v. 5, n. 1, 2025

Journal of Agricultural Sciences Research

Acceptance date: 22/01/2025

FORECASTING COFFEE EXPORTS IN BRAZIL USING HIERARCHICAL MODELS

Gabriel Amil Bastos
Universidade de São Paulo
https://orcid.org/0000-0001-6767-4075

José Erasmo Silva
Universidade Federal da Bahia
https://orcid.org/0000-0001-9075-3548

Abstract: The coffee export market in Brazil, the world's largest producer and exporter, faces significant challenges due to the volatility of international prices and variability in regional production. These factors make export forecasting a complex but crucial task for the economic sustainability of the sector. This study aimed to develop and apply a hierarchical model, integrating regional and temporal variables, to forecast coffee exports from Brazil. The research used data from six exporting states (Minas Gerais, Espírito Santo, São Paulo, Bahia, Paraná and Rio de Janeiro) from January 2012 to May 2022. For the analysis, two-level hierarchical models were used to capture fixed and random effects, comparing them with traditional multiple linear regression models. The methodology included transforming the data into natural logarithms, applying lags to dependent variables and using the regional economic activity index (IBCR) as a proxy for state GDP. The analysis was conducted using the R programming language, with specific packages for hierarchical modeling and descriptive statistics. The results indicated that the hierarchical model offered more accurate forecasts, excelling in capturing regional and temporal fluctuations. These forecasts improve planning and decision-making in the coffee sector, as well as reinforcing the need for public policies that consider regional particularities and help mitigate the effects of market volatility. The study confirmed that, contrary to initial expectations, the exchange rate was the main determinant of exports, while regional variables such as the Regional Economic Activity Index (IBCR) played a secondary, albeit significant, role. These findings suggest that economic and trade policies better adapted to regional realities could improve Brazil's competitiveness in the international coffee market.

Keywords: economic analysis; Brazilian coffee; hierarchical models; regional economic dynamics

INTRODUCTION

Technological developments, combined with advances in communication methods and networks, have significantly transformed the business environment, allowing large volumes of data to be integrated into the organizational routine. This development has made it possible to replace manual, static analysis with more sophisticated approaches based on advanced data processing. Data-driven decision making (DOD) has established itself as a superior alternative to intuitive practices, demonstrating substantial improvements in business results (Fawcett and Provost, 2016). However, despite these advances, many sectors still face significant challenges in applying these technologies effectively, especially in complex markets such as the export of agricultural commodities.

In this context, the need to ensure accurate forecasts that can guide strategic decisions is becoming increasingly evident. In Brazil, the coffee sector, one of the main pillars of agribusiness and fundamental to the national economy, stands out as a critical area for the application of these approaches. The literature points to the growing need for robust tools to help mitigate the risks associated with market volatility and production variability (Campos, 2022). Predictability in this sector is not just a technical issue, but also a strategic necessity for maintaining Brazil's global competitiveness in the coffee market.

Brazil - the world's largest coffee producer, with a 39.73% share of the international market, and the largest exporter, with 31.36% of global exports in the 2020/21 harvest (USDA, 2021) - is in a prominent position. In addition, unroasted coffee was the country's 11th most exported product in 2021, contributing 2.07% of national exports (Comex Stat, 2022; Paula et al., 2024; Haack et al., 2023). Given its importance, the coffee sector has become central to discussions on public policies, business strategies and commercial decisions, both for producing organizations and government bodies.

Forecasting exports, which is a significant source of income for coffee producers, is essential for making informed decisions. These forecasts help with production planning, price setting and stock management, as well as increasing the reliability of market indicators and futures contracts. Accurate information on future exports allows producers to optimize their operations and adapt their commercial strategies more effectively (Carrasco-Gutierrez and Almeida, 2013). Therefore, accurate forecasts not only improve internal management, but also ensure proper information governance, serving as a benchmark for the entire sector.

Although several predictive models are available in the literature, many focus on national-level analyses, limiting their applicability at smaller levels, such as companies or state governments (Tanahashi and Caldarelli, 2021). This study seeks to fill this gap by proposing a two-level hierarchical model capable of differentiating coffee export forecasts based on the interactions between producing states and periods of the year. The model was designed to capture the regional and temporal complexities of exports, allowing for more detailed analysis tailored to the specific needs of each context. In this way, the study aims to improve planning and decision-making in the Brazilian coffee sector, offering a robust and flexible tool that meets both national demands and the particularities of the main exporting states.

MATERIAL AND METHODS

To achieve the objectives of this study, variables from various sources were used, with the value exported serving as the dependent variable, calculated for each state in different months and years. This data was taken from the Comex Stat portal (2022) - under the responsibility of the Federal Government's Ministry of Industry, Foreign Trade and Services, linked to the Ministry of Economy.

The value exported, in real (R$), was deflated using the General Market Price Index (IGP-M), following the methodology proposed by Monteiro (2012). For the states in the second level, economic variables of state production were used, represented by the states' wealth expectations, based on the Regional Economic Activity Indices (IBCR), published by the Central Bank of Brazil (Bacen).

For both the dependent variable and the explanatory variables - with the exception of the "dummies" - a natural logarithm transformation process was applied, facilitating both interpretation, in order to visualize elasticity, and reducing the influence of extreme values when applying the models.

EMPIRICAL MODEL

The first level - corresponding to time - included variables that remain constant between states and products, but which show significant variations over time and can influence the results. These include: the dollar exchange rate, provided by the Central Bank of Brazil (Bacen), and the prices of coffee futures contracts traded on the B3 (Brazilian Stock Exchange). In addition, domestic coffee prices were considered, both Arabica and Robusta, expressed in reais (R$) and dollars (US$), available from the Center for Advanced Studies in Applied Economics (CEPEA) of the Luiz de Queiroz College of Agriculture (ESALQ) of the University of São Paulo (USP). The General Market Price Index (IGP-M), calculated by the Getúlio Vargas Foundation (FGV), was also used, both for inflationary adjustments in the series of exported values and in the statistics from the Central Bank of Brazil. Finally, the reference months were used as a proxy for seasonality and the years as a proxy for trend.

The variables selected were assessed on the basis of a literature review of other proposed studies on forecasting coffee exports (Table 1). Some variables were used as "proxies" for other proposals in the event that the same time breakdown was not available or other variables were not available for the second level context. These were the values referring to the state's wealth and production measured using the Regional Economic Activity Index (IBCR), replacing the Gross Domestic Product (GDP).

When proposing a model in the predictive context that varies over time (longitudinal data), combinations and transformations of variables can also be used, such as "lags" of the dependent variable, and in this case some were tested as explained in the literature. For the proposed models, data from January 2012 to May 2022 was considered.

Based on the construct proposed in this study, a hierarchical regression model was estimated to analyze Brazil's coffee exports, incorporating the nested structure of the data. The data was organized at different levels: exports by federal state over specific periods. Each state, in each period, was influenced both by state-specific variables and by consistent temporal variables between them. This multi-level approach made it possible to capture intra- and inter-state variations, offering a more detailed analysis of the dynamics of coffee exports. The use of the hierarchical model was essential to accommodate intra-group correlation and reflect the complexity of the data, composed of repeated measures organized under states over time.

The organization into distinct contexts made it possible to apply hierarchical (multilevel) models with repeated measures. These models offered advantages over linear models estimated by ordinary least squares (OLS), especially in their ability to estimate and analyze contexts, which are naturally incorporated into hierarchies. They made it possible to explore heterogeneities and include specific random components for each combination of data. In short, in addition to estimating the standard parameters, this method also facilitated the estimation of random effects and error terms. This was important for identifying whether, in the contexts analyzed, there were significant variations in the intercepts and slopes for specific groups or contexts (Fávero et al., 2023).

The problem presented, structured in two hierarchical levels, required the use of Hierarchical 2-Level Models with repeated measures (HLM 2). The time variable, transforming the dependent variable into a longitudinal vector, was presented as the first level, followed by the coffee producing and exporting states at level 2. As specified by Fávero et. al. (2023), HLM 2 was composed of 2 sub-models, in which there $t=1,...,T_i$ were years at level 1, which were nested in each $i=1,...,n$ state. Thus, for level 1, we have Equation (1):

$$y_{ti} = \pi_{0i} + \pi_{1i}.ano_{ti} + e_{ti} \quad (1)$$

where, π_{0i}: expected value of the exports variable (average) for each state i in the year/month; π_{1i}: growth rate of the exports variable for states i; $year_{ti}$: level 1 explanatory variable (repeated measure) corresponding to each year/month of the sample of exporting states; $t=1,...,T_i$ (years/months) and $i=1,...,n$ (producing states). In addition, it was assumed that the random term ~ N(0, σ^2).

According to Fávero et. al. (2023) each coefficient estimated at level 1 by , where is the index of the variable and *i* are the states that nest the information, becomes a dependent variable in the level 2 model. This gives Equation (2) for level 2

$$\pi_{pi} = \beta_{p0} + \sum_{q=1}^{Q_p} \beta_{pq}.x_{qi} + r_{pi} \quad (2)$$

where, β_{p0}: represents the intercepts of the level 2 estimates for each level 1 variable π_{pi}, with *p* being the index of the variable in the level 1 model, and *q=0* not being represented as the coefficient of a variable included in level 2;β_{pq}: represents the coefficient of the level 2 variables (states), where *q* is the index of the variable in the level 1 model, and *q* is the index of the predictor variables included in the level 2 model;

Variable	Type of variable	Description	Source	Reference
inflated_fob_value	R$	Value of coffee exports in real terms, deflated by the IGP-M	Comex	Monteiro, 2012
us	R$	Commercial dollar exchange rate at the end of the month	Bacen	Miranda; Coronel; Vieira, 2014
ibcr	Contents	Regional economic activity indices	Bacen	Carrasco-Gutierrez and Almeida, 2013
igpm	Contents	General Market Price Index	FGV	-
coffee_quotation	US$/Bag 60 kg	Closing price of coffee futures	B3	Miranda; Coronel; Vieira, 2014
price_arabic_dol	US$/Bag 60kg	Average domestic price of Arabica coffee in dollars	CEPEA / ESALQ / USP	Monteiro, 2012
price_robust_dol	US$/Bag 60kg	Average domestic price of Robusta coffee in dollars	CEPEA / ESALQ / USP	Monteiro, 2012
month	Binary	Dummy variables representing export months	Comex	Tanahashi and Caldarelli, 2021
year	Binary	Dummy variables representing years of export	Comex	Tanahashi and Caldarelli, 2021
vl_fob_def_lag1	R$	1-month lag of the value of exports by state in R$ at real values, deflated by the IGP-M	Comex	Tanahashi and Caldarelli, 2021
vl_fob_def_lag2	R$	2-month lag of the value of Exports by state in R$ at real values, deflated by the IGP-M	Comex	Tanahashi and Caldarelli, 2021
KG_Liquid_lag1	kg	1-month lag in the quantity of coffee exports	Comex	Tanahashi and Caldarelli, 2021
KG_Liquid_lag2	kg	2-month lag in the quantity of coffee exports	Comex	Tanahashi and Caldarelli, 2021

Table 1. Variables used and literature review

Source: Original research data

x_{qi}: is the vector of predictor variables of level 2 (states), where q is subscripted for predictor variables for each i state; r_{pi}: is the random effect of level 2. It is assumed that, for each unit i, the vector (r_{0i}, r_{1i}, ..., r_{pi}) has a multivariate normal distribution, in which each element r_{pi} has zero mean and variance $Var(r_{pi})=\tau_{\pi pp}$.

For the different methods, two forms of variable selection were used: the "stepwise" method for Ordinary Least Squares (OLS) and the "step-up" method for Hierarchical 2-Level Models with Repeated Measures (HLM 2). In the stepwise method, only variables that are statistically significant, using the Student's t-test, are selected in the estimation of the multiple regression models, leaving only variables that have an independent influence statistically different from zero. In the case of the "step-up" method, variables that alter the model can be evaluated, being inserted one by one, which makes it possible to discern their joint action through changes in the model's result (Armstrong and Hilton, 2010). These variables can remain in the model even when they are not independently statistically significant, especially in cases where the explanatory variables are correlated with each other, but not necessarily with the dependent variable. This evaluation makes it possible to more rigorously remove variables that make small changes to the model's fit or that remain when they change its fit but do not show statistically significant results from Student's t-tests. The step-up method is ideal for cases such as the HLM 2 model, since it has several nested estimates and the stepwise method is not applicable. In this study, the "step-up" method was evaluated by means of changes in the logarithm of the likelihood (LogLik) of the models.

In a hierarchical model such as HLM, LogLik is particularly useful for understanding how different levels of the model contribute to the variability in the data and for effectively adjusting the parameters at each level. LogLik helps assess whether including more levels or variables in the model actually improves the model's ability to explain the variability in the observed data.

The comparison between the models, carried out using the likelihood ratio test, followed the "step-up" approach recommended by Raudenbush and Bryk (2002), with the aim of identifying the model that maximizes the likelihood function ('LogLik').

The HLM has significant advantages for data with a grouped or hierarchical structure, as in the case of coffee exports organized by states in a country. This approach allows the model to accommodate variance at both the group and individual levels, more accurately reflecting the differences within each group. In addition, the HLM controls for heterogeneity between groups by incorporating random effects, which is particularly useful for data in which export behavior can vary substantially between states.

On the other hand, the OLS model is valued for its simplicity and ease of use, offering a quick and straightforward form of statistical modeling without the need for complex adjustments. Its wide applicability makes it a robust resource for initial and exploratory analyses. However, OLS operates under the assumption of independence between observations, which can result in biased and ineffective estimates in situations where the data is hierarchical or grouped.

In conclusion, the choice between HLM and OLS should be guided by the nature of the data and the specific needs of the analysis. Although both models show similar levels of accuracy in calculations, HLM excels in contexts with complex, structured data, while OLS is more suitable for direct analysis on less complicated data sets.

The main objective of this research was to understand which statistical model offered the best representation of the complexities of Brazilian coffee exports, taking into account variations both at state level and over time. To achieve this goal, different statistical models were compared to determine which one best fitted the data. The likelihood ratio test proved to be one of the most appropriate tools. This test makes it possible to assess whether the inclusion of more variables and details in the Hierarchical 2-Level Models with Repeated Measures (HLM 2) really improved the ability to predict export patterns compared to simpler approaches, such as the models estimated by ordinary least squares (OLS). In applying this test, we sought to identify the model that not only fitted the data adequately, but also effectively captured the nuances inherent in the object of study. The test was used following Equation (3) (Fávero et. al., 2023):

$$X^2_{1g.l.} = -2. \text{ (LL modelo final - LL modelo completo)} \quad (3)$$

where, $x^2_{1g.l.}$: represents the x^2 (chi-square) for 1 degree of freedom, in which we used, through the statistical tables, that $X^2=3,841$ (X^2 critical for 1 degree of freedom and for a significance level of 5%); *LL*: the log of the sum of the likelihood function, both in the final model and in the complete model.

In other words, this study chose to compare the Ordinary Least Squares (OLS) and Hierarchical Linear Multilevel Model (HLM 2) techniques, demonstrating a response to the complexity of data structured at different levels. This choice was influenced by Goldstein's (2011) discussion on the essentiality of considering the complex interactions between variables in multilevel studies, as well as Courgeau's (2003) argument, which emphasized the importance of transcending disciplinary barriers for a deeper understanding of the phenomena investigated, in order to overcome the limitations presented by traditional regression models.

In addition to evaluating the fit of the models using LogLik, the predictive power of the model was evaluated both using the fit graph and the metrics commonly used for this purpose, which are Mean Absolute Percentage Error (MAPE), Mean Absolute Error (MAE) and Root Mean Square Error (RMSE).

In this study, the R programming language and its available packages were used for both the estimates and the statistical tests.

RESULTS AND DISCUSSION

DESCRIPTIVE ANALYSIS

By evaluating the distribution, using the Kernel Density Estimate (KDE), it was possible to verify the existence of an asymmetrical distribution in the model's dependent variable, when in its initial proposed form in reais (R$) and in real values as of July 2022, deflated by the IGP-M (Figure 1).

The concentrated distribution close to zero, with a sharp peak at the lower end of the value scale, indicated that most states exported relatively small quantities of coffee (Figure 1). The density curve declined rapidly as the value of exports increased, suggesting that high export values are rare. This pattern indicated a concentration of smaller exporters in the market, with few states contributing large volumes to total coffee exports. Such a visualization highlights the importance of having economic support policies not only for large exporters, but also to improve the exporting capacity of the majority of states operating at lower export levels.

Kernel density estimation was applied to the logarithm of the value of coffee exports (Figure 2), providing an alternative view that balanced the variation on different scales of magnitude (Monteiro, 2012). The curve displayed a more complex profile with two main peaks, revealing a possible bimodal distribution. The first peak was located around a logarithmic value of 10, while the second slightly preceded the value of 20, suggesting the existence of two distinct groups of exporting states: one group with moderate export values and another with substantially higher values. This logarithmic representation highlighted nuances that are not apparent on the linear scale, emphasizing the diversity in export volumes and indicating the presence of different performance strata among exporting states. Understanding these dynamics is important for developing market strategies and economic policies targeted at each group of exporters.

Based on Figures 3 and 4, it is possible to understand the basis that supported part of the need to evaluate the performance of hierarchical models in the application of coffee export data.

Figure 3 illustrates the evolution of the Kernel density estimates of the logarithm of the value of coffee exports for Brazil's main exporting states, with the layers representing each year in the interval between 1997 and 2022. The color sequence shifted from lighter to darker tones as we moved from the past to the present.

At the start of the period (1997-2005), density was lower and more dispersed, indicating greater variability in export values. From 2006 onwards, there was a gradual increase in density and a concentration of values around a specific range of the log of the value (between log 15 and log 20). This trend suggested a growth in coffee exports and greater stability in values over the years, with 2022 standing out as the year with the highest density in the concentration range.

The pattern observed (Figure 3) provided a clear view of how the coffee export scenario has evolved in Brazil over the last 25 years, moving from a distribution with greater variability to a trend of concentration in higher export values, with notable growth in recent years. This pattern suggests a changing market, potentially impacted by factors such as the economic development of exporting

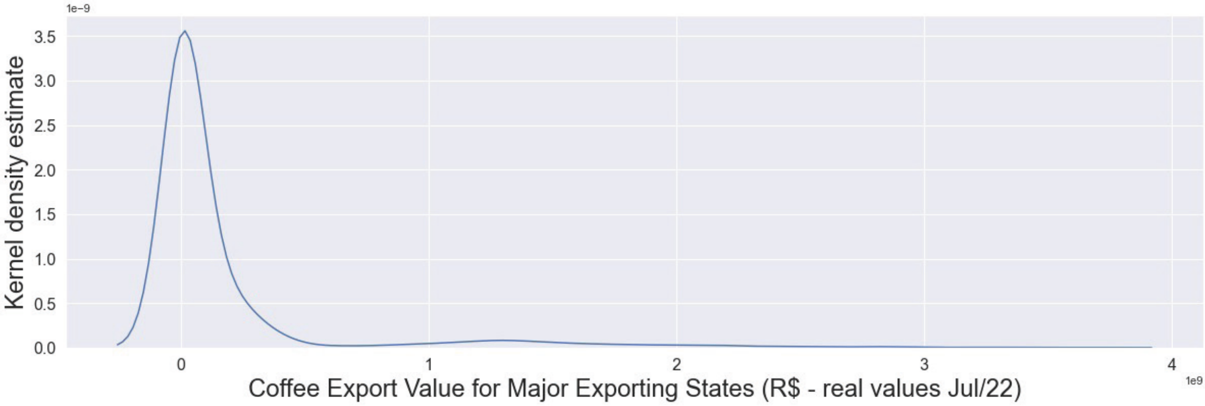

Figure 1 - Kernel density estimate of the value of coffee exports to the main exporting states (R$ - vl. real Jul/22)

Source: Original survey results; Ministry of Economy; Special Secretariat for Foreign Trade and International Affairs (2022)

Note: The values "1e-9" (y-axis) and "1e9" (x-axis) represent the scientific notation scales. The number after the "e" indicates the power of ten by which the number is to be multiplied. Therefore, "1e-9" represents 0.000000001 and "1e9" represents 1,000,000,000

Figure 2: Kernel density estimate of the log of the value of coffee exports to the main exporting states (R$ - vl. reais jul/22)

Source: Original survey results; Ministry of Economy; Special Secretariat for Foreign Trade and International Affairs (2022)

states, changes in global demand and the influence of national and international trade and economic policies (Volsi et. al., 2019).

Figure 4 provides a Kernel density estimate for the logarithm of the value of coffee exports, separated by Federative Unit, based on real values from July 2022. Each state has been represented by a different color, allowing visualization and comparison of the distribution of coffee export values between them.

It can be seen that the width of the curves varied between the states, and states with wider curves indicate a greater dispersion in export values, while narrower curves reflect a stronger concentration of values around a specific range. For example, a state with a wide, flat curve shows a greater diversity of export sizes, ranging from small to large exporters, while a state with a narrow, high peak shows a more homogeneous group of exporters operating in a similar value range.

The analysis presented in Figure 4 allows us to identify which states are the biggest contributors to coffee exports in the country and to understand the variability in exports within each state. It can be seen that Minas Gerais (MG) has a higher concentration of high export values, standing out as the main exporting state. On the other hand, states such as Espírito Santo (ES) and São Paulo (SP) show greater dispersion in values, suggesting greater diversity in the sizes of exporters. This regional differentiation is essential for the development of more targeted economic and trade policies, helping agribusiness decision-makers to devise strategies in line with the specific dynamics of the coffee export market (Moda et al., 2022).

When dealing with a time series used in the context of longitudinal data, it is important to assess the presence of seasonality and trend characteristics so that they can be included in the model. Figure 5 evaluates the historical series of the Value of Coffee Exports for the main exporting states (R$ - vl. reais jul/22) before the logarithmic transformation.

The time series was decomposed together (Figure 6) and this revealed the presence of a certain seasonal pattern and the existence of a trend component - which corroborated the difference in Kernel density estimates by year. The absence of a pattern in the residuals validated the use of a multiplicative decomposition method for this series

The presence of a trend component in the decomposition of the time series points to the impact of the constant search for productivity improvements in coffee plantations, evidenced not only in Brazil (Machado et al., 2024; Ventavoli et al., 2024; Covre et al., 2022), but also in other countries (Wambua et al., 2021; Saikai et al., 2023; Koutouleas et al., 2022). In addition to the trend, another point observed is that of seasonality, linked to what is known as bienniality. As a perennial crop, arabica coffee is significantly affected by bienniality, which has a significant impact on the annual income of farmers and the coffee industry. For this reason, it is essential to develop varieties that do not show bienniality, in addition to applying field management practices (Merga et al., 2023; Garcia et al., 2022).

From this decomposition we moved on to the partial and total autocorrelations in the dependent variable, proposed in order to find inputs for possible predictor variables (Figure 7). Using Total Autocorrelation (ACF) and Partial Autocorrelation (PACF), it was possible to observe that the two most recent periods in the lags had a strong correlation with the most current period, justifying their presence in the modeling prediction tests.

In addition to evaluating the evolution of the dependent variable over time, there was also an evaluation of its relationship with the context of the federal units over time, thus corroborating the differences already presented in the Kernel density estimation for the different contexts (Figure 4). It can be seen that there is a difference, since the state of MG showed a considerable part of the growth trend evolution - similar to its high representativeness - while other states, such as SP and ES, important producers, showed some differences, as did PR, RJ and BA (Figure 8).

Minas Gerais stands out as the largest coffee producer in Brazil, contributing significantly with Arabica varieties, especially in the South and Cerrado regions. The Cerrado Mineiro region, recognized for its denomination of origin, is known for its adoption of advanced technologies and a more mechanized approach to agriculture (Volsi et al., 2019).

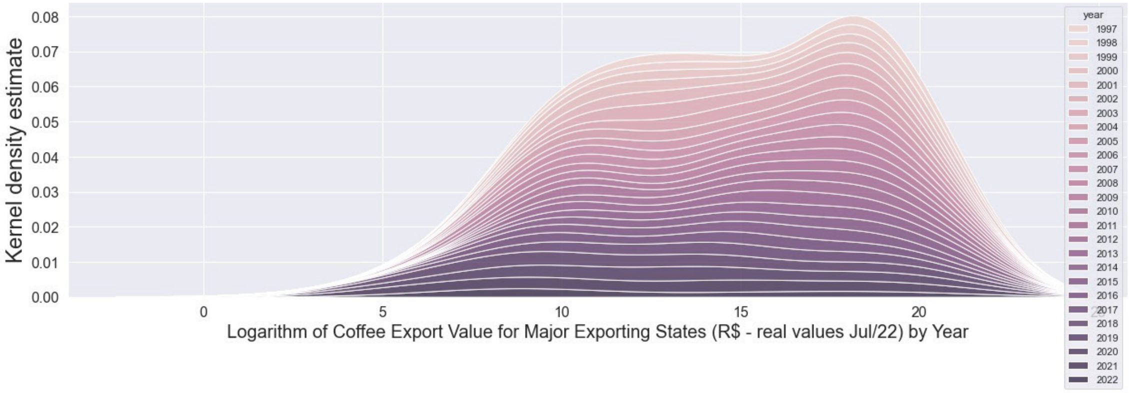

Figure 3 - Kernel density estimate of the log of the value of coffee exports to the main exporting states (R$ - vl. reais jul/22) per year

Source: Original survey results; Ministry of Economy; Special Secretariat for Foreign Trade and International Affairs (2022)

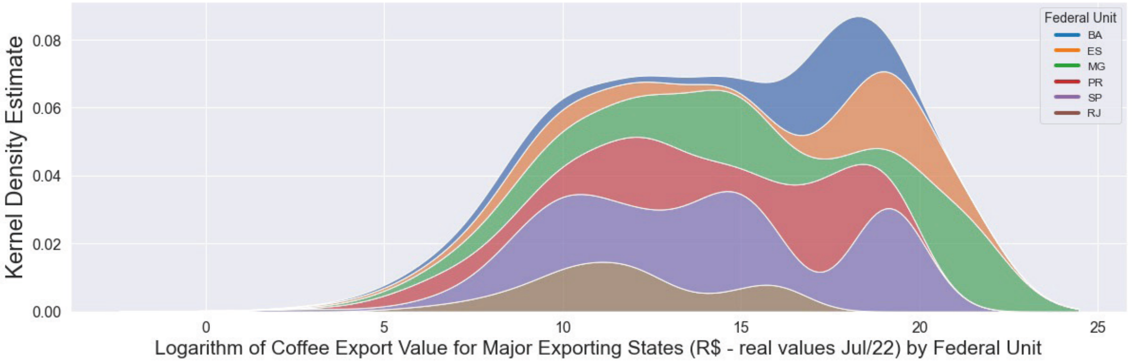

Figure 4 - Kernel density estimate of the log of the value of coffee exports to the main exporting states (R$ - vl. reais jul/22) by Federative Unit

Source: Original survey results / Ministry of Economy / Special Sec. for Foreign Trade and International Affairs (2022)

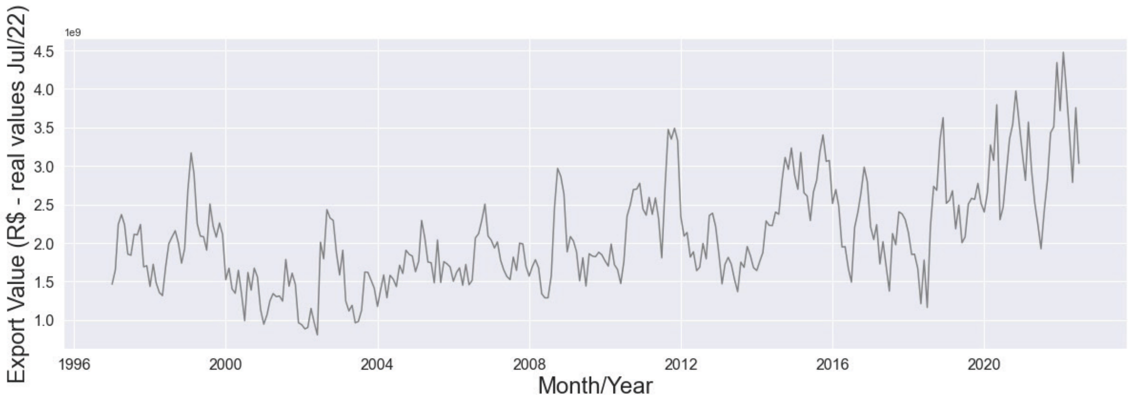

Figure 5 - Evolution of the value of coffee exports to the main exporting states (R$ - real value Jul/22), from January 1997 to May 2022

Source: Original survey results; Ministry of Economy; Special Secretariat for Foreign Trade and International Affairs (2022)

Figure 6. Multiplicative decomposition of the time series for the value of coffee exports to the main exporting states (R$ - vl. real Jul/22)

Source: Original survey results

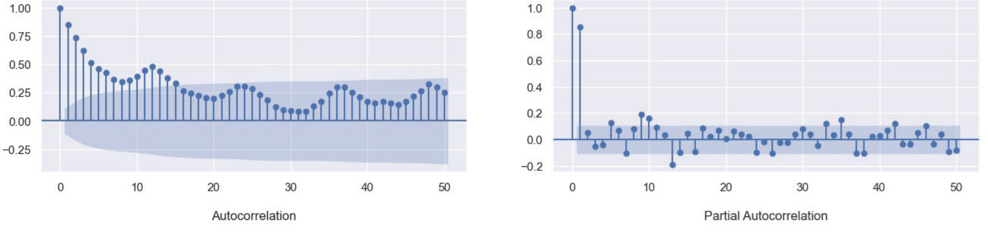

Figure 7. Autocorrelation and partial autocorrelation for the value of coffee exports to the main exporting states (R$ - vl. real Jul/22)

Source: Original survey results

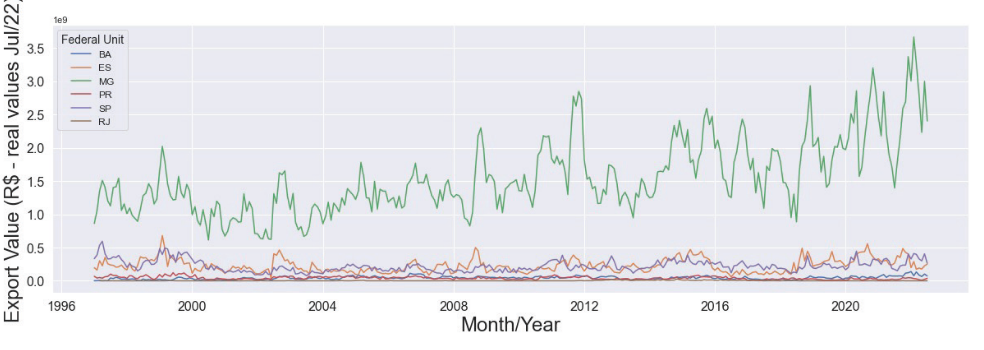

Figure 8. Evolution of the value of coffee exports to the main exporting states (R$ - vl. real jul/22) by federation unit - Jan/1997 to May/2022

Source: Original survey results; Ministry of Economy; Special Secretariat for Foreign Trade and International Affairs (2022)

MODELING RESULTS

Throughout the descriptive analyses, focused on the characteristics of the dependent variable, it was possible to observe its contexts and hypothesize predictor variables for validation. In the initial phase of statistical modeling, attention was paid to an in-depth descriptive analysis of the dependent variable, which consisted of quantifying the value of coffee exports. This preliminary examination was important for identifying the intrinsic characteristics of the data and formulating plausible hypotheses about the predictor variables that could influence export values.

In order to establish a reference point for subsequent analyses, a null model was built. This model does not include any of the potential predictor variables and assumes that the global average of coffee exports can be used to predict the exports of each state. The formalization of the null model in the context of this study is defined as:

$$Y_{ij} = \gamma_{00} + u_{0j} + r_{ij} \tag{4}$$

where, Y_{ij}: represents the value of coffee exports to state j in year i; Y_{00}: is the general intercept; u_{0j}: is the random effect associated with each state, and r_{ij}: is the random error for year i in state j.

This model was compared with an Ordinary Least Squares (OLS) regression model, which assumes the independence of the observations and does not take into account the grouped structure of the data. The likelihood ratio test was applied to compare the suitability of these two models. As shown in Table 2, the difference in LogLikelihood values between the null OLS model and the null HLM is statistically significant (χ^2 = -263825, p < 0.001), indicating that the hierarchical model offers a better fit to the data, even without the inclusion of predictor variables. This result highlights the importance of considering the grouped structure of the data, such as states, when modeling coffee exports in Brazil.

Method	Model	LogLik	Chisq	Pr(>Chisq)
OLS	OLS Null	-1769,5		
HLM 2	HLM 2 Nil	-1180,6	-263825	< 2,2e-16 ***

Table 2. Likelihood Ratio Test for the Null OLS and Null HLM 2 Models

Source: Original research data

Note: Signif,: 0 '***' 0,001 '**' 0,01 '*' 0,05 '.' 0,1 ' ' 1

In order to continue assessing the robustness and predictive power of the models, the explanatory variables presented in the materials and methods were included. The first model evaluated was an OLS using the "stepwise" technique to remove the variables that were not statistically significant, so that only the following remained: the closing value of the dollar per month; the arabica coffee price index in dollars per month; and the iteration variable between year and state.

As a comparison, three empirical strategies were adopted for multilevel modeling with random effects by state: complete model, without removing variables, only with random intercepts; complete model containing in the random effects function also the slope effects for the variables of closing value of the dollar per month, arabica coffee price index in dollars per month and IBCR at the state level; finally, a model was proposed containing the same variables in the random effects of intercept and slope, but keeping only as explanatory variables that would alter or reduce LogLik when removed, considering a "step-up" strategy. The variables that remained in the third model were: closing value of the dollar per month; arabica coffee price indices in dollars; IBCR for the states mentioned; dummies representing the years; and dummies for iterations between states and years. For these models, data from January 2012 to May 2022 was considered, taking into account the availability of all the selected variables.

The comparison between the statistical models showed that the OLS linear regression model, adjusted by the "stepwise" technique, performed better than the hierarchical model with repeated measures for states adjusted by "step-up", as shown by the likelihood ratio test (Table 3). The "stepwise" OLS obtained a LogLik of -663.06, against -701.64 for the "step-up" hierarchical model, with a statistically significant difference ($\chi^2 = 77156$, p = 5.221e-12).

Method	Model	LogLik	Chisq	Pr(>Chisq)
OLS	OLS stepwise	-663,06		
HLM 2	HLM 2 step-up	-701,64	77156	5,221e-12 ***

Table 3. Likelihood Ratio Test for HLM 2 "Step-up" and OLS "stepwise" Models

Source: Original survey results

Note: Signif,: 0 '***' 0,001 '**' 0,01 '*' 0,05 '.' 0,1 ' ' 1

In a more detailed analysis of the hierarchical models, the complete model (HLM 2 Complete), which includes random effects on the slopes of the main variables, showed a statistically significant decrease in LogLik compared to the model with only intercepts and random slopes (HLM 2 with Int. and Incl. Random), as can be seen in Table 4.

Method	Model	LogLik	Chisq	Pr(>Chisq)
HLM 2	HLM 2 with Int, and Incl, Aleat,	-722,56		
HLM 2	HLM 2 Complete	-704,92	35267	5,346e-05 ***

Table 4 - Likelihood Ratio Test for HLM 2 Models with Random Int. and Random Incl. and HLM 2 Complete

Source: Original research data

Note: Signif,: 0 '***' 0,001 '**' 0,01 '*' 0,05 '.' 0,1 ' ' 1

However, when comparing the more complex hierarchical models, the Complete HLM 2 and the Step-up HLM 2, no statistically significant differences were found between them (Table 5).

Method	Model	LogLik	Chisq	Pr(>Chisq)
HLM 2	HLM 2 Complete	-704,92		
HLM 2	HLM 2 Step-up	-701,64	6565	1

Table 5 - Likelihood Ratio Test for HLM 2 Full and HLM 2 Step-up Models

Source: Original research data

Note: Signif,: 0 '***' 0,001 '**' 0,01 '*' 0,05 '.' 0,1 ' ' 1

In view of these analyses, it was decided to proceed with the HLM 2 model by "step-up" for subsequent investigations. This model was preferred because of its more parsimonious use of correlated variables and because it maintains a comparatively low LogLik, implying a reasonable fit without overfitting. Therefore, the OLS "Stepwise" and HLM 2 "Step-up" models were selected for further validation, standing out as the most promising for representing the data in this study.

The analysis of the two selected models, which presented the lowest LogLik values, confirmed findings on the economic factors that influence the coffee export market. The exchange rate of the commercial dollar (US$) at the close of the month (Miranda et al., 2014) and the average domestic price of arabica coffee in dollars (price_arabic_dol) (Monteiro, 2012) were statistically significant in both models (Table 6). These variables highlighted the sensitivity of the value of coffee exports to fluctuations in the foreign exchange market and international coffee prices.

On the other hand, IBCR did not reach statistical significance (Carrasco-Gutierrez and Almeida, 2013). However, excluding this variable from the models significantly reduced the explanatory power, as evidenced by the changes in the LogLik values. This indicated that, despite its lack of statistical significance, the IBCR contributed to understanding the variability in coffee exports, possibly reflecting regional economic dynamics that were not captured by other variables.

The dummy variables representing the years (Tanahashi and Caldarelli, 2021) proved to be significant in the HLM 2 model adjusted by step-up, with the exception of 2015 and 2022. This observation suggested that specific annual variations may have influenced coffee exports, highlighting the need to investigate economic events or conditions that impacted the market in those specific years.

Due to the number of interactions present, the combinations between the years 2012 and 2022 and the states of ES, BA, RJ, PR, MG and SP are detailed in the Appendix, Table 8. In the OLS model, the interactions for Minas Gerais showed mostly positive and statistically significant effects, while Paraná and Rio de Janeiro showed predominantly negative effects. In the repeated measures model, the effects of PR and RJ over time were also statistically significant.

Some interpretations can be drawn from the data available for the random effects. The intraclass correlation coefficient (ICC), with a value of 1, indicates a high explanatory power of the random effects, representing the influence of state contexts on the inclinations of the variables and in explaining the clustering in the observations. In addition, the "Conditional R^2" measures the share of variance explained by the fixed and random effects together, while the "Marginal R^2" reflects the low explanatory power attributed only to the fixed effects, highlighting the importance of the random components in the modeling.

Despite the model's high ICC value, the adjusted ICC, equal to 1, does not represent the variance explained by the hierarchical context in the case of non-null models (Rabe-Hesketh and Skrondal, 2012). To this end, it was necessary to calculate the conditional ICC in order to obtain a more credible value for the intraclass correlation coefficient, which also takes into account the explanatory power of the fixed effects variables. In the model in question (Table 6), the conditional ICC showed a value of 0.527, which although lower than the adjusted ICC, demonstrated the importance of the explanation of the variance of the result by the context of the states alone, regardless of the fixed effects applied in the model.

The results of the estimated parameters for the random intercepts and the slopes of the variables included in the context of the states, shown in Figures 10 to 13 in the Appendix, showed similar behavior among the states, with the exception of Rio de Janeiro (RJ). The random intercepts represented the specific effects of each state, which were not captured by the fixed variables in the model. All the states, except Rio de Janeiro, had negative intercepts, which suggested that, keeping all the other variables constant, these states tended to have lower logarithmic values of coffee exports than the general average. On the other hand, after adjusting for the model's explanatory variables, Rio de Janeiro had a significant positive random intercept, indicating that its logarithmic coffee export values were higher than the overall average.

Although the positive random intercept for RJ suggested that its logarithmic coffee exports were higher than average, this did not mean that the state had the highest total export volumes. The model only indicated that, after adjusting for the explanatory variables, RJ stood out in the modeled context. It is important to note that the random intercept does not directly measure the total volume of exports, but rather the specific differences of each state in relation to the central trend captured by the fixed variables.

However, the results observed for RJ indicated that there may be unique factors - such as market conditions, regional policies, or marketing strategies - that the model did not fully capture. The presence of this significant positive intercept signaled that the model lacked additional variables to fully explain the state's coffee exports.

Thus, although random intercepts are not necessarily "good" or "bad", they signal that there is unexplained variability that merits further investigation. In the case of Rio de Janeiro, this may indicate regional characteristics or specific factors that the model has not adequately captured. Therefore, it would be desirable to consider including additional variables capable of capturing such singularities, ensuring that the model more accurately reflects regional nuances, especially in states like Rio de Janeiro.

Figure 9 compares the fit of the HLM 2 and OLS models, focusing on the predicted values *versus* the actual values of coffee exports. Specifically, for exports with lower logarithmic values, which correspond to lower volumes or monetary values of exports, there was a greater discrepancy between the predictions of the models and the actual data, particularly in the OLS model. For higher export values, the models tended to present more aligned predictions. This analysis suggested that the HLM 2 model may have a superior performance in capturing the variability in exports for lower export values, highlighting the advantage of hierarchical models in managing data with complex structures.

In order to assess the predictive power of the models, the following error metrics were calculated: Mean Absolute Percentage Error (MAPE), Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) (Table 7).

Error metrics	HLM 2 (Step-up)	OLS
MAPE	2,20%	2,30%
MAE	0,32	0,33
RMSE	0,565	0,586

Table 7. Metrics used to evaluate the forecasts made by the models

Source: Original survey results

Note: MAPE: ......; RMSE:.....

As can be seen from the fit graph (Figure 9) and the evaluation metrics (Table 7), the OLS and HLM models showed very close results.

The differences between the models were also relatively small when the forecasts were made at state level (Figure 10)

In the ASMs for forecasts by month/year, it was possible to see that the highest error rate was between the years 2018 (the start of the increase) and 2021, with a significant reduction in 2022 (Figure 11). The results were similar for the other metrics.

The analysis of the results, in comparison with the references on models for coffee exports, revealed a consensus regarding the impact of the exchange rate on export forecasts. Both in studies on national exports (Tanahashi and Caldarelli, 2021) and in the literature focused on state-specific forecasts, such as Espírito Santo (Monteiro, 2012), the exchange rate showed positive short-term effects on export results. However, when comparing the models, there was a difference in the importance attributed to this variable. As in the study on Espírito Santo (Monteiro, 2012), the model proposed in this paper identified the exchange rate as the main driver of coffee exports, even considering real prices adjusted by the IGP-M deflator and converted into reais (R$).

In the model proposed for Brazil by Tanahashi and Caldarelli (2021), Gross Domestic Product (GDP) was identified as the variable with the greatest impact on exports, highlighting the importance of the domestic market, while the exchange rate was less relevant. In this study, the Regional Economic Activity Index (IBCR) was used as a proxy for GDP in the HLM 2 model applied to the states. Although the IBCR was not statistically significant, its positive coefficient and inclusion in the second level model helped to keep LogLik at a lower level, showing a good predictive fit. In the Ordinary Least Squares (OLS) model, which used the stepwise technique, the IBCR was excluded because it was not statistically significant.

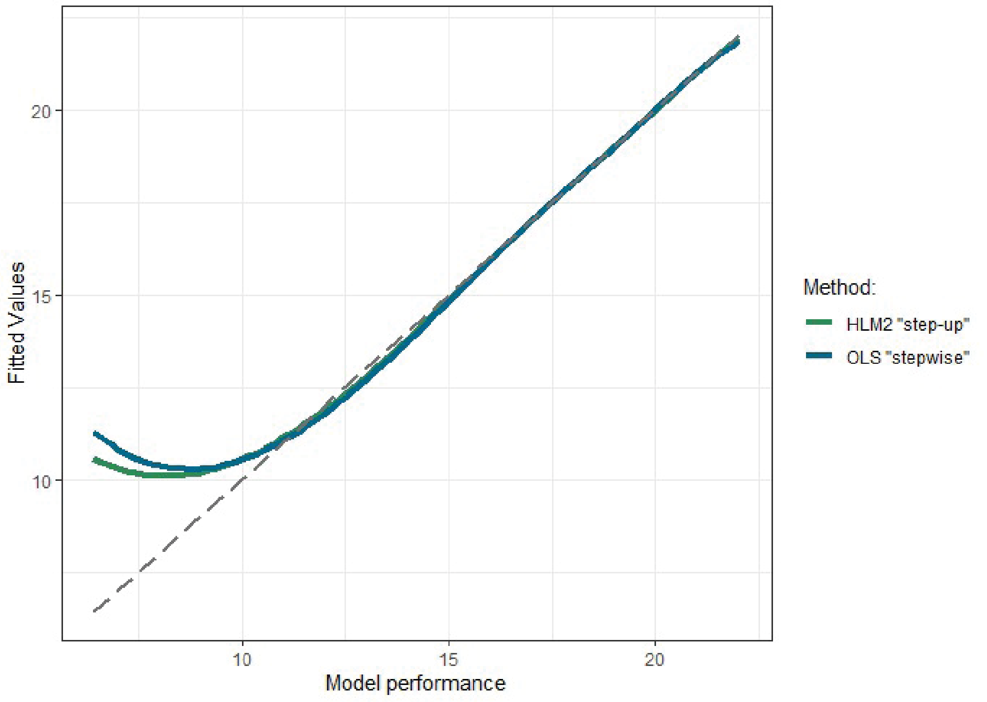

Figure 9. Comparison between actual and predicted values for HLM 2 model with step-up variable selection and OLS with stepwise variable selection.

Source: Original survey results

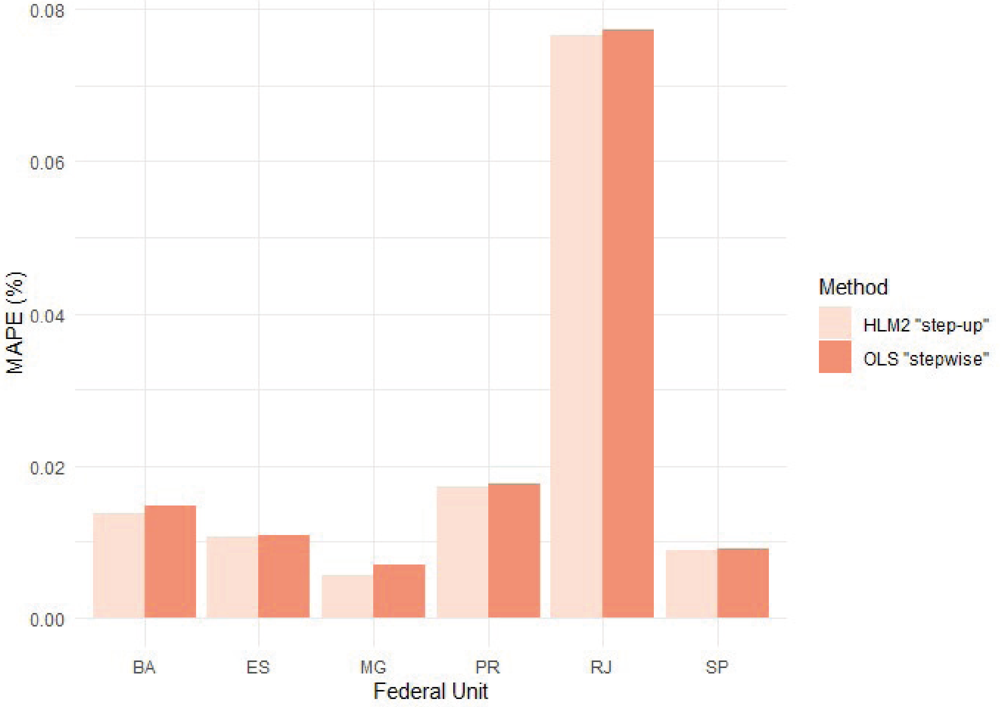

Figure 10. Comparison of MAPE for HLM 2 model with step-up variable selection and OLS with stepwise variable selection by state

Source: Original survey results

Journal of Agricultural Sciences Research ISSN 2764-0973 DOI https://doi.org/10.22533/at.ed.973512507019

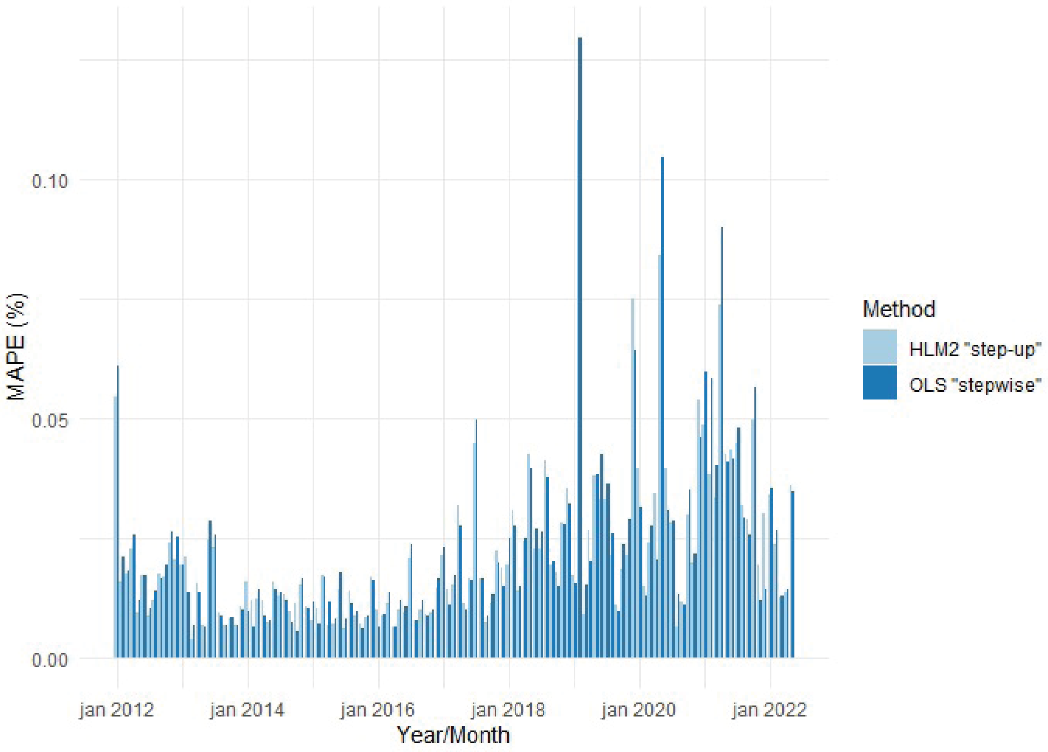

Figure 11. Comparison of MAPE for HLM 2 model with “Step-up” variable selection and OLS with “Stepwise” variable selection by Year and Month
Source: Original survey results

	HLM 2 Step-up		Stepwise OLS	
Predictors	Estimates	p	Estimates	p
us	1,81	<0,001	1,90	<0.001
price_arabic_dol	0,78	0,002	0,88	<0.001
ibcr	1,27	0,526		
Random Effects				
σ^2	0,35			
t00	0.33 state			
t11	0.82 state, us			
		0,15 state, price_arabic_dol		
		23.05 state, ibcr		
ρ_{01}	0,97			
	0,97			
	0,97			
ICC	1,00			
N	6 status			
Observations	749	749		
Marginal R^2 / Conditional R^2	0,473 / 1,000	0,948 / 0,943		
AIC	1.563.279	1.464.123		
log-likelihood	-701.640	-663.062		

Table 6. Summary - Dependent Variable: Log - Value of Coffee Exports (R$)
Source: Original survey results

The literature also discusses variables related to domestic coffee prices. In previous studies, both at national and state level, these prices showed statistically significant and negative effects on export values. In this study, the variable was analyzed for the two types of coffee exported in Brazil: robusta and arabica. While robusta coffee prices were not statistically significant, arabica coffee prices showed statistical significance with a positive coefficient, indicating a positive impact on exports. This result can be explained by the higher quality and price of Arabica coffee, possibly associated with a reduction in domestic consumption. Finally, other variables frequently used in the literature as “proxies” for seasonality (Tanahashi and Caldarelli, 2021) were also not statistically significant, corroborating the results obtained.

As its main contribution to the literature, the proposed model not only used longitudinal data, increasing the available sample, but also provided a more detailed understanding of the dynamics of coffee exports at a regional level. The use of a multi-level structure allowed the model to capture the specific effects of each producing and exporting state, while respecting their heterogeneities. This showed that coffee exports in Brazil are not homogeneous and are strongly influenced by specific regional factors such as market conditions, local policies and commercial strategies. The model revealed that the states had different inclinations over time, showing that the interactions between state and time, in the modeled hierarchy, are crucial to understanding how exports evolve.

From the analysis, it can be said that while the aggregate scenario of Brazil as an exporter is important, granular analysis by state is essential for more accurate decision-making. The flexibility offered by the random effects and the different regional inclinations highlighted the importance of targeted policies for each state, rather than a single approach for the whole country. Thus, the model not only contributes to more robust forecasts, but also to the formulation of export strategies that take regional peculiarities into account.

CONCLUSION

This study proposed a hierarchical model (HLM 2) to predict Brazilian coffee exports, taking into account both the national context and the particularities of state production and exports. The model proved to be effective in capturing the inherent differences between states, allowing independent identification of their results by estimating fixed and random effects. In addition, an Ordinary Least Squares (OLS) model was used for comparison, incorporating fixed effects through dummy variables and interactions.

Data from January 2012 to May 2022 was analyzed, covering the main coffee exporting states: Minas Gerais, Espírito Santo, São Paulo, Bahia and Paraná. In line with the literature, the exchange rate had a positive and statistically significant impact, reaffirming its decisive role in exports. However, the domestic prices of the Robusta and Arabica varieties produced different results. The price of Robusta coffee was not statistically significant, reflecting its smaller share of the international market. On the other hand, the price of Arabica coffee had a positive and significant impact, contrary to some expectations in the literature.

These findings suggest that the Arabica coffee export market may be influenced by emerging factors, such as a growing niche market or changes in global preferences for quality, which value Brazilian Arabica coffee. This highlights the need for further analysis of the dynamics of this market and strategic adjustments by exporters to exploit these opportunities. Although Robusta coffee has not shown a significant impact, it is recommended to monitor international consumption trends and prices in order to anticipate possible changes in demand.

The variable "state production", used as a "proxy", was not statistically significant in the HLM 2 and OLS models. However, its inclusion in the hierarchical model contributed to an improvement in predictive capacity, showing positive impacts as expected. Thus, the models proved to be flexible tools, suitable for both national and regional analysis, offering insights into local variations in coffee exports.

The inclusion of local characteristics, such as type of coffee and climate, can further enrich the analysis, expanding the model's explanatory capacity. With the use of panel data and a comprehensive history by state, the proposed model is applicable to sub-national entities, regional councils and producers looking for forecasts adjusted to regional specificities. As an extension of this study, we recommend incorporating climate variables, given their crucial impact on coffee production and exports.

REFERENCES

Armstrong, R.A.; Hilton, A.C. 2010. Stepwise Multiple Regression p.135-138. In: Armstrong, R.A.; Hilton, A.C. Statistical Analysis in Microbiology: Statnotes. Wiley-Blackwell.

Associação Brasileira da Indústria de Café [ABIC]. Sumário Executivo do Café, 2021. Disponível em: https://estatisticas.abic.com.br/wp-content/uploads/2021/12/2021.12.SumarioCafe.pdf. Acesso em: 14 abr. 2022.

B3. Market Data. 2022. Disponível em: https://www.b3.com.br/pt_br/market-data-e-indices/servicos-de-dados/market-data/cotacoes/. Acesso em: 14 abr. 2022.

Banco Central do Brasil [Bacen]. 2022. Sistema Gerenciador de Séries Temporais. Disponível em: https://www3.bcb.gov.br/sgspub/localizarseries/localizarSeries.do?method=prepararTelaLocalizarSeries. Acesso em: 17 de abr. 2022.

Campos, S.A.C. 2022. Análise da competitividade do setor cafeeiro brasileiro no mercado internacional. Economia & Região, 10(1): 123-143.

Carrasco-Gutierrez, C.E.; Almeida, F.M.M. 2013. Modelagem e previsão do preço do café brasileiro. Revista de Economia, 39(2). Disponível em: https://revistas.ufpr.br/economia/article/viewFile/23476/22506. Acesso em: 13 abr. 2022.

Courgeau, D. 2003. Methodology and epistemology of multilevel analysis. 1ed. Kluwer Academic Publishers, London, England. https://doi.org/10.1007/978-1-4020-4675-9.

Covre, A.M.; Oliveira, M.G.; Martins, L.D.; Bonomo, R.; Rodrigues, W.N.; Tomaz, M.A.; Vieira, H.D.; Paye, H.S.; Partelli, F.L. 2022. Como é modulado o desenvolvimento frutífero de Coffea canephora pelo abastecimento de água? Uma análise das curvas de crescimento para sistemas irrigados e não irrigado. Semina: Ciências Agrárias, 43(5): 2359-2374. https://doi.org/10.5433/1679-0359.2022v43n5p2359.

Paula, M.C.; Leão, C.; Santos, L.L. 2024. Café além-fronteiras: influência dos preços internacionais nos preços domésticos brasileiros no período de 2013 a 2022. Cuadernos de Educación y Desarrollo, 16(3). https://doi.org/10.55905/cuadv16n3-068.

Estatísticas de Comércio Exterior do Brasil [Comex Stat]. Comex Vis, 2022. Disponível em: http://comexstat.mdic.gov.br/pt/comex-vis. Acesso em: 15 abr. 2022.

Fávero, L.P.; Belfiore, P.; Souza, R.F. 2023. Data science, analytics and machine learning with R. Academic Press.

Fawcett, T.; Provost, F. 2016. Data Science para Negócios: O que você precisa saber sobre mineração de dados e pensamento analítico de dados. 1ed. Alta Books Editora, Rio de Janeiro, RJ, Brasil.

Garcia, G.M.; Orians, C.M. 2022. Reproductive tradeoffs in a perennial crop: Exploring the mechanisms of coffee alternate bearing in relation to farm management. Agriculture, Ecosystems & Environment, 340, 108151. https://doi.org/10.1016/j.agee.2022.108151.

Goldstein, H. 2011. Multilevel statistical models. 4ed. John Wiley & Sons, Chichester, West Sussex, England. DOI:10.1002/9780470973394.

Haack, D.M.P.; Ferreira, K.R.M; Gonçalves, L.B.; Azeredo, M.F.; Costa, S.R.R.; Santos, E.N.S.; Prell, B.P.; Nunes, L.B. 2023. Gestão da produção brasileira de café: uma questão de segurança alimentar em um cenário marcado pelas alterações do clima. GeSec: Revista de Gestão e Secretariado, 14(8): 13221-13235. https://doi.org/10.7769/gesec.v14i8.2597.

Koutouleas, A.; Sarzynski, T.; Bordeaux, M.; Bosselmann, A.S.; Campa, C.; Etienne, H.; Turreira-García, N.; Rigal, C.; Vaast, P.; Ramalho J.C.; Marraccini, P.; Ræbild, A. 2022. Shaded-coffee: A nature-based strategy for coffee production under climate change? A review. Frontiers in Sustainable Food Systems, 6, 877476. https://doi.org/10.3389/fsufs.2022.877476.

Machado, A.C.; Baronio, G.J.; Novaes, C.S.; Ollerton, J.; Torres, M. W.; Lopes, D.N.S.; Rech, A.R. 2024. Optimizing coffee production: Increases in floral visitation and bean quality at plantation edges with wild pollinators and natural vegetation. Journal of Applied Ecology. 61(3): 465-475. https://doi.org/10.1111/1365-2664.14591.

Merga, D.; Beksisa, L.; Merga, W.; Alemayehu, D. 2023. Genetic variability of coffee (*Coffea arabica* L.) germplasm in biennial bearing and its influence on selection efficiency. International Journal of Agronomy. https://doi.org/10.1155/2023/9995531.

Miranda, A.P.; Coronel, D.A.; Vieira, K.M. 2014. Previsão do mercado futuro do café arábica utilizando redes neurais e métodos econométricos. Estudos do CEPE, 38: 66-98. https://doi.org/10.17058/cepe.v0i0.3329.

Moda, L.R.; Spers, E.E.; Almeida, L.F; Alencar Schiavi, S.M. 2022. Brazilian Coffee Sustainability, Production, and Certification. In Sustainable Agricultural Value Chain. IntechOpen. DOI: 10.5772/intechopen.105135.

Monteiro, E. 2012. Exportações de café do Espírito Santo: aplicação da metodologia VAR. Revista de Política Agrícola, 21(4): 106-119.

Saikai, K.K.; Oduori, C.; Situma, E.; Njoroge, S.; Murunde, R.; Kimenju, J.W.; Miano, D.W.; Coyne, D. 2023. Biocontrol-based strategies for improving soil health and managing plant-parasitic nematodes in coffee production. Frontiers in Plant Science, 14, 1196171.

Raudenbush, S.W.; Bryk, A.S. 2002. Hierarchical linear models: applications and data analysis methods. 2ed. Sage Publications, Thousand Oaks, CA, USA.

Rabe-Hesketh, S; Skrondal, A. 2012. Multilevel and Longitudinal Modeling Using Stata. 3ed. StataCorp.

Tanahashi, A.; Canderelli, C. 2021. Determinantes da oferta de exportação brasileira de café. Revista de Política Agrícola. 30(1): 98-116.

United States Department of Agriculture [USDA]. Production Supply and Distribution Reports. Disponível em: https://apps.fas.usda.gov/psdonline/app/index.html#/app/home. Acesso em: 13 abr. 2022.

Ventavoli, D.M.A.; Rezende, D.C.; Moraes, C.S.B.; Jubrail, M.P. 2024. Sustainability and Quality: Analysis of the Benefits of Bee Pollination in Coffee. Revista de Gestão Social e Ambiental, 18(2): e04504. https://doi.org/10.24857/rgsa.v18n2-036.

Volsi, B.; Telles, T.S.; Caldarelli, C.E.; Camara, M.R.G.D. 2019. The dynamics of coffee production in Brazil. PloS one, 14(7), e0219742. https://doi.org/10.1371/journal.pone.0219742

Wambua, D.M.; Gichimu, B.M.; Ndirangu, S.N. 2021. Smallholder coffee productivity as affected by socioeconomic factors and technology adoption. International Journal of Agronomy, 2021, 1-8. https://doi.org/10.1155/2021/8852371

APPENDIX

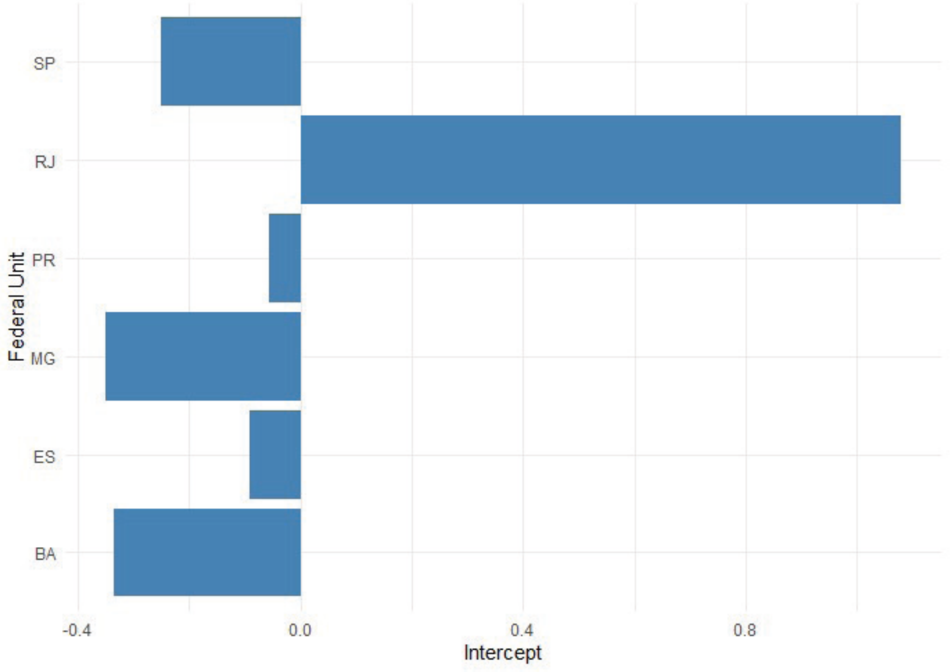

Figure 10. Random intercepts by state for HLM 2 model with step-up variable selection

Source: Own elaboration

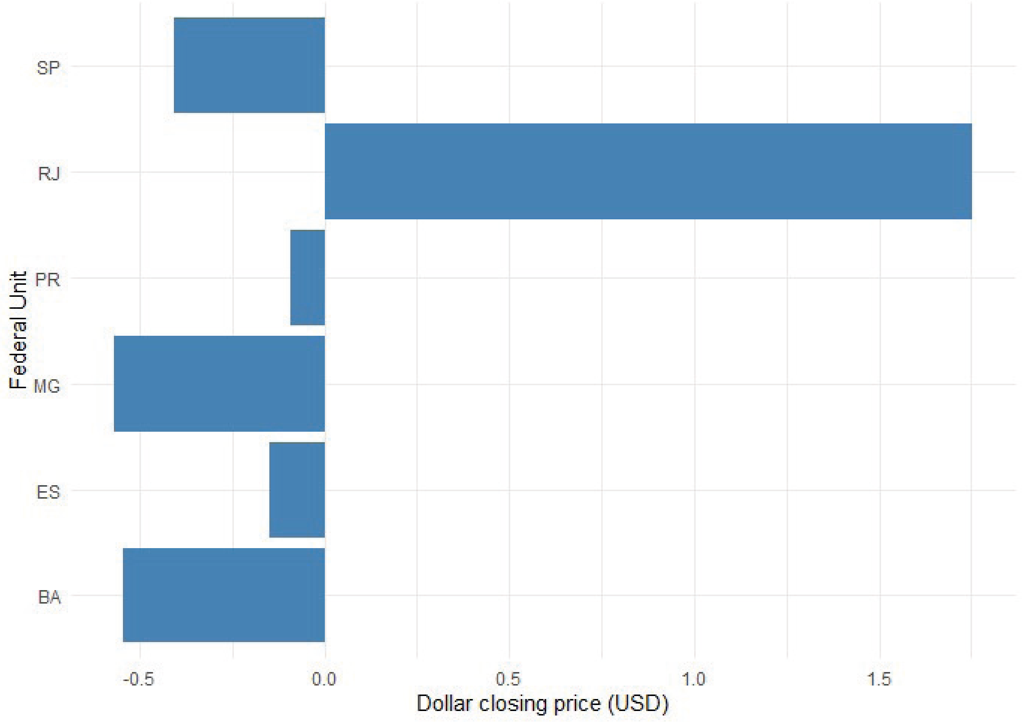

Figure 11. Random effects by state for dollar closing value variable for HLM 2 model with step-up variable selection

Source: Original survey results

Journal of Agricultural Sciences Research ISSN 2764-0973 DOI https://doi.org/10.22533/at.ed.973512507019

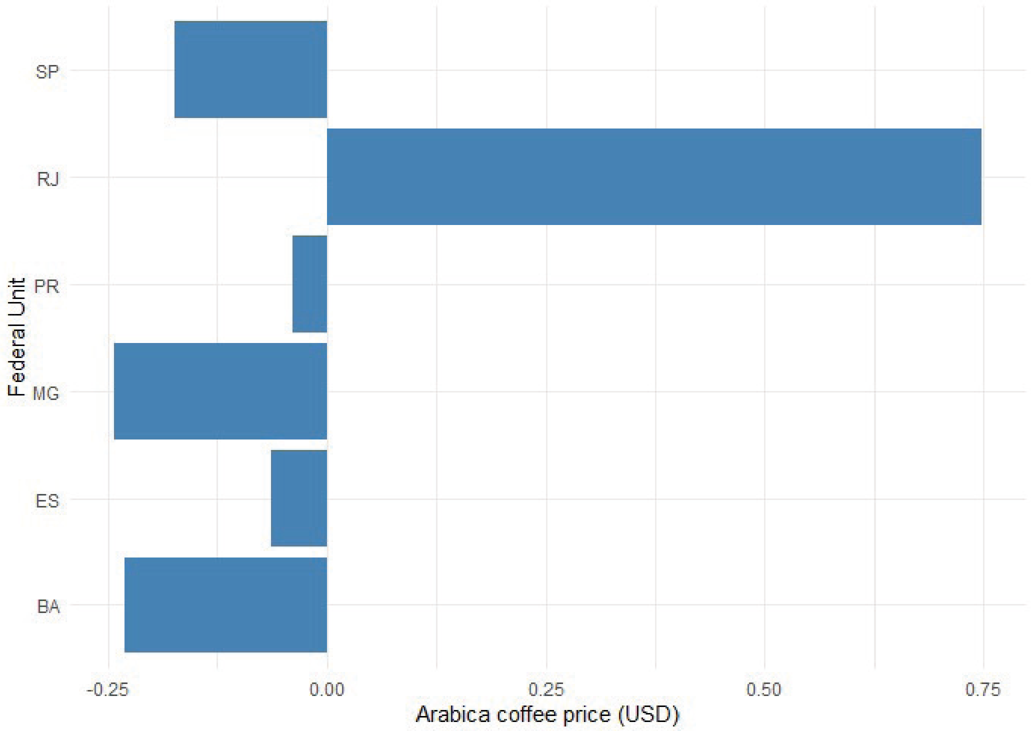

Figure 12. Random effects by state for arabica coffee price index variable in dollars for HLM 2 model with step-up variable selection.

Source: Original survey results

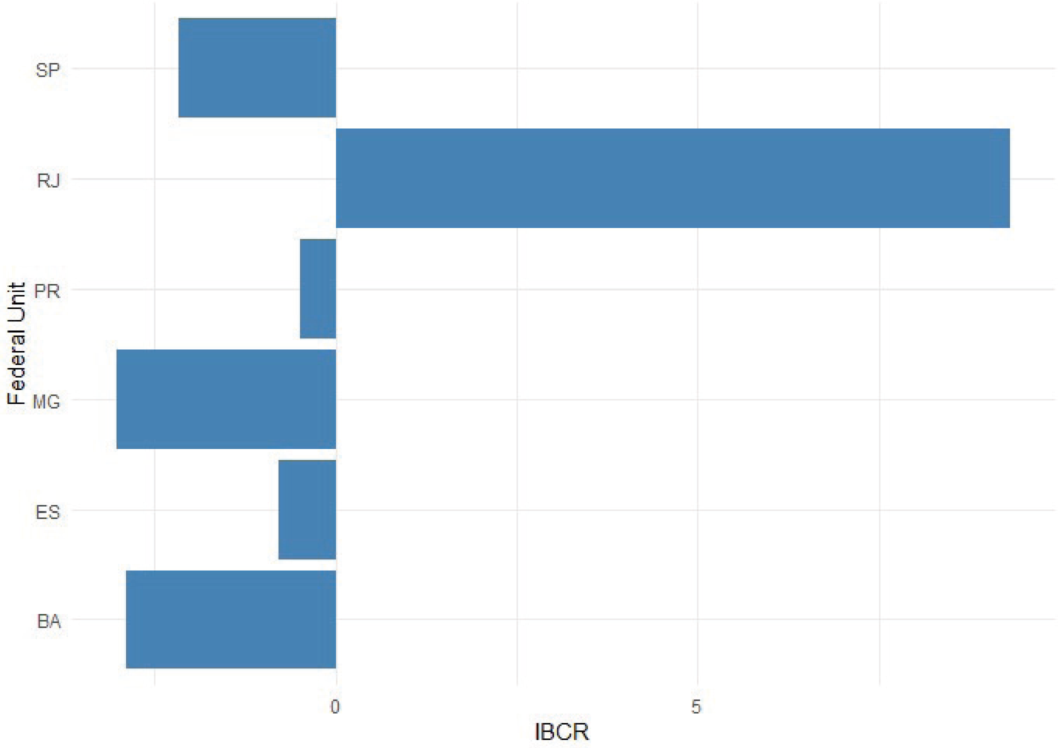

Figure 13. Random effects by state for IBCR variable for HLM 2 model with step-up variable selection

Source: Original survey results

Journal of Agricultural Sciences Research ISSN 2764-0973 DOI https://doi.org/10.22533/at.ed.973512507019

	HLM 2 Step-up		Stepwise OLS	
Predictors	Estimates	p	Estimates	p
(Intercept)	22.36	<0.001	11.74	<0.001
us	1.81	<0.001	1.90	<0.001
price_arabic_dol	0.78	0.002	0.88	<0.001
ibcr	1.27	0.526		
year [2013]	-0.82	0.001		
year [2014]	-0.58	0.024		
year [2015]	-0.43	0.160		
year [2016]	-1.17	<0.001		
year [2017]	-1.40	<0.001		
year [2018]	-1.46	<0.001		
year [2019]	-0.76	0.022		
year [2020]	-1.06	0.008		
year [2021]	-0.93	0.041		
year [2022]	-0.77	0.141		
year [2012] * factor(state)ES	-10.51	0.102	1.50	0.003
year [2013] * factor(state)ES	-9.76	0.129	1.47	0.005
year [2014] * factor(state)ES	-9.87	0.129	1.55	0.001
year [2015] * factor(state)ES	-10.20	0.119	1.33	0.001
year [2016] * factor(state)ES	-10.51	0.105	0.23	0.560
year [2017] * factor(state)ES	-10.28	0.112	0.24	0.568
year [2018] * factor(state)ES	-9.75	0.134	0.73	0.076
year [2019] * factor(state)ES	-10.13	0.119	1.02	0.012
year [2020] * factor(state)ES	-10.05	0.122	0.72	0.054
year [2021] * factor(state)ES	-10.69	0.105	0.15	0.646
year [2022] * factor(state)ES	-11.52	0.084	-0.55	0.158
year [2012] * factor(state)MG	4.09	0.561	3.46	<0.001
year [2013] * factor(state)MG	4.87	0.490	3.46	<0.001
year [2014] * factor(state)MG	4.67	0.511	3.37	<0.001
year [2015] * factor(state)MG	4.33	0.543	3.11	<0.001
year [2016] * factor(state)MG	4.83	0.495	2.90	<0.001
year [2017] * factor(state)MG	5.05	0.475	2.91	<0.001

year [2018] * factor(state)MG	5.03	0.479	2.80	<0.001
year [2019] * factor(state)MG	4.45	0.531	2.91	<0.001
year [2020] * factor(state)MG	4.60	0.519	2.62	<0.001
year [2021] * factor(state)MG	4.20	0.562	2.07	<0.001
year [2022] * factor(state)MG	4.22	0.563	2.13	<0.001
year [2012] * factor(state)PR	-14.10	0.027	-0.39	0.432
year [2013] * factor(state)PR	-13.44	0.035	-0.49	0.343
year [2014] * factor(state)PR	-13.67	0.034	-0.54	0.241
year [2015] * factor(state)PR	-13.68	0.034	-0.45	0.269
year [2016] * factor(state)PR	-13.26	0.039	-0.80	0.047
year [2017] * factor(state)PR	-14.25	0.026	-1.99	<0.001
year [2018] * factor(state)PR	-13.43	0.037	-1.22	0.003
year [2019] * factor(state)PR	-14.83	0.022	-1.90	<0.001
year [2020] * factor(state)PR	-14.54	0.025	-1.96	<0.001
year [2021] * factor(state)PR	-15.21	0.021	-2.54	<0.001
year [2022] * factor(state)PR	-15.50	0.020	-2.70	<0.001
year [2012] * factor(state)RJ	-70.00	<0.001	-2.34	<0.001
year [2013] * factor(state)RJ	-68.35	<0.001	-1.49	0.004
year [2014] * factor(state)RJ	-69.23	<0.001	-1.71	<0.001
year [2015] * factor(state)RJ	-70.21	<0.001	-2.32	<0.001
year [2016] * factor(state)RJ	-69.42	<0.001	-2.62	<0.001
year [2016] * factor(state)RJ	-69.42	<0.001	-2.62	<0.001
year [2017] * factor(state)RJ	-70.17	<0.001	-3.96	<0.001
year [2018] * factor(state)RJ	-70.91	<0.001	-4.67	<0.001
year [2019] * factor(state)RJ	-74.43	<0.001	-7.26	<0.001

year [2020] * factor(state)RJ	-74.02	<0.001	-7.06	<0.001
year [2021] * factor(state)RJ	-76.92	<0.001	-9.07	<0.001
year [2022] * factor(state)RJ	-78.54	<0.001	-10.22	<0.001
year [2012] * factor(state)SP	-2.64	0.724	1.55	0.002
year [2013] * factor(state)SP	-1.92	0.799	1.48	0.004
year [2014] * factor(state)SP	-2.04	0.787	1.49	0.001
year [2015] * factor(state)SP	-2.50	0.741	1.11	0.007
year [2016] * factor(state)SP	-2.07	0.783	0.83	0.040
year [2017] * factor(state)SP	-1.82	0.809	0.86	0.039
year [2018] * factor(state)SP	-1.98	0.794	0.61	0.138
year [2019] * factor(state)SP	-2.52	0.740	0.73	0.073
year [2020] * factor(state)SP	-2.40	0.752	0.45	0.231
year [2021] * factor(state)SP	-2.91	0.706	-0.12	0.727
year [2022] * factor(state)SP	-2.88	0.711		
year2012 *			0.19	0.703
year [2013] * factor(state)SP	-1.92	0.799	1.48	0.004
year [2014] * factor(state)SP	-2.04	0.787	1.49	0.001
year [2015] * factor(state)SP	-2.50	0.741	1.11	0.007
year [2016] * factor(state)SP	-2.07	0.783	0.83	0.040
year [2017] * factor(state)SP	-1.82	0.809	0.86	0.039
year [2018] * factor(state)SP	-1.98	0.794	0.61	0.138
year [2019] * factor(state)SP	-2.52	0.740	0.73	0.073
year [2020] * factor(state)SP	-2.40	0.752	0.45	0.231
year [2021] * factor(state)SP	-2.91	0.706	-0.12	0.727
year [2022] * factor(state)SP	-2.88	0.711		
year2012 *			0.19	0.703

factor(state)BA			
year2013 * factor(state)BA		-0.61	0.238
year2014 * factor(state)BA		-0.55	0.237
year2015 * factor(state)BA		-0.51	0.218
year2016 * factor(state)BA		-1.18	0.003
year2017 * factor(state)BA		-1.38	0.001
year2018 * factor(state)BA		-1.49	<0.001
year2019 * factor(state)BA		-0.81	0.048
year2020 * factor(state)BA		-1.20	0.001
year2021 * factor(state)BA		-1.29	<0.001
year2022 * factor(state)BA		-1.30	0.001
Random Effects			
σ2	0.35		
τ00	0.33 status		
τ11	0.82 state.us		
	0.15 state.price_ arabic_dol		
	23.05 state.ibcr		
ρ01	0.97		
	0.97		
	0.97		
ICC	1.00		
N	6 status		
Observations	749	749	
Marginal R2 / Conditional R2	0.473 / 1.000	0.948 / 0.943	
AIC	1.563.279	1.464.123	
log-likelihood	-701.640	-663.062	

Table 8. Complete Table - Dependent Variable: Log - Value of Coffee Exports (R$)

Source: Original survey results