# Revisiting Structural Breaks in the Terms of Trade of Primary Commodities (1900–2020)-Markov Switching Models and Finite Mixture Distributions

Armand Taranco, Vincent Geronimi

# Revisiting Structural Breaks in the Terms of Trade of Primary Commodities (1900–2020)—Markov Switching Models and Finite Mixture Distributions †

**Armand Taranco *** and **Vincent Geronimi**

Cemotev, University of Versailles St Quentin Paris Saclay, 78047 Guyancourt, France; vincent.geronimi@uvsq.fr

* Correspondence: armand.taranco@uvsq.fr

† Presented at the 7th International conference on Time Series and Forecasting, Gran Canaria, Spain, 19–21 July 2021.

**Abstract:** This paper presents an analysis of the long-term dynamics of the terms of trade of primary commodities (TTPC) using an extended data set for the whole period 1900–2020. Following our original contribution, we implement three approaches of time series—the finite mixture of distributions, the Markov finite mixture of distributions, and the Markov regime-switching model. Our results confirm the hypothesis of the existence of a succession of three different dynamic regimes in the TTPC over the 1900–2020 period. It seems that the uncertainty characterising the long-term dynamic analysis of TTPC is better taken into account with a Markov hypothesis in the transition from one regime to another than without this hypothesis. In addition, this hypothesis improves the quality of the time series segmentation into regimes.

**Keywords:** commodity prices; terms of trade; long-term fluctuations; structural breaks; finite mixture of distributions; finite Markov mixture of distributions; Markov switching models

## 1. Introduction

One of the main conclusions emerging from the abundant literature dedicated to the study of the long-term evolution of primary commodities' prices is that structural breaks constitute an essential characteristic for the comprehension of the long-term dynamics of terms of trade of primary commodities. Empirical studies of price volatility assess a high level of uncertainty, especially in the post-2008 boom research [1]. However, this literature appears inconclusive on the question of the identification of structural breaks. In this paper, we explore this question by implementing three time series approaches—that have not been, to our knowledge, considered in this literature—for detecting these breaks. We identify structural breaks as the endpoints of the time periods obtained by clustering the data (mixture distributions) or as the endpoints of the regimes (Markov switching regimes). Following our original contribution [2] to the empirical literature on the Prebisch–Singer hypothesis [3,4] of a secular decline in the terms of trade of primary commodities (TTPC), in this paper, we consider an extension of our approaches to the whole period of 1900–2020. The data correspond to the Grilli and Yang Index, here after $\{GY_t\}_{t=t_1, \dots, t_N}$, see [5,6]. The three approaches of time series we implement—the finite mixture of distributions, the Markov finite mixture of distributions and the Markov regime-switching model—converge in the detection of three different regimes over the 1900–2020 period.

The three following sections of the paper present the methodology and results of, respectively, a finite mixture of distributions approach (Section 2) a finite Markov mixture of distributions approach (Section 3) and a Markov switching model approach (Section 4). The last section is dedicated to the discussion and conclusion (Section 5).

## 2. A Finite Mixture Distributions Approach

To investigate the hypothesis that the time series $\{GY_t\}_{t=t_1, \ldots, t_N}$ follows different periods over 1900–2020, we first used a finite mixture of distributions with normal components, as a way of putting similar data points (years) together into clusters (which we call regimes). Clusters are represented by the components' distributions of the mixture. The idea is that the years that exhibit the same behaviour belong to the same cluster and come from the same distribution.

A very detailed account of the practical aspects of Markov Chain Monte Carlo (MCMC) for mixture of distributions is given in Frühwirth-Schnatter [7]. The Handbook of mixture analysis [8] provides an overview of the methods of mixture modelling.

### 2.1. Methodology

A finite mixture of normal distributions can be defined as follows:

$$f(y) = \sum_{i=1}^{K} \eta_i f_i\left(y, \ \mu_i, \ \sigma_i^2\right) \text{ with } \sum_{i=1}^{K} \eta_i = 1,$$

where:

$K$ is the number of components,
$\eta_i$ is the mixing weight of the $i$th component, $f_i$ is a normal component distribution of mean $\mu_i$ and variance $\sigma_i^2$.

In this approach, three kinds of statistical inference problems have to be considered:

- The specification of the number of components $K$,
- The component parameters $(\mu_i, \ \sigma_i^2)$ and the weight distribution $(\eta_1, \ \ldots, \ \eta_K)$ should be estimated from the data, Finally, we must assign each observation of the time series, $\{GY_t\}_{t=t_1, \ldots, t_N}$, to a certain component of the mixture model by making inference on a hidden vector indicator $\boldsymbol{S} = (S_{t_1}, \ \ldots, \ S_{t_N})$.

To estimate the parameters of the components and the weights, we use Bayesian estimation [9] with MCMC [10] and a two block Gibbs sampling algorithm [7]:

(1) Parameter simulation conditional on the classification $\boldsymbol{S} = (S_{t_1}, \ \ldots, \ S_{t_N})$:

    a. Sample the weights $\boldsymbol{\eta} = (\eta_1, \ \ldots, \ \eta_K)$ from a Dirichelet posterior $p(\boldsymbol{\eta}|\boldsymbol{S})$,

    b. Sample the variances $\sigma_i^2$ in each group $i$, from an inverted Gamma distribution $G^{-1}(c_i(\boldsymbol{S}), C_i(\boldsymbol{S}))$,

    c. Sample the means $\mu_i$ in each group $i$, from an inverted Gamma distribution $G^{-1}(b_i(\boldsymbol{S}), B_i(\boldsymbol{S}))$

The precise form of $b_i(\boldsymbol{S}), \ B_i(\boldsymbol{S}), c_i(\boldsymbol{S}), \ C_i(\boldsymbol{S})$ depends upon the chosen prior distribution family.

(2) Classification of each observation $y_i$ conditional on knowing $\boldsymbol{\mu} = (\mu_1, \ldots, \ \mu_K)$,

$$\sigma^2 = \left(\sigma_1^2, \ldots, \ \sigma_K^2\right) \text{ and } \boldsymbol{\eta} = (\eta_1, \ \ldots, \ \eta_K):$$

$$P\left(S_i = k \middle| \boldsymbol{\mu}, \ \sigma^2, \boldsymbol{\eta}, \ y_i\right) \propto \frac{1}{\sqrt{2\pi\sigma_k^2}} exp\left\{-\frac{(y_i - \mu_k)^2}{2\sigma_k^2}\right\}\eta_k$$

The number of components may be known or unknown. In our case, the number of components is unknown, and our model selection is based on marginal likelihood [11]. In the academic literature, the unknown number of regimes taken into account is three at most. To determinate the best model, we expand the number of potential regimes to five. Thus, we chose the model with the largest marginal likelihood, approximated by three estimators [7]:

- RI is the estimator obtained by reciprocal importance sampling,
- IS is the estimator obtained by importance sampling,

- BS is the estimator obtained by bridge sampling techniques.

  For computing purposes, we use the Matlab library Bayesf 2.0 in this publication.

*2.2. Results*

The results are presented in the following four sections. First, we confirm the existence of three different components in the mixture. Then, we present the statistical parameters (mean, standard deviation, and weight) of each distribution, associated to the correspondent regime (regime1: 1900–1921; regime 2: 1922–1985 and 2006–2020; regime 3: 1986–2005). The third sub-section presents a point representation of posterior draws. The fourth sub-section clusters the data based on MCMC draws.

In all Monte Carlo simulations using posterior draws, we use 1,000,000 draws after a burn-in of 100,000 draws.

2.2.1. The Choice of the Number of Components

If $K$ is not too large, the different estimators should approximatively agree. As $K$ increases, we observe that, the reciprocal importance sampling and the importance sampling estimators are less precise than the bridge sampling estimator, although all three select the same number of components: among the considered models (number of components $\leq$ 5), the model with the largest marginal likelihood is a mixture of three normal distributions.

Thus, the results (Table 1) for the mixture of distribution models confirm the accuracy of the hypothesis of the existence of three different components, as already established in our previous analysis for the 1900–2016 period.

**Table 1.** The choice of the number of components according to three estimators—Source: authors.

| Estimators | $K = 1$ | $K = 2$ | $K = 3$ | $K = 4$ | $K = 5$ |
|---|---|---|---|---|---|
| RI | −20.6488 | −21.8398 | −16.9335 | −17.3682 | −21.2979 |
| Standard error | $8.2511 \times 10^{-5}$ | $1.0581 \times 10^{-3}$ | $3.2659 \times 10^{-3}$ | $6.9481 \times 10^{-2}$ | $5.1914 \times 10^{-1}$ |
| IS | −20.6488 | −21.8456 | −16.9402 | −17.1843 | −19.2823 |
| Standard error | $8.0611 \times 10^{-5}$ | $2.6576 \times 10^{-3}$ | $4.2801 \times 10^{-3}$ | $1.0072 \times 10^{-1}$ | $1.199 \times 10^{-1}$ |
| BS | −20.6489 | −21.8402 | −16.9316 | −17.1489 | −17.7954 |
| Standard error | $5.581 \times 10^{-5}$ | $7.2275 \times 10^{-4}$ | $9.0533 \times 10^{-4}$ | $2.4614 \times 10^{-3}$ | $6.2907 \times 10^{-3}$ |

2.2.2. The Parameters of the Mixture of Three Normal Distributions

The components of the mixture differ mainly in the mean. Components 2 and 3 have nearly the same variance, whereas the first component has a variance that is slightly higher (Table 2).

**Table 2.** Weight, mean, and standard deviation of each distribution—Source: authors.

| Parameters of the $k$th Component | Distribution 1 | Distribution 2 | Distribution 3 |
|---|---|---|---|
| Weight | 0.3238 | 0.4945 | 0.1817 |
| Mean | 4.9091 | 4.5876 | 4.1829 |
| Standard deviation | 0.0246 | 0.0096 | 0.0153 |

2.2.3. The Point Process Representation of Posterior Draws

To produce sampling representations of the posterior draws, $\left\{ \mu^{(m)} \right\}_{m=1, \ldots, M}$ ($M$, the number of draws) is plotted against $\left\{ \sigma^{2(m)} \right\}_{m=1, \ldots, M}$. This scatter plot is closely related to the point process representation of the underlying mixture distribution. A finite mixture distribution from a fixed parametric family has a representation as a marked point process [12]. Here, we use point process representation (Figure 1) of draws from the posterior density. Three clusters of draws are distinguished, they will scatter around the

three points corresponding to the true point process representation, with the spread of the clouds representing the uncertainty of estimating the points (Figure 1).
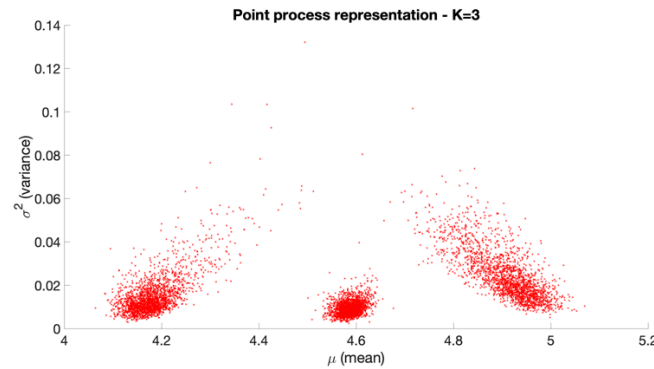


**Figure 1.** Point process representation for $K = 3$—Source: authors.

2.2.4. Clustering the Data

We perform clustering of the data into three groups (Figure 2) based on the MCMC draws. Three criteria are used:

- The Bayesian maximum a posteriori (MAP),
- The similarity matrix based on the posterior similarity,
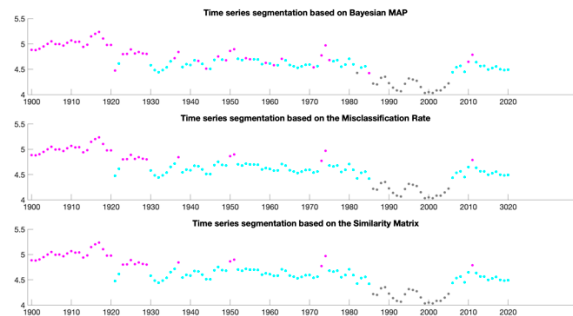- The misclassification rate.



**Figure 2.** Time-series segmentation according to three methods.

Three regimes are confirmed, (1900–1921; 1922–1985 and 1986–2020; 1986–2005). The second regime is interrupted by the regime 1986–2005, which represents the lowest level in the terms of trade of primary commodities (see Section 5). However, we observe that some years have an ambiguous cluster membership.

## 3. A Finite Markov Mixture Distributions Approach

### 3.1. Methodology

In the finite mixture models approach, we assign each observation of the time series $\{GY_t\}_{t=t_1, \ldots, t_N}$ to a certain component of the mixture model by making inference on a hidden vector indicator $S = (S_{t_1}, \ldots, S_{t_N})$. Now, we suppose that this allocation vector is a hidden Markov chain, $GY_t = \mu_{S_t} + \varepsilon_t$ where $\varepsilon_t$ is a zero-mean white noise process with variance $\sigma^2$, which is a special case of interest of finite Markov mixture of distributions. Now, the transition probability matrix $T$ of the hidden Markov chain $S = (S_{t_1}, \ldots, S_{t_N})$ is unknown and need to be estimated from the data. We suppose that the Markov chain is aperiodic and starts from its ergodic distribution $\eta = (\eta_1, \ldots, \eta_K)$:

$$P(S_N = k | T) = \eta_k$$

and the transition probability matrix $T$ is defined by:

$T_{ji} = P(S_{t+1} = j | S_t = i)$ for $i, j = 1, \ldots, K$ and $t = t_1, \ldots, t_N - 1$.

What is the relation between finite mixture distributions and finite Markov mixture distributions? In fact, every finite mixture of distributions may be considered of as a limiting case of a finite Markov mixture of the same family of distributions where $S = (S_{t_1}, \ldots, S_{t_N})$ is an i.i.d. random sequence and where the transition probabilities are all equal to $\eta_k$.

### 3.2. Results

The results are presented in the following three sub-sections. We present the statistical parameters (mean and standard deviation) of each distribution, associated to the correspondent regime (regime1: 1900–1921; regime 2: 1922–1985 and 2006–2020; regime 3: 1986–2005) and transition probabilities from one regime to another one. The second sub-section presents a point representation of posterior draws. The third sub-section clusters the data based on MCMC draws.

### 3.2.1. The Parameters of the Markov Mixture of Three Normal Distributions

The components of the mixture differ mainly in the mean but have nearly the same variance (Table 3).

**Table 3.** Mean and standard deviation of each distribution.

|  | Distribution 1 | Distribution 2 | Distribution 3 |
| --- | --- | --- | --- |
| Mean | 5.0099 | 4.6265 | 4.1822 |
| Standard deviation | 0.0110 | 0.0142 | 0.0134 |

The transition probabilities $T_{11}$, $T_{22}$, $T_{33}$ are high (Figure 3, Table 4), which indicate that is difficult to change from on regime to the other.
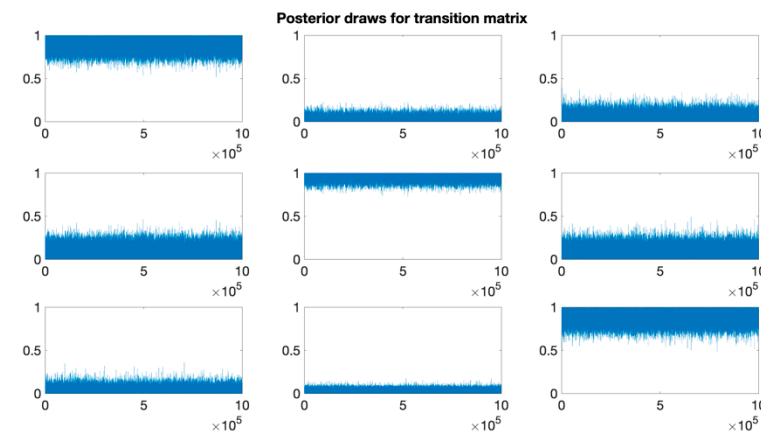


**Figure 3.** Posterior draws for transition probability matrix T—Source: authors.

**Table 4.** Transition probability matrix T from posterior draws—Source: authors.

|  | Regime 1, *t* | Regime 2, *t* | Regime 3, *t* |
| --- | --- | --- | --- |
| **Regime 1, *t* + 1** | 0.9384 | 0.0167 | 0.0284 |
| **Regime 2, *t* + 1** | 0.0483 | 0.9637 | 0.0571 |
| **Regime 3, *t* + 1** | 0.0133 | 0.0196 | 0.9146 |

### 3.2.2. Point Process Representation of Posterior Draws

We observe that this time, the clusters obtained with the point process representation of posterior draws in the case of a Markov finite mixture (Figure 4) are well-separated

and have less dispersion compared with that of the clusters obtained in the case of a finite mixture of distributions. The shapes of the clusters are also different.
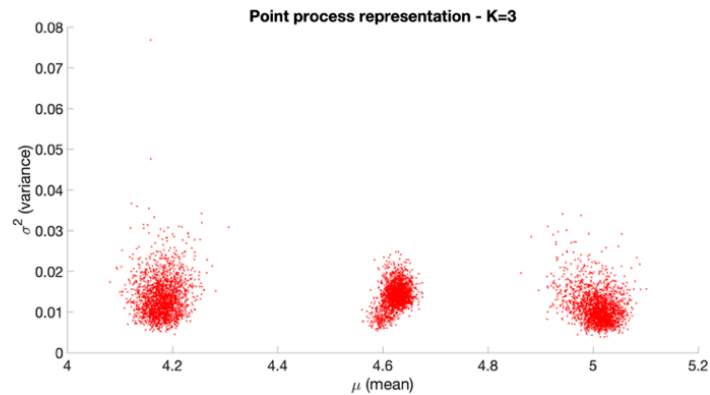


**Figure 4.** Point process representation for *K* = 3—Source: authors.

3.2.3. Clustering the Data

We confirm the existence of the three regimes previously found (1900–1921; 1922–1985 and 1986–2020; 1986–2005). This time, all the years have a perfect cluster membership. The periods of the regimes are well defined (Figure 5).
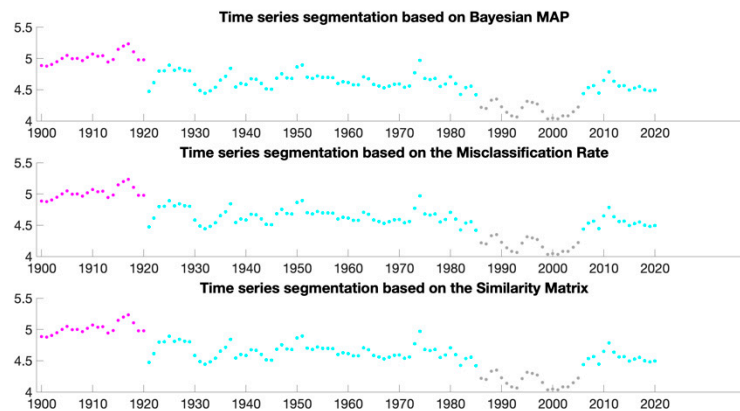


**Figure 5.** Time-series segmentation according to three methods.

## 4. A Markov Switching Model Approach

### 4.1. Methodology

A finite Markov switching (MS) model assumes that the dynamics of a data series, $\{y_t\}_{t=t_1, \, \ldots, \, t_N}$, depend on a discrete latent variable $S_t$, postulated to follow a Markov chain with realizations in $\{1, \ldots, K\}$. This model was popularized by Hamilton [13,14] who applied the Markov-switching approach to model the probability of a recession in the U.S. economy. In this model, the economy alternates between two unobserved states of economic expansion and recession according to a Markov chain process. The model assumes constant transition probabilities for the unobserved states, which, in turn, imply constant expected durations in the various regimes. A general representation is given by:

$$y_t = C_{S_t} + \sum_{i=1}^{p} \alpha_i X_t^{f,i} + \sum_{i=1}^{q} \beta_i(S_t) X_t^{r,i} + \sum_{i=1}^{r} \gamma_i(S_t) y_{t-i} + \varepsilon_t$$

where:

$y_t$ denotes the series observed,

$X_t^{f,i}$ are the independent regressors with fixed effects,

$X_t^{r,i}$ are the independent regressors with random effects,

$y_{t-i}$ these variables represent the autoregressive part of model,

$\varepsilon_t$ are independent variables with N $(0, \sigma^2_{\varepsilon,S_t})$ distribution,

$S_t$ is modelled by a homogeneous Markov chain with *K* states.

The transition probabilities verify:

$P(S_{t+1} = j | S_t = i) = P(S_2 = j | S_1 = i)$, for $t = t_1, \ldots, t_N - 1$ and for $i, j = 1, \ldots, K$ (homogeneity of the chain).

For $i = 1, \ldots, K$:

$$\sum_{j=1}^{K} P(S_{t+1} = j | S_t = i) = 1.$$

We consider only the case where there is no fixed or random effects and no autoregressive part in the model.

Essentially, two computational approaches can be used for the estimation of Markov-switching models. One approach involves maximising the log-likelihood, a function of the transition probabilities, subject to the constraint that the probabilities lie between 0 and 1 and sum to unity. This can be done with the EM algorithm [15], but the non-linear programming approach [16] can also be used. We mobilise this last approach implemented in Oxmetrics. An alternative approach involves using Bayesian estimators with MCMC methods.

### 4.2. Results

The results of a three-regime model based on the terms of trade of commodities are shown in Tables 5 and 6, and Figure 6. There is a perfect match with the previous results, notably concerning the identification of three regimes over the exact same sub-periods.

**Table 5.** Statistical characteristics of regimes—Source: authors.

| Regimes | Coefficient | Standard Error | *t*-Value | *p*-Value |
|---------|-------------|----------------|-----------|-----------|
| Regime 1 | 5.01765 | 0.02014 | 249. | 0.000 |
| Regime 2 | 4.63002 | 0.01365 | 339. | 0.000 |
| Regime 3 | 4.18067 | 0.02514 | 166. | 0.000 |

**Table 6.** Transition probability matrix.

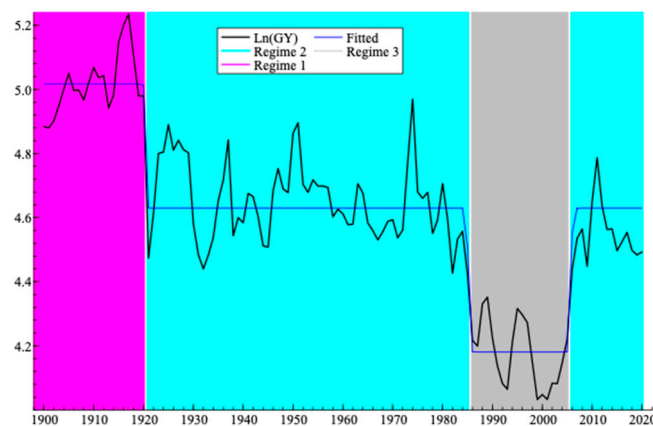|  | Regime 1, *t* | Regime 2, *t* | Regime 3, *t* |
|--|---------------|---------------|---------------|
| **Regime 1, *t* + 1** | 0.95451 | 0.0000 | 0.0000 |
| **Regime 2, *t* + 1** | 0.045485 | 0.98722 | 0.048562 |
| **Regime 3, *t* + 1** | 0.0000 | 0.012781 | 0.95144 |



**Figure 6.** Change of regime in the evolution of the terms of trade for primary commodities (Neperian logarithm, 1900–2020)—Source: authors.

## 5. Discussion and Conclusions

The existence of different regimes appears robust to various changes in the data span. Indeed, considering data from 1900 to 2010, or 1900 to 2014, or 1900 to 2016 or 1900 to 2020 on the same $\{GY_t\}_{t=t_1, \ldots, t_N}$ index leads to the same representation, with the same break dates (1921, 1986, 2006). The approach using a Markov finite mixture of distributions and the approach using a Markov switching model give very similar results. These two methods differ essentially in the computational aspects. The former uses Bayesian estimation with MCMC and the later involves maximising the log-likelihood. The fact that each observation of the time series $\{GY_t\}_{t=t_1, \ldots, t_N}$ is assigned to a certain component of the Markov mixture model by making inference on a hidden Markov vector indicator, improves the results obtained with the finite mixture model. This time, all years have a perfect cluster membership.

These three approaches applied to the extended 1900–2020 data set confirm the identification of a succession of three different dynamic regimes in the TTPC over the 1900–2020 period. The third regime (1986–2005) is still characterized by the lowest level of terms of trade of the whole period, and the return to the second regime after 2005 is associated with a price significantly higher (56.7% higher). Such an upward shift in primary commodities' prices is unprecedented at the scale of the 20th century and questions more specifically the hypothesis of a secular decline in the terms of trade of primary commodities. Indeed, the entry into a higher level of prices contradicts the hypothesis of a secular decline. However, from 1900 to 2006, this decline manifested itself through the succession of regimes with a lower average level of primary commodity terms of trade, but not in a continuous way. Moreover, data from 2020 for TTPC do not exhibit a specific pattern, leaving open the question of the effect of COVID on the long-term dynamics of primary commodity prices. Therefore, the dynamics behind the evolution of primary commodities in the long-run call for alternative explanations and a change of perspective.

This paper contributes to this change of perspective by considering (and confirming) the existence of three changes in regime in the long term (121 years). Yet, an operational theory of long-term dynamic regime change in primary commodities' terms of trade is still to be constructed.

Following the methodologies used in this present paper, a promising perspective appears to be the introduction of explanatory variables (such as the GDP of main countries, the share of emerging countries in the global GDP, and various indices of real interest rate and exchange rates) in a Markov switching model, in order to identify the incidence of these covariates on the dynamic regimes.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Data Available at https://www.cemotev.uvsq.fr/base-de-donnees.

## References

1. Prakash, A. *Safeguarding Food Security in Volatile Global Markets*; FAO: Rome, Italy, 2011.
2. Geronimi, V.; Taranco, A. Revisiting the Prebisch-Singer hypothesis of a secular decline in the terms of trade of primary commodities (1900–2016). A dynamic regime approach. *Resour. Policy* **2018**, *59*, 329–339. [CrossRef]
3. Prebisch, R. The Economic Development of Latin America and its Principal Problems. *Econ. Bull. Lat. Am.* **1962**, *7*, 1–22.
4. Singer, H.W. U.S. Foreign Investment in Underdeveloped Areas: The distribution of gains between Investing and Borrowing countries. *Am. Econ. Rev. Pap. Proc.* **1950**, *40*, 473–485. [CrossRef]
5. Grilli, E.R.; Yang, M.C. Primary commodity prices, manufactured goods prices, and the terms of trade of developing countries: What the long run shows. *World Bank Econ. Rev.* **1988**, *2*, 1–47. [CrossRef]
6. Geronimi, V.; Anani, E.T.G.; Taranco, A. Notes on updating prices indices and terms of trade for primary commodities (No. 3–2017). *Cahier CEMOTEV* **2017**, *2017*, 3.
7. Frühwirth-Schnatter, S. *Finite Mixture and Markov Switching Models; Springer Series in Statistics*; Springer: New York, NY, USA, 2006. [CrossRef]
8. Celeux, G.; Frühwirth-Schnatter, S.; Robert, C.P. (Eds.) *Handbook of Mixture Analysis*; Chapman and Hall/CRC: Boca Raton, FL, USA, 2018. [CrossRef]

9.    Diebolt, J.; Robert, C. *Bayesian Estimation of Finite Mixture Distributions, Part I: Theoretical Aspects*; Rapport Technique LSTA; Université Paris VI: Paris, France, 1990; Volume 110.

10.   Robert, C.; Casella, G. A short history of Markov chain Monte Carlo: Subjective recollections from incomplete data. *Stat. Sci.* **2011**, *26*, 102–115. [CrossRef]

11.   Frühwirth-Schnatter, S. Keeping the balance—Bridge sampling for marginal likelihood estimation in finite mixture, mixture of experts and Markov mixture models. *Braz. J. Probab. Stat.* **2019**, *33*, 706–733. [CrossRef]

12.   Stephens, M. Bayesian analysis of mixture models with an unknown number of components-an alternative to reversible jump methods. *Ann. Stat.* **2000**, *28*, 40–74. [CrossRef]

13.   Hamilton, J.D. A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica* **1989**, *57*, 357. [CrossRef]

14.   Hamilton, J.D. Regime-switching models. In *Palgrave Dictionary of Economics*; Durlauf, S., Blume, L., Eds.; Palgrave McMillan Ltd.: London, UK, 2005. [CrossRef]

15.   Dempster, A.P.; Laird, N.M.; Rubin, D.B. Maximum likelihood from incomplete data via the EM algorithm (with discussion). *J. R. Statist. Soc. Ser. B* **1977**, *39*, 1–38. [CrossRef]

16.   Lawrence, C.T.; Tits, A.L. A computationally efficient feasible sequential quadratic programming algorithm. *Siam J. Optim.* **2001**, *11*, 1092–1118. [CrossRef]