameter.

Expectations are assumed to be rational, or model consistent, exploiting all information available until the present period \( \Omega_t \)

\[
p'_t |_{t+1} = E[p_{t+1} | \Omega_t]. \tag{24}
\]

The model is closed by the market clearing condition

\[
q_t + s_{t-1} = c_t + s_t. \tag{25}
\]

The equations for the stochastic processes of the exogenous variables (18)-(20) complete the model.

### 3.1 Market fundamentals

We now solve the above structural system into a solved or reduced form in price and stocks depending on a reduced set of state variables. To this end, following Gilbert (1995), we make use of two constructed variables measuring market imbalance in the short and long term. As argued before, the rationale is that agents will be willing to hold stocks if the short-term fundamental is weak, as excess supply prevails in the market, but not if the long-term fundamental is weak too.

As competitive markets always clear, we need to define the market imbalances at a reference price \( \bar{p} \). This can be thought of as the price which on average clears the market or the steady state equilibrium price.

**Proposition 1. Short term market fundamental**

Given the structural model (21)-(25), the short-term market fundamental \( z_{1t} \) is given by the excess supply calculated at the reference price \( \bar{p} \) as

\[
z_{1t} = q_t - \beta_p(L)(p_t - \bar{p}) - c_t + \varphi_p(p_t - \bar{p}) + (s_{t-1} - \bar{s}). \tag{26}
\]

We can interpret \( z_{1t} \) as the net addition to stocks were the market clearing price equal to the reference price. It is obtained from the market clearing condition (25) by adjusting supply and demand with the response to the price differential \( (p_t - \bar{p}) \). Notice how, through the use of the lag polynomial \( \beta_p(L) \), we have allowed for possible long-term supply effects on the current market imbalance. Letting \( y_{1t} = p_t - \bar{p}, y_{2t} = s_t - \bar{s} \) and using the definition of \( z_{1t} \), we can write the market clearing condition as

\[
y_{1t} = -\lambda(z_{1t} - y_{2t}) \tag{27}
\]

where \( \lambda = (\beta_p(L) - \varphi_p)^{-1} \). Such expression will prove useful in the following derivations.

**Proposition 2. Long term market fundamental**
We define the long-term market fundamental as the gap between production at the reference price \( \bar{p} \) and the trend in consumption \( \hat{c}_t \)
\[
 z_{2t} = q_t - \beta_p(L)(p_t - \bar{p}) - \hat{c}_t. 
\] (28)

The possibility of defining a steady-state equilibrium price relies eventually on the stationarity of the price stochastic process, so that shocks are ultimately transitory. On the other hand, permanent shifts in supply or demand due to technical change or sustained investment may well determine structural breaks in the series, thus changing the long-run equilibrium price. Therefore, the price to which reference is made in calculating market imbalances has to be updated according to the structural evolution of the system.\(^{13}\)

**Proposition 3. Dynamics of short term fundamental**

Substituting from equations (22)-(25) into (26), set at time \( t + 1 \), the evolution of the short-term market fundamental is derived as
\[
 z_{1,t+1} = \xi_{1t} + \theta_{12} z_{2t} + \theta_{13} x_{1t+1} + \theta_{14} y_{2,t-1} + \phi_{12} y_{1,t-1} + \epsilon_{1,t+1} 
\] (29)

where
\[
 \xi_{1t} = \beta_p(L)\bar{p} - \varphi_p \bar{p} + \mu^q_{t+1} - \varphi_x \gamma_{10} + \beta_x \gamma_{20} + \beta_q \hat{c}_t \\
 \epsilon_{1,t+1} = u_{2,t+1} - u_{1,t+1} + \beta_x u_{2,t+1} - \varphi_x w_{1,t+1} \\
 \theta_{12} = \beta_q, \quad \theta_{13} = -\varphi_x, \quad \theta_{14} = \beta_x, \quad \varphi_{11} = \beta_q \beta_p(L), \quad \phi_{12} = 1.
\]

**Proposition 4. Dynamics of long term fundamental**

The evolution of the long-term market fundamental is obtained in turn by substituting from equations (22)-(25) into (28) at time \( t + 1 \)
\[
 z_{2,t+1} = \xi_{2t} + \theta_{22} z_{2t} + \theta_{24} x_{2,t-1} + \varphi_{21} y_{1,t-1} + \epsilon_{2,t+1} 
\] (30)

where
\[
 \xi_{2t} = \mu^q_{t+1} + \beta_p(L)\bar{p} + \beta_x \gamma_{20} + \beta_q \hat{c}_t - \hat{c}_{t+1} \\
 \epsilon_{2,t+1} = u_{2,t+1} + \beta_x u_{2,t+1} \\
 \theta_{22} = \beta_q, \quad \theta_{24} = \beta_x, \quad \varphi_{21} = \beta_q \beta_p(L).
\]

Stacking \( z_{1,t+1} \) and \( z_{2,t+1} \) along with the stochastic processes of the exogenous variables we get the dynamic system

---

\(^{12}\)In the empirical application we shall use the fitted consumption values as a proxy for consumption trend.

\(^{13}\)Such considerations suggest using a moving average of appropriate window size in computing the reference price, as it has been done in the empirical application.
deviations from the respective reference values,\( y \)

endogenous variables\( y \) \( y \)

The linearity of the system allows to solve it with respect to the vector of the 3.2 The model solution

properties of the statistical series, as investigated in Section 6.

\[ \beta \]

where\( \beta \) to the following solution of a short-run model,

\[ (2 \times 1) \]

that is the solved price and stock equations

\[ y \]

\[ (y_{1,t-1}) \]

\[ y_{2,t-1} \]

\[ y_{1,t-1} \]

Then, write equations (27) and (23) in system form as follows

\[ z_{t+1} = \xi_t + \Phi_0 y_t + \Phi_1 y_{t-1} + \Theta z_t + \epsilon_{t+1} \] (32)

where\( z_t \) is a vector of the new state variables, \( y_t \) is the (2 \times 1) vector of price and stocks expressed in terms of deviations from the respective reference values, \( y_{t-1} \) is a (2 \times 1) vector of the lagged endogenous, \( \xi_t \) is a (5 \times 1) vector of intercepts, \( \epsilon_{t+1} \) a (5 \times 1) vector of disturbances, while \( \Theta \) (5 \times 5), \( \Phi_0 \) (5 \times 2) and \( \Phi_1 \) (5 \times 2) are matrices of intermediate parameters.

The explicit inclusion in the above system of the one-lagged vector of the endogenous variables is functional to the following solution of a short-run model, where \( \beta_p(L) \) is restricted to \( \beta_{1p} \). This choice is motivated by the dynamic properties of the statistical series, as investigated in Section 6. 3.2 The model solution

The linearity of the system allows to solve it with respect to the vector of the endogenous variables \( y_t \) to obtain

\[ y_t = a_0 + A_0 z_t + A_1 y_{t-1} + v_t \] (33)

that is the solved price and stock equations

\[ p_t = a_{10} + a_{11} z_{1t} + a_{12} z_{2t} + a_{13} x_{1t} \]

\[ s_t = a_{20} + a_{21} z_{1t} + a_{22} z_{2t} + a_{23} x_{1t} + a_{24} x_{2,t-1} + a_{25} r_t + a_{26} s_{t-1} + v_{2t} \] (35)

In order to derive an explicit representation for the matrices \( A_0 \) and \( A_1 \), first substitute (32) into \( y_{t+1} \) and take the expectation to get

\[ E y_{t+1} = a_0 + A_0 \xi_t + A_0 \Theta z_t + A_0 \Phi_0 y_t + (A_0 \Phi_1 + A_1) y_{t-1}. \] (36)

Then, write equations (27) and (23) in system form as follows
\[
\begin{pmatrix}
1 & -\lambda \\
\eta_e & 1 - \eta_e L
\end{pmatrix}
(y_t - a_0) = \begin{pmatrix}
0 & 0 \\
\eta_e & 0
\end{pmatrix} E y_{t+1} + \begin{pmatrix}
-\lambda & 0 & 0 & 0 \\
0 & 0 & 0 & -\eta_e
\end{pmatrix}
(z_t - \zeta)
\]
(37)

which in matrix form is
\[
Hy_t = JE y_{t+1} + K z_t + k
\]
(38)

where \( k = H a_0 - K \zeta \). Substituting (36) into the above modified structural system and factoring similar terms we get
\[
(H - JA_0 \Phi_0) y_t = (K + JA_0 \Theta) z_t + J(A_0 \Phi_1 + A_1) y_{t-1} + \tau_0
\]
(39)

where \( \tau_0 = k + J(a_0 + A_0 \xi_t) \). It follows that an implicit expression for matrices \( A_0 \) and \( A_1 \) can be obtained from
\[
y_t = W(K + JA_0 \Theta) z_t + WJ(A_0 \Phi_1 + A_1) y_{t-1} + W \tau_0.
\]
(40)

where \( W = (H - JA_0 \Phi_0)^{-1} \). The coefficients of the response matrix \( A_0 \) thus obtained are a combination of structural and reduced form parameters, and by manipulating it a set of restrictions stemming from the rational expectations hypothesis can be derived, as it will be shown in the next section.

Before proceeding further, the explicit representation of the \( A_0 \) matrix is provided by following the steps just illustrated. Therefore, from equation (34) we compute the rational expectations of \( E p_{t+1} \) as
\[
E p_{t+1} = a_{10} + a_{16} p_t + a_{13} \gamma_{10} + a_{14} \gamma_{20} + a_{15} \gamma_{30} + z_{2t} (a_{11} \theta_{12} + a_{12} \theta_{22}) + x_{1t} (a_{13} + a_{11} \theta_{13}) + x_{2,t-1} (a_{14} + a_{11} \theta_{14} + a_{12} \theta_{24}) + a_{11} \xi_{1t} + a_{12} \xi_{2t} + a_{15} \rho_{3t} (s_t - \bar{s}) a_{11} \phi_{12} + (p_{t-1} - \bar{p}) (a_{11} \phi_{11} + a_{12} \phi_{21})
\]
(41)

and substitute it into the modified system (37) to derive the reduced form in price and stocks (40), whose slope coefficients form the \( A_0 \) and \( A_1 \) matrices. The elements of the second column of \( A_0 \) are for instance given by
\[
a_{12} = \frac{\lambda (-a_{11} \eta_e \theta_{12} - a_{12} \eta_e \theta_{22})}{\Delta}, \quad a_{22} = -\frac{a_{11} \eta_e \theta_{12} + a_{12} \eta_e \theta_{22}}{\Delta}
\]

where \( \Delta = \eta_e (a_{16} \lambda - \lambda + a_{11} \phi_{12}) - 1 \).

### 3.3 The REH restrictions

One peculiar feature of rational expectations models is that they yield a set of non-linear cross equations restrictions linking structural and reduced form parameters.

As discussed in Gilbert (1995), a first set of restrictions verify the logical
consistency of the estimated price and stock equations with the first stage production and consumption estimates. They allow to infer the coefficients of the price equation from those of the stock equations and vice versa.

**Proposition 5. Logical restrictions**

Given the matrix $A_0$ of the reduced form slope coefficients and the parameter $\lambda$, which is a function of the price response coefficients in the structural supply and demand equations, the logical restrictions are given by

$$a_{21} = 1 + \frac{1}{\lambda}a_{11}, \quad a_{2j} = \frac{1}{\lambda}a_{1j}, \quad j = 2, 3, 4, 5. \quad (42)$$

A second set of restrictions provide a check of the consistency of the estimated price equation (34) with the rational expectation of price in the stock demand equation (23).

**Proposition 6. REH restrictions**

Given the matrix $A_0$ of the reduced form slope coefficients, the parameter $\lambda$, the logical restrictions (42) and the matrices of intermediate parameters $\Phi_0$ and $\Theta$, the rational expectations restrictions are given by

$$\frac{a_{22}}{a_{21}} = \frac{(a_{16}a_{12} + \theta_{22}a_{12} - a_{12} + a_{11}\theta_{12})}{a_{11}(a_{16} - 1)} \quad (43)$$

$$\frac{a_{23}}{a_{21}} = \frac{(a_{13}a_{16} + a_{11}\theta_{13})}{a_{11}(a_{16} - 1)} \quad (44)$$

$$\frac{a_{24}}{a_{21}} = \frac{(a_{14}a_{16} + a_{11}\theta_{14} + a_{12}\theta_{24})}{a_{11}(a_{16} - 1)} \quad (45)$$

$$\frac{a_{25}}{a_{21}} = \frac{(a_{16}a_{15} + \rho_3a_{15} - a_{15} - 1)}{a_{11}(a_{16} - 1)}. \quad (46)$$

Proofs of Proposition 5 and 6 are given in Appendix A. As shown in the appendix, an intermediate step of the derivation of the REH restrictions allows to recover the key parameter $\eta_e$, given by

$$\eta_e = \frac{a_{21}}{a_{11}(a_{16} + a_{21}\phi_{12} - 1)}. \quad (47)$$

Furthermore, the parameter $\eta_s$ can be derived through an analogous procedure as

$$\eta_s = \frac{(a_{16} - 1)(a_{11}a_{26} - a_{16}a_{21})}{a_{11}(a_{16} + a_{21}\phi_{12} - 1)}. \quad (48)$$

By using estimates of the reduced form parameters, such restrictions can be tested through Wald tests. Rejection of the restrictions implies the rejection of the rational expectations provided the model and the expectations are well specified. In fact, the rejection may well be regarded as a rejection of the model.
conditional on the validity of the REH.

Furthermore, as shown in Appendix B, the logical and REH restrictions allow to recover the structure of the model from the estimated solved form and to investigate the qualitative effects of changes in the state variables on price and stocks, as illustrated in Appendix C.

4 The dynamic analysis of the system

In this section we provide a brief description of the conditions determining the dynamic stability of the system, using a short run model where further restrictions are imposed. In particular, the price lags in the supply equation are reduced to one, $\beta_p(L) = \beta_{1p}$, and $\eta_s = 0$. The equilibrium analysis of the model is indeed important as it may permit to verify whether the theoretical model is able to generate a dynamic path consistent with the observed realization of the stochastic process of the main endogenous variable of the system.

Solving equations (21)-(25) with respect to $p_t$, we get the stochastic price difference equation

$$p_t = -\frac{\eta_e}{\eta_e - \varphi_p} p_t^e + \frac{\eta_e}{\eta_e - \varphi_p} p_{t+1}^e + \frac{\eta_e - \beta_{1p}}{\eta_e - \varphi_p} p_{t-1} + X_t$$

(49)

where $X_t$ includes the exogenous variables and all disturbances. For illustration purposes, we confine our analysis to a deterministic setting, which in a rational expectations framework is equivalent to perfect foresight, i.e. $p_{t+1}^e = p_{t+1}$. Thus, substituting the expectational variables with observed realizations we get the second order non-homogeneous difference equation

$$p_{t+1} - \phi_1 p_t - \phi_2 p_{t-1} = Z_t$$

(50)

where

$$\phi_1 = \frac{2\eta_e - \varphi_p}{\eta_e}, \quad \phi_2 = \frac{\beta_{1p} - \eta_e}{\eta_e}, \quad Z_t = \frac{\mu_t + \omega_t}{\eta_e}$$

and

$$\mu_t = -\mu_t^e + \mu_t^q - \Delta \mu_t^s$$

$$\omega_t = -u_t^e + u_t^q - \Delta u_t^s + \beta_x \Delta x_{t-1} + \eta_e \Delta r_t - \varphi_x x_{t1}$$

By the superposition principle the general solution $p_t$ of a non-homogeneous difference equation can be written as the sum of the general solution to the homogeneous equation $p_t^{(g)}$ and any particular solution to the non-homogeneous equation $p_t^{(p)}$.

A particular solution of interest, when is well defined ($1 - \phi_1 - \phi_2 \neq 0$), is the steady state equilibrium, where all stochastic variables have been set to their long-term values $\bar{Z}$

$$\bar{p} = -\frac{\eta_e \bar{Z}}{\beta_{1p} - \varphi_p}$$

(51)
It can be shown (Neusser, 2009) that this equilibrium is asymptotically stable (i.e. all solutions converge to $\bar{p}$) if and only if the following conditions are satisfied

\begin{align*}
(i) & \quad 1 - \phi_1 - \phi_2 > 0 \quad \varphi_p - \beta_{1p} > 0 \\
(ii) & \quad 1 + \phi_1 - \phi_2 > 0 \quad \text{or} \quad -\beta_{1p} + 4\eta_e - \varphi_p > 0 \quad (52) \\
(iii) & \quad 1 + \phi_2 > 0 \quad \beta_{1p} > 0.
\end{align*}

In general, the dynamic stability of the system depends on the roots of the characteristic equation

$$\lambda^2 - \phi_1 \lambda - \phi_2 = 0 \quad (53)$$

which in terms of the original parameters are given by

$$\lambda_{1,2} = \frac{\varphi_p - \sqrt{\varphi_p^2 - 4\eta_e \varphi_p + 4\beta_{1p}\eta_e}}{2\eta_e} \quad (54)$$

whose nature is determined by the sign of the discriminant $\Delta = \varphi_p^2 - 4\eta_e \varphi_p + 4\beta_{1p}\eta_e$.

It turns out that the signs of the structural parameters predicted by the theory ($\beta_{1p} > 0$, $\varphi_p < 0$, $\eta_e > 0$) are such that $\Delta > 0$, therefore we are given two real and distinct roots. In this case, the general solution to the second order non-homogeneous difference equation is given by

$$p_t^{(g)} = c_1\lambda_1^t + c_2\lambda_2^t + p_t^{(p)} \quad (55)$$

which can be written as

$$p_t^{(g)} = \lambda_2^t \left[ c_2 + c_1 \left( \frac{\lambda_1}{\lambda_2} \right)^t \right] + p_t^{(p)}. \quad (56)$$

If without loss of generality we assume that $|\lambda_2| > |\lambda_1|$, then the asymptotic behaviour of the solution depends on the value of the larger root $\lambda_2$, since $(\lambda_1/\lambda_2)^t \to 0$ as $t \to \infty$. Then, six possible cases emerge depending on the value of $\lambda_2$:

1. $\lambda_2 > 1$, $c_2\lambda_2^t$ diverges (the system is unstable);
2. $\lambda_2 = 1$, $c_2\lambda_2^t$ remains constant at $c_2$;
3. $0 < \lambda_2 < 1$, $c_2\lambda_2^t$ decreases monotonically to zero (the system is stable);
4. $-1 < \lambda_2 < 0$, $c_2\lambda_2^t$ oscillates around zero, alternating in sign, but converges;
5. $\lambda_2 = -1$, $c_2\lambda_2^t$ alternates between the values $c_2$ and $-c_2$;
6. $\lambda_2 < -1$, $c_2\lambda_2^t$ alternates in sign but diverges.
At this regard, it is immediate to verify that the stability condition (i) is not satisfied and therefore the system is unstable as can be seen also from the values of $\lambda_2$ reported in Table 1, computed for plausible ranges of values of the structural parameters, namely $\varphi_p \in [0.05, 0.4]$, $\beta_{1p} \in [0.05, 0.4]$, and $\eta_e \in [3, 2.5]$.

Since we have postulated the rational expectations hypothesis, agents are assumed to be forward looking and incorporate future developments of shocks $\omega_t$ into their decisions. Therefore, following Neusser (2009), we conjecture that a sensible particular solution can be given by

$$p_t^{(p)} = \sum_{j=-\infty}^{\infty} \psi_j \omega_{t-j}. \quad (57)$$

We proceed using the method of undetermined coefficients\textsuperscript{14} and thus substituting (57) into (50)

$$\sum_{j=-\infty}^{\infty} \psi_j \omega_{t+1-j} = \phi_1 \sum_{j=-\infty}^{\infty} \psi_j \omega_{t-j} + \phi_2 \sum_{j=-\infty}^{\infty} \psi_j \omega_{t-1-j} + Z_t. \quad (58)$$

and equating terms we get

$$\begin{align*}
\cdots
\psi_0 &= \phi_1 \psi_{-1} + \phi_2 \psi_{-2} \\
\psi_1 &= \phi_1 \psi_0 + \phi_2 \psi_{-1} + \frac{1}{\eta_e} \\
\psi_2 &= \phi_1 \psi_1 + \phi_2 \psi_0 \\
\cdots
\end{align*}$$

The coefficients $\psi_j$’s follow second order homogeneous difference equations

$$\begin{align*}
\psi_j &= \phi_1 \psi_{j-1} + \phi_2 \psi_{j-2} & j \geq 1 \\
\psi_{j+1} &= \phi_1 \psi_{j+1} + \phi_2 \psi_{j} & j \geq 1
\end{align*}$$

whose solutions are given by

$$\begin{align*}
\psi_j &= d_1 \lambda_1^j + d_2 \lambda_2^j \\
\psi_{j+1} &= e_1 \lambda_1^j + e_2 \lambda_2^j
\end{align*}$$

Table 1 shows that, for plausible values of the structural parameters, the roots are on either side of the unit circle. Therefore, if we look for a stationary solution, we have to rule out the explosive part by setting $d_2 = 0$ and $e_2 = 0$. Then, noting that the two solutions coincide for $j = 0$, which implies $d_1 = e_1$, we calculate the common constant $d$ by exploiting the fact that both solutions

\textsuperscript{14} An extensive review of different solution methods to linear rational expectations models is given in Pesaran (1987).
must satisfy the initial value condition \( \psi_1 = \phi_1 \psi_0 + \phi_2 \psi_{-1} + 1/\eta_e \). Substituting for \( \psi_1, \psi_0 \) and \( \psi_{-1} \), leads to

\[
d\lambda = \phi_1 d + \phi_2 d\lambda + \frac{1}{\eta_e} \Rightarrow d = \frac{\eta_e^{-1}}{\lambda(1 - \phi_2) - \phi_1}
\]  

(59)

where \( \lambda \equiv \lambda_1 \) and, as a consequence, \( \lambda_2 = -\phi_2 \lambda^{-1} \). The general solution to the non-homogeneous difference equation (50) is thus given by

\[
p_t = c_1 \lambda^t - c_2 \phi_2 \lambda^{-t} + d \sum_{j=-\infty}^{\infty} \lambda^j \omega_{t-j}
\]  

(60)

whereas in order to impose a boundedness condition we must set \( c_2 = 0 \) and assume that the infinite sum converges as \( t \to \infty \). In what follows, we simulate the response of price and stocks to shocks using some combinations of the structural parameters reported in Table 1.

<table>
<thead>
<tr>
<th>( \varphi_p )</th>
<th>( \eta_e )</th>
<th>( \beta_p )</th>
<th>( \Delta )</th>
<th>( \lambda_1 )</th>
<th>( \lambda_2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.05</td>
<td>0.3</td>
<td>0.05</td>
<td>0.1225</td>
<td>1.6667</td>
<td>0.5000</td>
</tr>
<tr>
<td>-0.05</td>
<td>0.3</td>
<td>0.4</td>
<td>0.5425</td>
<td>2.3109</td>
<td>-0.1442</td>
</tr>
<tr>
<td>-0.05</td>
<td>2.5</td>
<td>0.05</td>
<td>1.0025</td>
<td>1.2102</td>
<td>0.8098</td>
</tr>
<tr>
<td>-0.05</td>
<td>2.5</td>
<td>0.4</td>
<td>4.5025</td>
<td>1.4344</td>
<td>0.5856</td>
</tr>
<tr>
<td>-0.4</td>
<td>0.3</td>
<td>0.05</td>
<td>0.7000</td>
<td>3.0611</td>
<td>0.2722</td>
</tr>
<tr>
<td>-0.4</td>
<td>0.3</td>
<td>0.4</td>
<td>1.1200</td>
<td>3.4305</td>
<td>-0.0972</td>
</tr>
<tr>
<td>-0.4</td>
<td>2.5</td>
<td>0.05</td>
<td>4.6600</td>
<td>1.5117</td>
<td>0.6483</td>
</tr>
<tr>
<td>-0.4</td>
<td>2.5</td>
<td>0.4</td>
<td>8.1600</td>
<td>1.6513</td>
<td>0.5087</td>
</tr>
<tr>
<td>-0.21</td>
<td>2.35</td>
<td>0.06</td>
<td>2.5821</td>
<td>1.3866</td>
<td>0.7028</td>
</tr>
</tbody>
</table>

For simplicity, we drop the interest rate in the stockholding equation and express price exclusively as a function of past and expected future shocks by setting also \( c_1 = 0 \). The stochastic term \( \omega_{t-j} \) in equation (60) is a composite disturbance as it includes shocks arising from various parts in the system, including exchange rate, interest rate and demand shocks, which can have offsetting effects. Though, given the relevance of supply shocks in generating price cycles in agricultural markets we may focus on \( u_{t-j} \).

Assume that a unitary unexpected transitory supply shock hits the market at \( t = 0 \). As can be seen in Figure 1 (bottom part), in all scenarios the price immediately falls, while inventories (upper part) rise as prices are expected to increase in the future due to the transitory nature of the shock. Then price smoothly goes up to the previous level as stocks are gradually run down. The different simulations use values of the structural parameters typical of perennial

\footnote{The constant \( c_1 \) could have been otherwise initialized at \( p_0 \) by solving an initial value condition.}
crop markets, characterized by low short-term supply and demand elasticities. Though, we observe how in scenario S3 (solid red/black lines), characterized by a high value of the parameter $\eta$, the price fall is less pronounced than in scenario S1 (dashed red/black line), where the stock elasticity is low. A similar smooth pattern is observed also in scenario S9 (solid/dashed green line), which differs from S3 practically only for a higher demand elasticity. We may argue that a greater responsiveness of stocks to expected price gains seems to have in this model a stabilizing effect on price, a relatively larger value of the small root increasing the persistency of both price and stocks series.

Suppose now that a positive supply shock is expected to hit the market in period 0. As shown in Figure 2, at time $t-5$, when the shock is announced, market participants expect the price to fall in the future and therefore want to get rid of their stocks trying to sell them already at this time. As a result, the price starts to fall before the shock actually occurs. When finally the positive shock takes place in period 0, the excess supply sharply increases stocks, depressing price at its lowest level. Inventories are then decumulated to their initial levels, but only gradually, as agents expect the price to move up again. Beginning from period 0, the market adjusts like in the previous case as shocks are again assumed to be transitory. Similar considerations apply here concerning the different simulation scenarios. The price decrease at time 0 in scenario S1 is much more pronounced than in the other cases and the low root implies a relatively higher discounting of future events, so that adjustments occur more rapidly.

Figure 1: Unanticipated supply shock at $t=0$
5 The world cocoa market

This section gives a brief description of the cocoa market, illustrating the major past events and some recent developments occurring along the supply chain.

Cocoa production is concentrated in countries located in the equatorial belt (within 10°N and 10°S of the equator) where the climate conditions are favourable for growing the cocoa tree (Theobroma cacao). The natural habitat is the lower storey of the evergreen rainforest which provides the necessary humidity, temperature (it should be not less than 18-21 degrees C. on average) and shade. Rainfall is the main climatic factor affecting yields and must be abundant and well distributed throughout the year (preferably between 1500mm and 2000mm of annual rainfall level).

The cocoa tree flowers in two cycles of six months, yielding two harvests per year. In most African countries the main harvest lasts from October to March, while the mid harvest (typically much less abundant) from May to August. It takes from three to five years of gestation before the plant starts bearing fruits, after that yields rapidly increase to reach a peak after 8 or 11 years, depending on the varieties. Then, yields remain constant until 20-25 years before steadily declining, though the tree is productive for about 40 years.

Western and Central Africa are the most important producing regions with four countries (Côte d’Ivoire, Ghana, Nigeria, Cameroon) providing about 64% of world production (crop year 2009/2010). Côte d’Ivoire alone accounts for 34% of total supply, while other major producers include Indonesia (17%), Brazil (4%), Ecuador (4%) and Malaysia which is though recently withdrawing from cocoa production.

Despite a remarkable increase of grindings at the origin (from 33.6% in 2001/02 to about 40% in 2009/10), mostly located in Côte d’Ivoire and Malaysia (together 47% of total origin grindings), cocoa processing continues to be mainly undertaken in importing countries, notably Europe (41%), the Netherlands
(13%) being the world largest cocoa-processing country, and the United States
(10%).

As concerns demand, measured in terms of apparent consumption, given
by the sum of grindings plus net imports of cocoa products, either final or
semi-processed, converted in beans equivalent, it is concentrated in developed
countries, mainly Western Europe and North America. United States, Germany,
France and United Kingdom are, in the order, the single largest consumer coun-
tries (average 1997/98-2005/06).

Figure 3 shows the pattern of world production, grindings and stocks over
the last fifty years. After averaging at about 1.5 million tonnes until the mid-
-eighties, production went on a constant growth path, despite the prolonged
descending trend in world prices which followed the price spikes in the late sev-
enties, caused by the frosts in Brazil (see Figure 4). The sustained growth was
partly due to the investments occurred worldwide following the earlier price in-
creases, the arrival of Indonesia as a major producer and the continued support
to domestic producer prices in major producing countries in spite of the falling
world prices.\textsuperscript{16}

Though, such pricing policies became soon no longer sustainable, as the
attempts of raising world prices through the purchases of buffer stocks under
several International Cocoa Agreements turned out to be ineffective, so that
most state-owned marketing agencies went bankrupt. Under the pressure of
international donors, a liberalization process started in most African countries
leading to the progressive dismantlement of existing marketing boards or caisse
systems\textsuperscript{17}.

Despite the increasing trend in consumption, the excess supply determined
a structural break in stocks and the stocks-to-use ratio jumped to levels perma-
nently above 0.4, as can be seen from Figure 4. The stocks-to-use ratio (SUR)
is an overall indicator of world cocoa availability widely used in the industry,
as there is a fairly steady relationship between world market prices and this
ratio, as appears also in Figure 5. Apart from the shift in the intercept occurring
in the late seventies, a negative relationship between the world price and the
SUR indicator clearly emerges. This evidence motivates investigating further
the empirical relationship between stocks and prices as we try to pursue in the
next sections.

As to the industry structure, cocoa is a typical smallholder crop as almost
90\% per cent of cocoa production worldwide comes from smallholdings below 5
hectares. At the end of the 1990s, the global number of cocoa producers was
estimated at about 14 millions, about 75\% of whom in Africa. The international
trade of cocoa is instead dominated by a limited number of large multinationals.

\textsuperscript{16}In most African countries, the whole cocoa domestic chain was heavily controlled by
the government through state-owned marketing agencies. They either directly handled the
physical delivery of the produce from the farm gate to the ports (marketing board system),
or set all prices and margins along the supply chain, releasing export licences (caisse system).

\textsuperscript{17}The first African country to liberalize its cocoa sector was Nigeria in 1986-1987, followed
by Cameroon in stages, during 1989-1991 and 1995. In Côte d'Ivoire the disengagement of the
State began in 1994 and was not complete until 1999, while Ghana is still not fully liberalized.
The most remarkable events in recent years have been the processes of vertical integration along the supply chain and horizontal concentration, with mergers of large multinationals. Major international traders have in fact started to vertically integrate upstream, taking over local exporters, in order to secure supply and downstream, engaging in the first stages of cocoa processing. At present, some two-thirds of total grindings is done by the top ten firms, with three large multinationals (ADM, Cargill and Barry Callebaut) dominating the market. If the market power of such multinationals is balanced downstream by the strength of large chocolate manufacturers, the upstream integration has raised concerns of a possible abuse of buying power against a multitude of unorganized local producers.

6 The dynamic characteristics of the series

This section analyzes the statistical properties of the variables used in the model in order to identify the order of integration of the series for a correct dynamic specification. In a rational expectations framework, it is furthermore important to find a proper representation of the stochastic process of the exogenous variables, given the role they play as predictors of future dated variables.

As a preliminary operation, all price series are deflated using the US CPI, a convenient deflator used in Deaton and Laroque (2003). The same deflator is used to obtain the real income, computed as the consumption weighted GDP of the main consuming countries. Further details on definitions and data sources are given in the Appendix.

Table 2 reports the results of different unit root tests on the series trans-
Figure 4: World cocoa price and stock-to-use ratio (1)

Figure 5: World cocoa price and stock-to-use ratio (2)
formed in logarithms over the sample period 1970-2010. The ADF test (Dickey and Fuller, 1979) and the PP test (Phillips and Perron, 1988) share the same null hypothesis of the presence of a unit root, but adopt different methods to account for serial correlation. In the ADF test a sufficient number of lagged first differences is included to ensure that residuals are innovations\textsuperscript{18}, while in the PP test Newey-West consistent estimate correction for autocorrelation is implemented. Conversely, in the KPSS test (Kwiatkowski et al., 1992) the null is the trend (or level) stationarity of the series. Inference using this test is therefore complementary to that obtainable from the former tests.

Whereas the series exhibit clear upward trends, as in the case of production, consumption and real income, a linear time trend has been included in the regressions. The test statistics and the corresponding MacKinnon approximate p-values from both tests indicate that, at 5% significance level, the null of unit root is not rejected for all series except for income (ADF) and production (PP), for which the alternative of trend stationarity is accepted\textsuperscript{19}. Coherently the null of trend (level) stationarity is rejected by the KPSS test for all series, included income and production actually. All series become stationary after first differencing (results are not reported), so they appear to be integrated of order one $I(1)$.

The theoretically appealing stationarity of the price series does not emerge from the tests, as suggested also by visual inspection of the real price series of cocoa, coffee, rubber and palm oil (Figure 6). Along with the well documented comovement of commodity prices, the picture shows the long descending trend in prices started in the 1980s and partially offset by the increases of the last decade\textsuperscript{20}. In general, the combination of long price cycles, arising from the complex dynamics of perennial crops, with potentially multiple structural breaks in long annual time series, make it difficult to draw conclusive answers from unit root tests.

A problem of these tests is the lack of power in the presence of structural breaks in the series, so that the null hypothesis of unit root is overly accepted. Several tests have been devised to address this problem, where the unit root hypothesis is tested allowing for a change in the mean or trend (or both) of the series (Perron, 1989; Perron and Vogelsang, 1992; Zivot and Andrews, 1992) and where the optimal breakpoint is endogenously determined. Table 2 presents results from two of such tests, the Zivot and Andrews (ZA) test, where the unit root hypothesis is tested allowing for a single endogenously determined structural break in the series, either in the intercept or trend (or both), and the Clemente et al. (1998) test (CMR), which extends the test of Perron and Vogelsang (1992) to the case of multiple structural breaks, more precisely a double shift in the mean in the additional outlier version (AO2). In Table 2 are reported the minimum t-statistics, the 5% critical value of the left-tailed test and

\textsuperscript{18} The Pormanteau (Q) test has been applied to the ADF residuals; for almost all series a single lagged difference is sufficient to generate approximately white noise residuals.

\textsuperscript{19} Though, at 10% significance level, the ADF test does reject the unit root for the exchange rate and interest rate series.

\textsuperscript{20} Stationarity can be achieved by reducing the sample to the period 1986-2010.
the endogenous optimal break(s). The ZA test does not reject the null of unit root for all but the production series, and the same outcome hold for the CMR test, with the exclusion of the interest rate series. Moreover, both tests seem to suggest the existence of a structural break in both quantity and price series in the mid-eighties, possibly in 1986. We will try to exploit this fact using a time dummy in the following estimation.

7 Estimation methods and results

In Section 3 we illustrated how the structural model (21)-(25) can be solved, under the rational expectations hypothesis, through the use of the short and long term market fundamentals, \( z_{1t} \) and \( z_{2t} \) respectively. The solved or reduced form in price and stocks (61)-(62) is what we shall estimate in this section, whereas the original set of explanatory variables is substituted by a new vector of state variables \( (z_{1t}, z_{2t}, x_{1t}, x_{2,t-1}, r_t) \) plus the lagged endogenous

\[
p_t = a_{10} + a_{11} z_{1t} + a_{12} z_{2t} + a_{13} x_{1t} + a_{14} x_{2,t-1} + a_{15} r_t + a_{16} p_{t-1} + v_t
\]

\[
s_t = a_{20} + a_{21} z_{1t} + a_{22} z_{2t} + a_{23} x_{1t} + a_{24} x_{2,t-1} + a_{25} r_t + a_{26} s_{t-1} + v_{2t}.
\]

The first step in the estimation process is the computation of the market fundamentals. We report the definition of the short run excess supply at the reference price which is given by

\[
z_{1t} = q_t - \tilde{\beta}_1 p_t (p_t - \bar{p}_t) - c_t - \tilde{\varphi}_p (p_t - \bar{p}_t) + (s_{t-1} - \bar{s}_t).
\]
Table 2: Unit root tests

<table>
<thead>
<tr>
<th></th>
<th>q</th>
<th>c</th>
<th>s</th>
<th>x₁</th>
<th>x₂</th>
<th>r</th>
<th>p</th>
<th>p₁ᵇ</th>
<th>pᵱ</th>
<th>pᵲ</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADF p-value</td>
<td>-3.131ᵇ</td>
<td>-2.304ᵇ</td>
<td>-1.593</td>
<td>-3.834ᵇ</td>
<td>-1.413ᵃ</td>
<td>-2.594</td>
<td>-1.82</td>
<td>-1.494</td>
<td>-1.913</td>
<td>-2.010</td>
</tr>
<tr>
<td>KPSS p-value</td>
<td>0.178ᵇ</td>
<td>0.198ᵇ</td>
<td>1.11</td>
<td>0.211ᵇ</td>
<td>1.03</td>
<td>0.495</td>
<td>0.705</td>
<td>0.888</td>
<td>0.957</td>
<td>0.712</td>
</tr>
<tr>
<td>PP p-value</td>
<td>-3.848ᵇ</td>
<td>-2.332ᵇ</td>
<td>-1.497</td>
<td>-2.635ᵇ</td>
<td>-1.389</td>
<td>-1.462</td>
<td>-1.573</td>
<td>-1.316</td>
<td>-1.653</td>
<td>-1.919</td>
</tr>
</tbody>
</table>

Notes: Sample 1970-2010. (a) a drift term is included; (b) a linear trend is included; (c) only intercept included; default is both intercept and trend.
One first empirical question regards the choice of the reference price \( \bar{p}_t \) at which the market imbalance is computed. As discussed in the previous chapter, a moving average of appropriate length would seem appropriate. We experimented with different window sizes and eventually decided for a five years moving average, so that the relevant price is compared with the five preceding years when supply response is evaluated. The same rationale applies to the calculation of the reference stocks level \( \bar{s}_t \).

In order to proceed with the computation, we also need consistent estimates of the price coefficients in the structural supply and demand equations. As concerns consumption (21), we postulated that world grindings respond instantaneously to current price and income, and we need therefore to instrument price because of the endogeneity with the quantity demanded. The additional instruments we selected are the lagged weighted GDP of major producing countries and the lagged fCFA/US$ exchange rate, all in logarithms. The overall fit of the two-step efficient GMM estimation of the consumption equation is good \( (R^2 = 0.96) \), the coefficients are significant at 5% level, display the expected signs and economically meaningful magnitudes. The price elasticity \( (\varphi_p = -0.21) \) is slightly higher than in previous works, while the income elasticity is not far from unity \( (\varphi_x = 0.84) \). Hansen J-statistic \( \chi^2(1) = 0.269 \) (p-val 0.6042) confirms the validity of the instruments as the orthogonality conditions are met.

Turning to the supply equation (22), in our short-run model a single (one period lagged) price coefficient \( \hat{\beta}_1p \) is estimated. The OLS estimates are quite satisfactory \( (R^2 = 0.96) \) as all coefficients show the expected signs, even though most are significant only at 10% level. The short-run price elasticity is low \( (\hat{\beta}_p = 0.06) \) and the effect of lagged production \( (\hat{\beta}_q = 0.82) \) is positive and relevant, both expected results for a perennial crop, while the positive sign of the exchange rate coefficient \( (\hat{\beta}_x = 0.10) \) is coherent with the effect on production of a depreciation of the local currency. Finally, the step-type time dummy variable \( D_{86} \) included in the regression to account for the structural break identified in the previous section is positive and significant \( (\hat{\mu}_q = 0.30) \).

If we were to estimate a long-run version, in order to decide on the lag length, we know from the age yield profile of the cocoa tree that typically four to five years pass before the plant starts bearing fruits, then yields rapidly increase to reach a peak at about eight or eleven years for hybrid and traditional varieties, respectively.

The long-run market fundamental has been defined as the difference between production at the reference price and the consumption trend. We use as a proxy for consumption trend the fitted values from the estimated consumption equation \( \hat{c}_t \)

\[ z_{2t} = q_t - \hat{\beta}_1pL(p_t - \bar{p}_t) - \hat{c}_t. \]  

\[ (64) \]

\( \text{21} \) Notice how the use of a time varying reference value requires the addition of a time subscript in the notation.

\( \text{22} \) For the list of countries see Appendix D.

\( \text{23} \) Other trend extraction techniques are available, from polynomial and ARMA models to other filters used in macroeconometrics (Hodrick and Prescott, 1997; Baxter and King, 1999).
Figure 7 shows the pattern of the short and long term market fundamentals over the sample period 1970-2010.

Furthermore, from the supply and demand equations we get all the necessary estimates to compute the intermediate parameters $\phi_{ij}$ and $\theta_{ij}$ (the elements in the first two rows of the matrices $\Phi_0$ and $\Theta$ of the previous chapter) appearing in the evolution of the market fundamentals, $z_{1,t+1}$ and $z_{2,t+1}$, respectively, and in the REH restrictions.\footnote{An alternative procedure is to re-estimate the supply and demand equations jointly with the solved price and stock equations, thus getting new estimates of the parameters to be used in testing the restrictions, with likely efficiency gains (Gilbert, 1995).}

7.1 The reduced form estimates

7.1.1 GMM estimation

Once computed the market fundamentals, we can now turn to the estimation of the solved price and stock equations. Since $z_{1t}$ and $z_{2t}$ are, by construction, endogenous regressors, we have to use instrumental variable methods. This fact is, as expected, confirmed by Wu-Hausman endogeneity tests applied to both equations.

As it has been shown before (see Figure 5), a contemporaneous (negative) relation exists between price and stocks. Therefore, it seems appropriate to estimate the solved form as a system as the estimates would benefit in terms of efficiency. As to the estimation method, Perali and Pieroni (2004) used three
Stage least squares (3SLS).\textsuperscript{25} Peculiar features of 3SLS estimation are that the full set of instruments is used for all equations and the error terms are assumed to be \textit{iid}. Therefore we decided to pursue GMM estimation as it allows the use of equation specific instruments and produces efficient estimates that are robust to heteroskedasticity (and possibly autocorrelation if HAC correction is employed). Moreover, if the number of moment conditions available are in excess with respect to the number of parameters to be estimated, tests of overidentifying restrictions allow to check the validity of instruments and in general the specification of the model.

One major drawback from using instruments which are not relevant is that GMM estimates can be inconsistent. In order to check on instruments relevance, after controlling pairwise correlations between excluded instruments and endogenous regressors, we have run auxiliary first-stage regressions of market fundamentals on the full set (included and excluded) instruments. Low values of the partial $R^2$ and of the F-test on the excluded instruments may all signal possible problems of weak identification. Moreover, non-negligible discrepancies between the partial $R^2$ and the Shea partial $R^2$ may suggest possible collinearity between the instruments. In all these cases parsimony in the use of instruments is a good strategy.

We then performed single equation GMM estimates running C (or difference in Sargan) tests to decide on the best set of instruments to use for each equation. At this stage, we also checked the Kleibergen-Paap F statistic to further control for possible weak identification issues; the not particularly high values suggest that there is probably scope for further improvements in this direction. After these checks, the following sets of additional instruments have been selected: for the price equation, the one-period lagged values of stocks level $s_{t-1}$, palm oil price $p_{oji}^{o}$ and a production weighted food production index for major cocoa producers $FPI_{t-1}$; for the stock equation, the one-period lagged values of stock and production levels, coffee price $p_{cf}^{o}$ and cocoa price $p_{t-1}$, and the production weighted population in major producing countries $POP_{t-1}$. As we shall see, the tests conducted on the system GMM estimates confirm the validity of such instruments.

Table 3 reports the results of the iterated GMM estimates of the solved form, for the base model and one including a time dummy for the structural break in 1986, using the sample 1970-2010.

The overall fit of the GMM system estimation of the price and stock equation in both base and break models is quite good. The Hansen-J statistics, distributed as a $\chi^2(4)$ (four being the number of overidentifying restrictions), confirm the validity of the instruments (i.e. uncorrelated with the error terms and correctly excluded instruments). As a general remark, we observe that the autoregressive terms tend to absorb much of the explanatory power in the regression, as they are large and highly significant (this is true in particular for the lagged price) and display positive signs, thus showing a positive short-run

\textsuperscript{25}Gilbert (1995) first ran single equation regressions and then employed an iterated NL3SLS algorithm for the full system imposing different set of restrictions.
### Table 3: GMM estimates of reduced form

<table>
<thead>
<tr>
<th></th>
<th>price base</th>
<th>price break</th>
<th>stocks base</th>
<th>stocks break</th>
</tr>
</thead>
<tbody>
<tr>
<td>cons</td>
<td>-0.529</td>
<td>-8.723</td>
<td>-15.599</td>
<td>-3.522</td>
</tr>
<tr>
<td></td>
<td>(-0.10)</td>
<td>(-1.74)</td>
<td>(-1.45)</td>
<td>(-0.54)</td>
</tr>
<tr>
<td>z_{1t}</td>
<td>-0.414</td>
<td>-0.124</td>
<td>1.164</td>
<td>0.679</td>
</tr>
<tr>
<td></td>
<td>(-1.74)</td>
<td>(-0.60)</td>
<td>(2.23)</td>
<td>(2.11)</td>
</tr>
<tr>
<td>z_{2t}</td>
<td>-0.855</td>
<td>-1.730</td>
<td>-1.974</td>
<td>-0.544</td>
</tr>
<tr>
<td></td>
<td>(-0.87)</td>
<td>(-1.91)</td>
<td>(-0.95)</td>
<td>(-0.40)</td>
</tr>
<tr>
<td>x_{1t}</td>
<td>0.129</td>
<td>0.458</td>
<td>0.905</td>
<td>0.325</td>
</tr>
<tr>
<td></td>
<td>(0.56)</td>
<td>(1.97)</td>
<td>(1.55)</td>
<td>(0.95)</td>
</tr>
<tr>
<td>x_{2,t-1}</td>
<td>-0.242</td>
<td>-0.140</td>
<td>0.187</td>
<td>0.081</td>
</tr>
<tr>
<td></td>
<td>(-1.55)</td>
<td>(-1.00)</td>
<td>(0.96)</td>
<td>(0.64)</td>
</tr>
<tr>
<td>r_{t}</td>
<td>-0.055</td>
<td>-0.157</td>
<td>-0.285</td>
<td>-0.130</td>
</tr>
<tr>
<td></td>
<td>(-0.52)</td>
<td>(-1.38)</td>
<td>(-1.7)</td>
<td>(-1.31)</td>
</tr>
<tr>
<td>p_{t-1}</td>
<td>0.891</td>
<td>0.941</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(7.72)</td>
<td>(7.58)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>s_{t-1}</td>
<td></td>
<td></td>
<td>0.157</td>
<td>0.357</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.36)</td>
<td>(1.44)</td>
</tr>
<tr>
<td>D_{86}</td>
<td>-0.317</td>
<td>0.356</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-2.51)</td>
<td>(3.10)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>\eta_e</td>
<td>-2.662</td>
<td>-8.790</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-1.76)</td>
<td>(-0.61)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>\eta_s</td>
<td>-0.274</td>
<td>-0.518</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-0.89)</td>
<td>(-0.41)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Hansen’s J \(\chi^2(4)\): 1.180, 2.577

p-value: (0.7578), (0.4616)

persistence in the series. Using standard errors robust to heteroskedasticity, few other coefficients are actually significant at 5% significance level. The signs are the same in both models, while the inclusion of the time dummy improves the significance of coefficients in the price equation (the reverse is true for the stock equation).

In the price equation, short and long-term fundamentals display the expected negative sign,26 as excess supply tends to depress current prices, an high $z_{2t}$ reducing in particular the incentive to carry stocks forward, even though it is not significant in the base model (at 10% in the break one). The absolute value of $z_{2t}$ is rather high in the base model, approximately twice as large as $z_{1t}$ (significant at 10%). The positive sign of $z_{1t}$ (significant at 5%) in the stock equation correctly reflects the immediate impact of excess supply on stocks level, while the negative sign of $z_{2t}$ is in our opinion consistent with the reduced incentive to accumulating stocks in face of a weak long-term fundamental.

Income elasticity is correctly signed in both equations, but only significant in the break model. The negative sign of interest rate in the stock equation is coherent with a speculative stockholding explanation as postulated in our model, as an increased cost of storage would reduce the incentive to carry over stocks, in turn depressing current price. Conversely, an interpretation which exploits the countercyclical nature of commodity prices could be offered to a possible positive sign of the interest rate in the price equation. In fact, interest rates tend to be high in late expansion and early recession phases, exactly when commodity prices are typically high. The opposite signs of the exchange rate in the two equations look consistent with an increased supply by major producers following a competitive devaluation, in turn possibly depressing world prices and increasing stock levels. Though, both interest rate and exchange rate variables are not significant in both models. Finally, the time dummy included to account for the structural break in 1986 is significant in both equations and display the expected signs, increasing the mean stock level and depressing the mean price.

As concerns the free structural parameters $\eta_e$ and $\eta_s$, they can be recovered from the estimated solved form by manipulating the response matrix $A_0$, as previously discussed. The estimates of $\eta_e$ from both models are not satisfactory, as display negative signs and not much plausible large values, even though in general we would expect the speculative component of stock demand to be very responsive to even small expected capital gains. The sign of $\eta_s$ is negative as well and not significant. Such results deserve further investigation.

7.1.2 VAR models

This section presents the estimation of different VAR models in the attempt to more fully account for the dynamics of the system. As discussed so far, the role of stocks is crucial in explaining the, at least, short term dynamics of world prices. Therefore, we shall continue to analyze the relationship between

---

26 A useful comparison can be drawn with the signs derived from the theoretical model in Appendix C through a comparative statics analysis.
price and stocks in a substantially bivariate framework, keeping the specification close to the derived solved form. In our exercise we shall move from an atheoretical specification, where only lagged values of price and stocks are used as explanatory variables, to a more structured one where additional exogenous regressors from the solved form are included. Other alternative specifications were of course available, increasing for instance the VAR dimension or modelling explicitly long-run relations in VECM form. Nevertheless, that would have implied specifying ultimately a different model, possibly in a non rational expectations framework (Deaton and Laroque, 2003) and we decided to leave it for future work.

A first choice to make concerns whether to estimate a model in levels or first-differences. Results from the preliminary analysis on the stationarity of the series do not provide definitive answers. Even though the unit root hypothesis cannot be rejected for both series over the whole sample, if we account for the structural break in 1986 and restrict accordingly the sample period, the price and stocks variables can be considered $I(0)$ processes. In the end, we decided to keep the variables in (log) levels, using the sample 1970-2010 and comparing benchmark estimates from various models with those including a time dummy to control for the shift in the mean.\footnote{However, the point estimates from an integrated VAR are consistent, as long as the dynamics is correctly specified, and can be used for forecasting purposes (Sims et al., 1990). Moreover, the lag selection criteria are still valid.}

In order to decide on the VAR length, lag selection criteria (FPE, AIC, SBIC, HQIC) have been used. As can be seen from Table 4, they unambiguously seem to suggest that a single lag is sufficient to control for the persistency of the processes. Hence, a VAR(1) provides our benchmark model, even though from the univariate analysis on the price series and a priori knowledge on the age-yield profile of the cocoa tree we know that other lags might be relevant. The general formulation of the VAR can be written as

$$ y_t = \mu + B(L)y_t + B_0x_t + v_t $$  \hspace{-2em}

where $y_t$ is a $K \times 1$ vector of endogenous variables, $B(L)$ is a matrix polynomial of order $p$, $x_t$ is an $L \times 1$ vector of exogenous regressors, $B_0$ a $K \times L$ matrix of coefficients, $\mu$ a $K \times 1$ vector of constants and $v_t$ a $K \times 1$ vector of disturbances. We assume that a sufficient number of lags have been included so that the $v_t$ can be considered as white noises. In our basic specification $y_t = (p_t, s_t)'$ is a $2 \times 1$ vector while $B_0$ is restricted to be 0. As to the lag polynomial, after checking the lag selection criteria and exploiting a priori information, we can restrict it to $B_1L$ in a short-run specification (model A1).

In the extended specifications we add exogenous variables from the solved form, whereas the fitted values of the market fundamentals, $\tilde{z}_{1t}$ and $\tilde{z}_{2t}$, substitute the original variables because of the known endogeneity problems. The full vector of the exogenous variables thus becomes $x_t = (\tilde{z}_{1t}, \tilde{z}_{2t}, x_{1t}, x_{2,t-1}, r_t, D_{86})'$, as the time dummy has been included. We estimate a model including the entire set of exogenous variables (model C1) and a restricted version which excludes
the fundamentals (model B1), in both cases with or without the time dummy.

Table 5 reports the estimates of the VAR(1) models presented above. The overall fit of the models, as measured by the log-likelihood, in general improves as we move from simple to more structured specifications and a further improvement is achieved by including the time dummy, whose effect is negative in the price equation, as it captures the declining trend in prices, and positive in the stocks equation reflecting the increased stocks level.

Turning to the single coefficients, the autoregressive terms in model A1(d) are positive and significant in both equations, and the magnitude and significance is largely preserved in models B1 and C1. Conversely, cross lagged endogenous variables, $s_{t-1}$ negatively impacting on current price and $p_{t-1}$ positively affecting stocks, lose relevance and significance in models with other exogenous regressors. In particular, the inclusion of $z_{1t}$ in model C1, which is negative and significant, seems to make the effect of $s_{t-1}$ negligible, not surprisingly as by construction it includes lagged stocks. The same considerations made above concerning the expected signs of the coefficients should hold here. In the stock equation, the signs of $x_{1t}$, $x_{2,t-1}$ and $r_t$ remain the same in models B1 and C1, while the same is not true in the price equation, where the coefficients are however poorly estimated.

We then compare the GMM estimates of the reduced form with those from a restricted VAR(C1r) model including the same exogenous variables, where the cross lagged endogenous variables have been dropped, as reported in Table 6. As a general remark, we observe again that also in the VAR estimates the autoregressive components, positive and rather close to unity (also for stocks in this case), seem to absorb much of the explanatory power as the other exogenous regressors, apart from $z_{1t}$, $z_{2t}$ (in the price equation) and marginally $x_{1t}$, result not significant in both equations. The signs of the GMM and VAR estimates to a very large extent coincide in both models, except for the positive sign of $z_{2t}$ in the VAR stock equation. Also the magnitudes of coefficients are largely comparable, in particular in the price equation, while discrepancies are larger in the stock equation, notably for short and long-run fundamentals. The same considerations made before concerning the time dummy and the derived parameters $\eta_c$ and $\eta_s$ hold here. The analysis of residuals (not reported) from the preceding estimated models suggest that disturbances can be considered as innovations. In particular, the Breush-Godfrey tests exclude the presence of residual autocorrelation, while the Jarque-Bera tests do not reject the hypothesis of normality of residuals. Furthermore, the VAR models satisfy the stability conditions as all eigenvalues of the companion matrix lie inside the unit circle.

7.2 Testing the REH restrictions

A peculiar feature of rational expectations models is that they yield a set of non-linear cross equations restrictions linking structural and reduced form parameters. Tests of such restrictions provide a means of indirectly verifying the rational expectations hypothesis conditional on the validity of the model or
Table 4: Lag order selection criteria

<table>
<thead>
<tr>
<th>lag</th>
<th>LL</th>
<th>LR</th>
<th>FPE</th>
<th>AIC</th>
<th>HQIC</th>
<th>SBIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>-45.7986</td>
<td>0.03529</td>
<td>2.33164</td>
<td>2.36208</td>
<td>2.41523</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>40.67</td>
<td>172.94*</td>
<td>0.00063*</td>
<td>-1.69122*</td>
<td>-1.59991*</td>
<td>-1.44045*</td>
</tr>
<tr>
<td>2</td>
<td>43.6973</td>
<td>6.05460</td>
<td>0.00066</td>
<td>-1.64377</td>
<td>-1.49158</td>
<td>-1.22583</td>
</tr>
<tr>
<td>3</td>
<td>45.5666</td>
<td>3.73860</td>
<td>0.00074</td>
<td>-1.53983</td>
<td>-1.32677</td>
<td>-0.95471</td>
</tr>
<tr>
<td>4</td>
<td>47.9156</td>
<td>4.69810</td>
<td>0.00081</td>
<td>-1.45930</td>
<td>-1.18535</td>
<td>-0.70700</td>
</tr>
</tbody>
</table>

Notes: * denotes lag chosen by the criterion. LL (Loglikelihood), LR (Likelihood Ratio), FPE (Final Prediction Error), AIC (Akaike), HQIC (Hannan-Queen), SBIC (Schwartz).

equivalently, and perhaps preferably, of testing the validity of model specification under the REH. In the previous chapter we have illustrated the analytical derivation of the restrictions which we now test using the estimates of the intermediate and reduced form parameters from both base and break models.

Since we are using estimates from a short-run model, and in particular short-run supply and demand elasticities have been employed in constructing the market fundamentals, what we are testing is in fact whether stockholders form their expectations rationally, or consistently with the postulated model, but using a short-term information set, neglecting events far in the past affecting supply conditions and potentially generating future market imbalances. Although the restrictions stem from and are theoretically consistent with the solved form GMM estimates, the specification used in the restricted VAR(C1r) model matches the derived reduced form, allowing us to run the tests using the VAR estimates of the response matrix $A_0$, as from Table 6.

Table 7 reports the Wald statistics of the logical restrictions from the four estimated models over the period 1970-2010. Considering the GMM estimates, the single restrictions easily pass the tests in all cases and the overall set is accepted in both models, even though more weakly in the break model. The single restrictions are not rejected also using the VAR(C1r) estimates, except for $z_{2t}$ and the overall set in the break model. This result does not come as a surprise since the GMM and VAR estimates are quite close.

Similar considerations can be made by looking at the results from the tests of the REH restrictions, which provide a check of the consistency of the estimated price equation with the price expectation in the stock demand equation (23).

Table 8 presents the Wald statistics of the four REH restrictions given above. Using either the GMM or the VAR(C1r) estimates, the single restrictions and the overall sets are not rejected for both models, the break one performing slightly better. Such results are encouraging as suggest a substantially correct basic specification of the model and are supportive of the view that the relevant information set of stockholders is quite limited in time.

37
<table>
<thead>
<tr>
<th></th>
<th>A1</th>
<th>B1</th>
<th>C1</th>
<th>A1d</th>
<th>B1d</th>
<th>C1d</th>
</tr>
</thead>
<tbody>
<tr>
<td>cons</td>
<td>3.9079</td>
<td>-1.762</td>
<td>3.655</td>
<td>3.377</td>
<td>-7.616</td>
<td>-6.92</td>
</tr>
<tr>
<td></td>
<td>(2.22)</td>
<td>(-0.26)</td>
<td>(0.49)</td>
<td>(1.84)</td>
<td>(-1.04)</td>
<td>(-0.76)</td>
</tr>
<tr>
<td>$p_{t-1}$</td>
<td>0.7014</td>
<td>0.6761</td>
<td>0.815</td>
<td>0.691</td>
<td>0.634</td>
<td>0.834</td>
</tr>
<tr>
<td></td>
<td>(5.59)</td>
<td>(3.9)</td>
<td>(4.1)</td>
<td>(5.49)</td>
<td>(3.76)</td>
<td>(4.35)</td>
</tr>
<tr>
<td>$s_{t-1}$</td>
<td>-0.226</td>
<td>-0.345</td>
<td>0.1125</td>
<td>-0.12</td>
<td>-0.25</td>
<td>0.091</td>
</tr>
<tr>
<td></td>
<td>(-1.9)</td>
<td>(-1.64)</td>
<td>(0.36)</td>
<td>(-0.78)</td>
<td>(-1.19)</td>
<td>(0.3)</td>
</tr>
<tr>
<td>$x_{1t}$</td>
<td>-0.171</td>
<td>-0.313</td>
<td>-0.185</td>
<td>-0.28</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-1.02)</td>
<td>(-1.76)</td>
<td>(-1.14)</td>
<td>(-1.64)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$x_{2,t-1}$</td>
<td>0.3416</td>
<td>-0.048</td>
<td>0.596</td>
<td>0.424</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.87)</td>
<td>(-0.11)</td>
<td>(1.48)</td>
<td>(0.87)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$r_t$</td>
<td>-0.073</td>
<td>0.0242</td>
<td>-0.073</td>
<td>-0.07</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-0.78)</td>
<td>(0.2)</td>
<td>(-0.8)</td>
<td>(-0.56)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\hat{z}_{1t}$</td>
<td>-0.667</td>
<td>-0.48</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-1.99)</td>
<td>(-1.41)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\hat{z}_{2t}$</td>
<td>0.0194</td>
<td>-1.38</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td>(-1.04)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$D_{86}$</td>
<td>-0.16</td>
<td>-0.325</td>
<td>-0.39</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-1.01)</td>
<td>(-1.87)</td>
<td>(-1.89)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>A1</th>
<th>B1</th>
<th>C1</th>
<th>A1d</th>
<th>B1d</th>
<th>C1d</th>
</tr>
</thead>
<tbody>
<tr>
<td>cons</td>
<td>-2.29</td>
<td>-3.077</td>
<td>-7.598</td>
<td>-1.34</td>
<td>2.814</td>
<td>2.15</td>
</tr>
<tr>
<td></td>
<td>(-1.92)</td>
<td>(-0.64)</td>
<td>(-1.44)</td>
<td>(-1.18)</td>
<td>(0.58)</td>
<td>(0.35)</td>
</tr>
<tr>
<td>$p_{t-1}$</td>
<td>0.2007</td>
<td>0.2293</td>
<td>0.1743</td>
<td>0.219</td>
<td>0.272</td>
<td>0.157</td>
</tr>
<tr>
<td></td>
<td>(2.36)</td>
<td>(1.88)</td>
<td>(1.25)</td>
<td>(2.8)</td>
<td>(2.42)</td>
<td>(1.21)</td>
</tr>
<tr>
<td>$s_{t-1}$</td>
<td>1.1073</td>
<td>1.0991</td>
<td>0.7791</td>
<td>0.922</td>
<td>1.003</td>
<td>0.799</td>
</tr>
<tr>
<td></td>
<td>(13.71)</td>
<td>(7.41)</td>
<td>(3.59)</td>
<td>(9.48)</td>
<td>(7.19)</td>
<td>(3.99)</td>
</tr>
<tr>
<td>$x_{1t}$</td>
<td>0.0179</td>
<td>0.1194</td>
<td>0.032</td>
<td>0.091</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.15)</td>
<td>(0.96)</td>
<td>(0.3)</td>
<td>(0.79)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$x_{2,t-1}$</td>
<td>0.0244</td>
<td>0.3162</td>
<td>-0.231</td>
<td>-0.12</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.09)</td>
<td>(1.03)</td>
<td>(-0.86)</td>
<td>(-0.36)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$r_t$</td>
<td>-0.031</td>
<td>-0.124</td>
<td>-0.031</td>
<td>-0.04</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-0.46)</td>
<td>(-1.46)</td>
<td>(-0.51)</td>
<td>(-0.42)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\hat{z}_{1t}$</td>
<td>0.4632</td>
<td>0.287</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.97)</td>
<td>(1.26)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\hat{z}_{2t}$</td>
<td>-0.543</td>
<td>0.747</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-0.67)</td>
<td>(0.83)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$D_{86}$</td>
<td>0.291</td>
<td>0.328</td>
<td>0.36</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.92)</td>
<td>(2.84)</td>
<td>(2.59)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Log likelihood: 40.670, 49.089, 52.340, 46.538, 53.695, 56.272  

### Table 6: GMM vs VAR(C1r) estimates of reduced form

<table>
<thead>
<tr>
<th>GMM</th>
<th>VAR(C1r)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>price equation</strong></td>
<td><strong>base</strong></td>
</tr>
<tr>
<td>cons</td>
<td>-0.529</td>
</tr>
<tr>
<td></td>
<td>(-0.1)</td>
</tr>
<tr>
<td>( z_{1t} )</td>
<td>-0.414</td>
</tr>
<tr>
<td></td>
<td>(-1.74)</td>
</tr>
<tr>
<td>( z_{2t} )</td>
<td>-0.855</td>
</tr>
<tr>
<td></td>
<td>(-0.87)</td>
</tr>
<tr>
<td>( x_{1t} )</td>
<td>0.129</td>
</tr>
<tr>
<td></td>
<td>(0.56)</td>
</tr>
<tr>
<td>( x_{2,t-1} )</td>
<td>-0.242</td>
</tr>
<tr>
<td></td>
<td>(-1.55)</td>
</tr>
<tr>
<td>( r_t )</td>
<td>-0.055</td>
</tr>
<tr>
<td></td>
<td>(-0.52)</td>
</tr>
<tr>
<td>( p_{t-1} )</td>
<td>0.891</td>
</tr>
<tr>
<td></td>
<td>(7.72)</td>
</tr>
<tr>
<td>( s_{t-1} )</td>
<td>-0.317</td>
</tr>
<tr>
<td></td>
<td>(-2.51)</td>
</tr>
<tr>
<td>( D_{86} )</td>
<td>0.356</td>
</tr>
<tr>
<td></td>
<td>(3.1)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>stock equation</strong></th>
<th><strong>base</strong></th>
<th><strong>break</strong></th>
<th><strong>base</strong></th>
<th><strong>break</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>cons</td>
<td>-15.599</td>
<td>-3.522</td>
<td>-5.997</td>
<td>3.760</td>
</tr>
<tr>
<td></td>
<td>(-1.45)</td>
<td>(-0.54)</td>
<td>(-1.35)</td>
<td>(0.68)</td>
</tr>
<tr>
<td>( z_{1t} )</td>
<td>1.164</td>
<td>0.679</td>
<td>0.386</td>
<td>0.225</td>
</tr>
<tr>
<td></td>
<td>(2.23)</td>
<td>(2.11)</td>
<td>(2.03)</td>
<td>(1.19)</td>
</tr>
<tr>
<td>( z_{2t} )</td>
<td>-1.974</td>
<td>-0.544</td>
<td>0.256</td>
<td>1.488</td>
</tr>
<tr>
<td></td>
<td>(-0.95)</td>
<td>(-0.4)</td>
<td>(0.38)</td>
<td>(1.89)</td>
</tr>
<tr>
<td>( x_{1t} )</td>
<td>0.905</td>
<td>0.325</td>
<td>0.332</td>
<td>-0.109</td>
</tr>
<tr>
<td></td>
<td>(1.55)</td>
<td>(0.95)</td>
<td>(1.49)</td>
<td>(-0.41)</td>
</tr>
<tr>
<td>( x_{2,t-1} )</td>
<td>0.187</td>
<td>0.081</td>
<td>0.038</td>
<td>0.021</td>
</tr>
<tr>
<td></td>
<td>(0.96)</td>
<td>(0.64)</td>
<td>(0.32)</td>
<td>(0.2)</td>
</tr>
<tr>
<td>( r_t )</td>
<td>-0.285</td>
<td>-0.130</td>
<td>-0.040</td>
<td>0.040</td>
</tr>
<tr>
<td></td>
<td>(-1.7)</td>
<td>(-1.31)</td>
<td>(-0.56)</td>
<td>(0.54)</td>
</tr>
<tr>
<td>( p_{t-1} )</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( s_{t-1} )</td>
<td>0.157</td>
<td>0.357</td>
<td>0.743</td>
<td>0.752</td>
</tr>
<tr>
<td></td>
<td>(0.36)</td>
<td>(1.44)</td>
<td>(6.61)</td>
<td>(6.95)</td>
</tr>
<tr>
<td>( D_{86} )</td>
<td>0.356</td>
<td>( D_{86} )</td>
<td>0.380</td>
<td>( D_{86} )</td>
</tr>
<tr>
<td></td>
<td>(3.1)</td>
<td>(2.7)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Hansen’s \( J \chi^2(4) \): 1.180 \( \chi^2(4) \): 2.397  

Table 7: Wald tests of logical restrictions

<table>
<thead>
<tr>
<th>j</th>
<th>variable</th>
<th>GMM base</th>
<th>VAR(C1r) base</th>
<th>GMM break</th>
<th>VAR(C1r) break</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>z_{1t}</td>
<td>0.28</td>
<td>0.15</td>
<td>0.47</td>
<td>0.13</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.5992)</td>
<td>(0.6986)</td>
<td>(0.492)</td>
<td>(0.7201)</td>
</tr>
<tr>
<td>2</td>
<td>z_{2t}</td>
<td>0.33</td>
<td>0.20</td>
<td>0.16</td>
<td>0.20</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.565)</td>
<td>(0.6556)</td>
<td>(0.6892)</td>
<td>(0.406)</td>
</tr>
<tr>
<td>3</td>
<td>x_{1t}</td>
<td>0.20</td>
<td>0.23</td>
<td>0.74</td>
<td>0.70</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.6556)</td>
<td>(0.6318)</td>
<td>(0.3904)</td>
<td>(0.4031)</td>
</tr>
<tr>
<td>4</td>
<td>x_{2,t-1}</td>
<td>1.68</td>
<td>0.53</td>
<td>0.00</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.1955)</td>
<td>(0.4731)</td>
<td>(0.9567)</td>
<td>(0.4031)</td>
</tr>
<tr>
<td>5</td>
<td>r_{t}</td>
<td>0.23</td>
<td>0.51</td>
<td>0.70</td>
<td>0.17</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.6518)</td>
<td>(0.4731)</td>
<td>(0.6792)</td>
<td>(0.7066)</td>
</tr>
<tr>
<td>Overall (\chi^2) (5)</td>
<td>2.23</td>
<td>9.36</td>
<td>8.42</td>
<td>17.89</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.8162)</td>
<td>(0.0955)</td>
<td>(0.1346)</td>
<td>(0.0031)</td>
</tr>
</tbody>
</table>

Notes: t-statistics reported are \(\chi^2(1)\). P-values are given in parentheses.

Table 8: Wald tests of REH restrictions

<table>
<thead>
<tr>
<th>equation</th>
<th>GMM base</th>
<th>GMM break</th>
<th>VAR(C1r) base</th>
<th>VAR(C1r) break</th>
</tr>
</thead>
<tbody>
<tr>
<td>16</td>
<td>0.28</td>
<td>0.11</td>
<td>0.15</td>
<td>0.13</td>
</tr>
<tr>
<td></td>
<td>(0.5984)</td>
<td>(0.7407)</td>
<td>(0.6986)</td>
<td>(0.7201)</td>
</tr>
<tr>
<td>17</td>
<td>0.49</td>
<td>0.16</td>
<td>0.20</td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td>(0.4851)</td>
<td>(0.6906)</td>
<td>(0.6523)</td>
<td>(0.6984)</td>
</tr>
<tr>
<td>18</td>
<td>0.51</td>
<td>0.14</td>
<td>0.19</td>
<td>0.11</td>
</tr>
<tr>
<td></td>
<td>(0.4731)</td>
<td>(0.7066)</td>
<td>(0.6618)</td>
<td>(0.7376)</td>
</tr>
<tr>
<td>19</td>
<td>0.53</td>
<td>0.17</td>
<td>0.2</td>
<td>0.14</td>
</tr>
<tr>
<td></td>
<td>(0.4676)</td>
<td>(0.6792)</td>
<td>(0.6552)</td>
<td>(0.7047)</td>
</tr>
<tr>
<td>Overall (\chi^2) (4)</td>
<td>1.12</td>
<td>1.37</td>
<td>0.68</td>
<td>0.46</td>
</tr>
<tr>
<td></td>
<td>(0.8908)</td>
<td>(0.8499)</td>
<td>(0.9532)</td>
<td>(0.9777)</td>
</tr>
</tbody>
</table>

Notes: t-statistics reported are \(\chi^2(1)\). P-values in brackets.
8 Concluding remarks

In this work we specified a model for the global market of a perennial crop, namely cocoa, under linear rational expectations. One feature common to primary commodity markets is the lagged supply response by producers to market signals. In fact, domestic policies, imperfect price transmission along the supply chain and substantial adjustment costs under uncertain market conditions make supply quite unresponsive in the short run to world market prices. This is most true for perennial crops where the biological gestation lag further accentuates such characteristics.

As a result, given the low demand elasticity, stocks fluctuate widely in response to supply shocks and a fairly stable relationship between stock levels and world price can be observed in the market. Nevertheless, stocks may fail to absorb market imbalances and the use of commodities as a financial asset in a balanced portfolio may also determine temporary departures of price from its equilibrium time path, exacerbating price spikes and producing bubble effects with possibly destabilizing long-run effects on the market.

Such stylized facts led us on the one hand to pay attention to the derivation of the supply equation and on the other hand to focus on the role of stocks, by including a speculative stockholding equation in the structural model and deriving a dual solved or reduced form in price and stocks. The solution method follows Gilbert (1995) in reducing the state variables to a reduced set including, along with world income, interest rate and exchange rate, two constructed variables representing short and long run market imbalances, whose inclusion appears theoretically appealing for the reasons explained above.

The consideration, anticipated in Ghosh et al. (1987) and stressed in Gilbert (1995), that the implications of REH may be limited for long-term market developments because agents do not possess relevant information, motivated us to extend the definition of market fundamental accounting for past supply shocks, given that the forecasting horizon for perennial crops should be longer than for other commodities.

The rational expectations hypothesis generates a set of nonlinear cross equation restrictions linking structural and reduced form parameters, which provide indirect tests of the REH and in general of model specification. We used such restrictions to recover the structure of the model and carry out a comparative statics in order to assess the model predictions about the effects of changes in the state variables on price and stocks.

Furthermore, we investigated, in a deterministic setting, the stability conditions of a short-run version of the model, deriving the price rational expectations solution. We noticed how the model generates one explosive and one stable root, as usual in a rational expectations framework. Therefore, in order to obtain a bounded time path for the price solution, the explosive root must be ruled out and stationarity conditions imposed. We further simulated the impact of a positive supply shock on price and stocks under different scenarios given by alternative combinations of values for the structural parameters. Results suggest that low values of supply and stock demand elasticities, generating in turn
relatively higher values of the stable root, tend to increase the persistency of the series. A greater sensitivity of stock demand to expected capital gains seems to have a stabilizing effect on price, deflating the explosive root and making adjustments to long-term equilibria more gradual. Conversely, a high demand elasticity would tend to amplify market imbalances.

In the empirical section of the paper, we estimated the solved or reduced form in price and stocks of the short-run rational expectations model of the world cocoa market. As already discussed, the choice of focusing on these two variables is motivated by the apparent stable relationship over time existing between world price, as measured by the ICCO indicator price, and the stocks-to-use ratio, an indicator of cocoa availability at global level monitored by industry analysts.

Hence, a reduced form in price and stocks, using as state variables the market fundamentals and exogenous variables such as world income, exchange rate and interest rate, has been estimated using the generalized method of moments (GMM) to account for the endogeneity in the fundamentals. The GMM estimates, robust to possible heteroskedasticity, are quite satisfactory. The estimated coefficients present theoretically consistent signs, as derived in the qualitative analysis, and meaningful magnitudes, apparently confirming the hypothesized effect of market fundamentals on current price and stocks. The lagged endogenous variables are mostly significant, confirming the validity of their inclusion in the statistical model. A model including a time dummy to account for a structural break has also been estimated yielding similar results and a slightly superior fit. Conversely, the derived estimates of the stock elasticities are not satisfactory suggesting further investigation in this direction.

The restrictions stemming from the hypothesized rational expectations have been tested using estimates from both models. The single restrictions and the overall set are not rejected, suggesting an acceptable model specification and the validity of the model consistent expectations.

At this regard, we remark that the failure at rejecting the REH restrictions using estimates from short-run models seems to imply that the information set used by market participants in forming their expectations about future market developments is essentially limited in time. Our initial guess that past events, mostly supply related, might be incorporated in expectations would not seem to be supported by empirical evidence, at least as far as stockholding decisions are concerned. Nevertheless, based on preliminary results from long lag structure VAR models, testing using also estimates from a long-run version model could possibly corroborate these findings.

An alternative approach, aiming at possibly providing a better representation of the underlying statistical model, in particular the short-run dynamics, has also been pursued by specifying different vector autoregressive models in the price and stocks dimension. Nevertheless, specification tests pointed to a VAR(1) as a suitable representation of price and stocks dynamics which we augmented in alternative versions with additional exogenous variables from the solved form, supposed to have a contemporaneous impact on the endogenous variables and thus paralleling the GMM estimation of the reduced form.

The GMM estimates of the solved form and those from a restricted VAR
model, with a matching specification, have been then compared. The results are to a large extent similar, also as concerns the acceptance of the REH restrictions, confirming the validity of the short-run statistical model and the relevance of the autoregressive components which tend though to crowd out other state variables in the model.

Finally, we observe that the ongoing process of vertical integration, both upstream and downstream, and the increased interest in supply conditions by international traders, may justify a modelling exercise focused on market fundamentals or tentatively exploiting the information set used by global players, despite speculative bubbles on financial markets might turn away prices from market fundamentals for prolonged periods and even generate permanent effects.
Appendices

A Derivation of REH restrictions

This appendix illustrates the derivation of the logical and REH restrictions, linking structural and reduced form parameters, as defined in Proposition 5 and 6.

Proposition 5. Logical restrictions
Given the matrix $A_0$ of the reduced form slope coefficients and the parameter $\lambda$, which is a function of the price response coefficients in the structural supply and demand equations, the logical restrictions are given by

$$a_{21} = 1 + \frac{1}{\lambda} a_{11}, \quad a_{2j} = \frac{1}{\lambda} a_{1j}, \quad j = 2, 3, 4, 5.$$  \hspace{1cm} (A.1)

Proof 5.
Dividing the second row elements of $A_0$ by the corresponding elements of the first row it is easy to verify that

$$\frac{a_{22}}{a_{12}} = -\frac{\eta_e (a_{11} \theta_{12} + a_{12} \theta_{22})}{\lambda \eta_e (a_{11} \theta_{12} + a_{12} \theta_{22})} = \frac{1}{\lambda},$$
$$\frac{a_{23}}{a_{13}} = -\frac{\eta_e (a_{13} + a_{11} \theta_{13})}{\lambda \eta_e (a_{13} + a_{11} \theta_{13})} = \frac{1}{\lambda},$$
$$\frac{a_{24}}{a_{14}} = -\frac{\eta_e (a_{14} + a_{11} \theta_{14} + a_{12} \theta_{24})}{\lambda \eta_e (a_{14} + a_{11} \theta_{14} + a_{12} \theta_{24})} = \frac{1}{\lambda},$$
$$\frac{a_{25}}{a_{15}} = \frac{\eta_e (1 - a_{15} \rho_3)}{\lambda \eta_e (1 - a_{15} \rho_3)} = \frac{1}{\lambda}.$$

The first logical restriction is instead derived as follows

$$\frac{a_{21} - 1}{a_{11}} = \frac{\lambda \eta_e (a_{16} - 1)}{\lambda (1 - a_{11} \eta_e \phi_{12})} = \frac{1}{\lambda}. \quad \square$$

A second set of restrictions provide a check of the consistency of the estimated price equation (34) with the rational expectation of price in the stock demand equation (23).

Proposition 6. REH restrictions
Given the matrix $A_0$ of the reduced form slope coefficients, the parameter $\lambda$, the logical restrictions (A.1) and the matrices of intermediate parameters $\Phi_0$ and $\Theta$, the rational expectations restrictions are given by

\begin{align*}
\frac{a_{22}}{a_{21}} &= \frac{(a_{16} a_{12} + \theta_{22} a_{12} - a_{12} + a_{11} \theta_{12})}{a_{11} (a_{16} - 1)} \quad \text{(A.2)}
\end{align*}
Proof 6.
In what follows we illustrate the analytical derivation of the REH restriction (A.2). The other restrictions are analogously obtained. Substituting the rational expectation of price (41) into the stock demand equation (23), and equating the coefficient of \( z_{2t} \) with the corresponding coefficient in the reduced form stock equation (35), we get

\[
\frac{a_{23}}{a_{21}} = \frac{(a_{13}a_{16} + a_{11}\theta_{13})}{a_{11}(a_{16} - 1)} \tag{A.3}
\]

\[
\frac{a_{24}}{a_{21}} = \frac{(a_{14}a_{16} + a_{11}\theta_{14} + a_{12}\theta_{24})}{a_{11}(a_{16} - 1)} \tag{A.4}
\]

\[
\frac{a_{25}}{a_{21}} = \frac{(a_{16}a_{15} + \rho_{3}a_{15} - a_{15} - 1)}{a_{11}(a_{16} - 1)} \tag{A.5}
\]

An intermediate step of the proof thus consists in deriving an analytical expression for the free parameter \( \eta_{e} \). This is achieved by repeating the first step above for the \( z_{1t} \) coefficients and combining the resulting expression with the first logical restriction as given below

\[
a_{21} = \frac{\lambda (-a_{11}\eta_{e}\theta_{12} - a_{12}\eta_{e}\theta_{22})}{\eta_{e}(a_{16}\lambda - \lambda + a_{11}\phi_{12}) - 1}, \quad a_{21} = 1 + \frac{1}{\lambda} a_{11}.
\]

After cancelling out \( \lambda \) and solving for \( \eta_{e} \) we get

\[
\eta_{e} = \frac{a_{21}}{a_{11}(a_{16} + a_{21}\phi_{12} - 1)}. \tag{A.7}
\]

We can now substitute for \( \eta_{e} \) into (A.6), use the logical restriction \( a_{22} = \lambda^{-1} a_{12} \) for cancelling out \( \lambda \), and solve for \( a_{22} \) to get upon simplifying

\[
\frac{a_{22}}{a_{21}} = \frac{(a_{16}a_{12} + \theta_{22}a_{12} - a_{12} + a_{11}\theta_{12})}{a_{11}(a_{16} - 1)}.
\]

\[\square\]
B Recovery of structural parameters

The structure of the model can be recovered using the logical and REH restrictions which link the structural and reduced form parameters. From the definitions of the intermediate parameters in equations (29) and (30) we note that \( \theta_{12} = \theta_{22} = \beta_q \). Thus, solving (A.2) with respect to \( \beta_q \) we get

\[
\beta_q = -\frac{(a_{16} - 1)(a_{12}a_{21} - a_{11}a_{22})}{(a_{11} + a_{12})a_{21}}.
\]  

(B.1)

Similarly, noting that \( \theta_{13} = -\varphi_x \) and \( \theta_{24} = \beta_x \), we can solve (A.3) and (A.4) for \( \beta_x \) and \( \varphi_x \), respectively, to obtain

\[
\beta_x = -\frac{a_{13}a_{16}a_{21} - a_{11}a_{23} + a_{11}a_{16}a_{23}}{a_{11}a_{21}},
\]  

(B.2)

\[
\varphi_x = -\frac{a_{14}a_{16}a_{21} - a_{11}a_{24} + a_{11}a_{16}a_{24}}{(a_{11} + a_{12})a_{21}}.
\]  

(B.3)

In order to recover the price supply elasticity \( \beta_p \), since \( \varphi_{11} = \beta_q \beta_p \), we first derive the intermediate parameter \( \varphi_{11} \). To this end, we equate the reduced form coefficients of \( p_{t-1} \) in (34) and (40)

\[
a_{16} = \frac{\lambda(-a_{11}\eta_c \varphi_{11} - a_{12}\eta_c \varphi_{11})}{(a_{16}\lambda - \lambda + a_{11})\eta_c - 1}
\]

and substituting \( a_{12} = \lambda a_{22} \) and (A.7) into the above expression to eliminate \( \lambda \) and \( \eta_c \) we get

\[
\varphi_{11} = -\frac{a_{16}(-a_{12}a_{21} + a_{12}a_{16}a_{21} + a_{11}a_{22} - a_{11}a_{16}a_{22})}{(a_{11} + a_{12})a_{12}a_{21}}.
\]  

(B.4)

We can now substitute (B.4) and (B.1) into \( \varphi_{11} = \beta_q \beta_p \) to get, upon simplifying,

\[
\beta_p = \frac{a_{16}}{a_{12}}.
\]  

(B.5)

Finally, in order to derive the price demand elasticity \( \varphi_p \), we substitute for \( \lambda \) and \( \beta_p \) into \( \lambda^{-1} = \beta_p - \varphi_p \) to obtain

\[
\varphi_p = -\frac{a_{16} + a_{22}}{a_{12}}.
\]  

(B.6)
C Qualitative analysis

This appendix investigates the qualitative effects of changes in the state variables on the dependent variables of the model, by inferring the signs of the $A_0$ matrix coefficients of the solved form in price and stocks from previous knowledge of the signs of the structural parameters based on economic theory.

The adopted solution method, based on market fundamentals, prevent us from carrying out a traditional comparative statics analysis by differentiating the solved form (40), as the resulting terms are a combination of reduced form and structural parameters. Nevertheless, from plausible conjectures about the values of the structural parameters and the reduced form lagged endogenous coefficients, based on statistical grounds, we can tentatively predict the solved form signs using the logical and REH restrictions and the expressions derived in Appendix B. In what follows, we illustrate through propositions the procedure used with reference to market fundamentals as the effects of changes in the other state variables, reported in Table 9, are analogously obtained.

The first step consists in determining the signs of the structural parameters from the causal effects suggested by economic theory in the structural relationships.

Assumption 1. Structural parameters signs
The expected signs of the structural parameters in system (21)-(25) are the following:

\[
\varphi_p < 0, \quad \varphi_x > 0, \quad \beta_p > 0, \quad \beta_x > 0, \quad \beta_q > 0, \quad \eta_e > 0.
\]

These signs are justified by observing that the effect on demand of own price $\varphi_p$ and income $\varphi_x$ are respectively negative and positive. As concerns supply, the own price elasticity $\beta_p$ is expected to be positive, as such is the effect on production of a competitive devaluation of the exchange rate $\beta_x$. The positive sign of $\beta_q$ is related to the observed autocorrelation in production, as a result of the slow changes in the tree stock due to high adjustment costs, a typical feature of perennial crops. Finally, the positive stock demand elasticity to expected capital gains, $\eta_e$, is postulated by the theory of speculative storage.

Since $\lambda^{-1} = \beta_p - \varphi_p$, from ASS.1 it follows that $\lambda > 0$. This in turn allows to make the following statement

**Proposition B.1. Symmetry condition**
Given ASS.1 and the logical restrictions (42), the signs of corresponding coefficients in the price and stock equations coincide

\[
\text{sign}[a_{1j}] = \text{sign}[a_{2j}] \quad \text{for} \quad j = 2, 3, 4, 5.
\]

Assumption 2. Stationarity condition
The effect of previous period price on current price in the solved form is such that $0 < a_{16} < 1$. 

47
This assumption is based on the statistical analysis developed in the following chapter and on the desirable stationary property of the price series, which lead us to rule out unit and explosive roots.

**Proposition B.2. Long term fundamental effect**
Given ASS.1, ASS.2, PROP.B.1 and equation (B.5), the effect of the long-term market fundamental $z_{2t}$ on price and stocks is negative.

**Proof B.2**
Using ASS.1 and ASS.2 we can sign the parameters in equation (B.5) as follows

\[ \beta_p = \frac{a_{16}}{a_{12}} \]

from which it follows that $a_{12} < 0$ and, by PROP.B.1, $a_{22} < 0$.\(^{28}\)

**Proposition B.3. Short term fundamental effect**
Given ASS.1, ASS.2, PROP.B.1, PROP.B.2, equation (B.1) and the first logical restriction, the predicted effect of the short-term market fundamental $z_{1t}$ is negative on price and positive on stocks.

**Proof B.3**
These results hinge on further mild assumptions about the values of the structural parameters. In fact, the latter result follows from the first logical restriction, $a_{21} = 1 + \lambda^{-1}a_{11}$, by observing that, for plausible values of $\lambda$ and $a_{11}, a_{21} > 0$ regardless of the sign of $a_{11}$. In order to sign $a_{11}$, consider equation (B.1)

\[ \beta_q = \frac{a_{16} - 1}{a_{11} + a_{12}} \frac{a_{16}a_{21} - a_{11}a_{22}}{a_{21}} \]

where from PROP.B.1 we know that $a_{12} < 0$ and $a_{22} < 0$. We can note that the signs of the left and right-hand side members are coherent, without imposing any further assumptions on the values of the reduced form coefficients, only if $a_{11} < 0$. The same result holds also if we check the consistency of the signs in equation (A.3), assuming $a_{23} > 0$ and $a_{13} > 0$.\(^{\square}\)

Given these signs, it is easy to verify, using the REH restrictions (A.3) and (A.4), the positive effect of income $x_{1t}$ and exchange rate $x_{2,t-1}$ on both price and stocks. The predicted negative effect of $r_{t}$ in both equations, through (A.5), is instead a likely outcome for plausible values of $\rho_{3}$ and $a_{15}$. Finally, we observe how the consistency of the signs in (A.7) implies the further restriction on the reduced form parameters $a_{16} + a_{12} < 1$.

\(^{28}\)This result in turn implies that, for (B.6) to be valid, $|a_{16}| > |a_{22}|$.\(^{48}\)
Table 9: Comparative statics

<table>
<thead>
<tr>
<th>$z_{1t}$</th>
<th>$z_{2t}$</th>
<th>$x_{1t}$</th>
<th>$x_{2,t-1}$</th>
<th>$r_t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>↓</td>
<td>↓</td>
<td>↑</td>
<td>↑</td>
<td>↓</td>
</tr>
<tr>
<td>↑</td>
<td>↓</td>
<td>↑</td>
<td>↑</td>
<td>↓</td>
</tr>
</tbody>
</table>

D Data sources

Data used in the estimation have been collected from various sources. Figures on world production, grindings, stocks, apparent consumption and the world indicator price come from the International Cocoa Organization (ICCO). Data on area harvested, production, trade and producer prices for individual countries are from FAOSTAT database (FAO). Macroeconomic data such as GDP, exchange rates, interest rates, consumer price indexes and other international commodity prices come from the International Monetary Fund (IMF) and the World Bank (WB). We now provide definitions and further details on the single series.

Data on world production, grindings and stocks of cocoa beans are measured in thousand tonnes. The stock series is computed from the annual supply/demand balance and from two base year stocks estimates. The ICCO Secretariat uses an estimate of 325,000 tonnes in 1973/74 and of 1,682,000 tonnes in 2003/04 as the base year figures. World end-of-season stocks of cocoa beans are calculated as current production adjusted for loss in weight (net production) minus seasonal grindings plus previous stocks. Source: Quarterly Bulletin of Cocoa Statistics (ICCO).

World income is calculated as a weighted average of the GDP (in current US dollars) of major consuming countries. The weights are given by average apparent domestic consumption computed over the period 1997/1998 through 2005/2006. Source: World Development Indicators (WB)).

Apparent domestic consumption is given by the sum of grindings plus net imports of cocoa products, either finished or semiprocessed, expressed in beans equivalent. For the purpose of determining the beans equivalent of cocoa products, the following conversion factors are used: cocoa butter 1.33, cocoa cake and powder 1.18, cocoa paste/liquor and nibs 1.25. Source: Quarterly Bulletin of Cocoa Statistics (ICCO).

The world cocoa price is the crop year average of the ICCO daily price for cocoa beans, unit US$ per tonne. The latter price is calculated as the average of the quotations of the nearest three active futures trading months on NYSE Liffe Futures and Options and ICE Futures US at the time of London close.

29 The most important countries in terms of domestic apparent consumption are the United States, Germany, France, Belgium-Luxemburg, UK, Italy, Russian Federation, Japan, Brazil, Spain, Canada, Poland.
The international commodity prices used have been converted in US$ per tonne. Rubber and palm oil prices are for production originating from Malaysia, the coffee price is for the Robusta variety (Uganda origin), the one cultivated mostly in cocoa producing countries. *Source: International Financial Statistics (IMF).*

The exchange rate is given by the annual average of the franc CFA per US$ used in fourteen African countries, included Côte d’Ivoire and Cameroon. *Source: International Financial Statistics (IMF).*

The interest rate used is the annual average return on US three-month Treasury Bills. *Source: International Financial Statistics (IMF).*

The price deflator is the US consumer price index, all items (100=2004). *Source: International Financial Statistics (IMF).*

The food production index and the population figures are weighted averages for major producing countries. The weights used are the time series of the production shares for those countries. *Source: World Development Indicators (WB).*

---

30 The list of countries includes Côte d’Ivoire, Ghana, Nigeria, Cameroon, Indonesia, Malaysia, Brazil, Ecuador, Dominican Republic and Papua New Guinea.
References


P. Perron. The great crash, the oil price shock, and the unit root hypothesis. 


53