Universidad Pontificia Comillas

Facultad de Ciencias Económicas y Empresariales (ICADE)

COFFEE AS A COMMODITY: MODELLING THE COFFEE MARKET

A Structural Vector Autoregressive Analysis of the Real Price of Coffee

Author: Manuel de Vega Moreno

Director: Elena María Díaz Aguiluz

Madrid | June 2026

Abstract

This thesis develops a structural Vector Autoregressive (VAR) model of the global coffee market in order to quantify the dynamic interactions between the real price of Arabica coffee and nine fundamental drivers identified in de Vega Moreno (2026): rainfall in producing countries, Brazilian temperature, the Global Supply Chain Pressure Index, the World Industrial Production Index, energy prices, fertiliser prices, the average minimum wage of producing countries, and the bilateral exchange rates of Brazil and Colombia against the US dollar. The model is estimated on a balanced monthly panel of 300 observations covering January 1999 to December 2023 and is identified through a recursive Cholesky decomposition whose ordering runs from the most exogenous drivers (climate) through global activity and input costs to the real coffee price and, finally, to a domestic block of country-specific exchange rates and wages. The empirical results show that the own price shock dominates the variance of the coffee price, falling from about 97% at the impact horizon to roughly 63% at twenty-four months, while the nine fundamentals jointly account for around 37% of the long-run forecast error variance. Global industrial activity, rainfall, minimum wages and supply-chain pressure emerge as the most influential drivers, a ranking that is consistent with the descriptive evidence of de Vega Moreno (2026). The work illustrates how a descriptive framework can be enriched with a structural econometric model that quantifies how much, how fast, and through which channels fundamentals transmit to coffee prices.

Keywords: *coffee; commodity prices; vector autoregression; Cholesky identification; impulse response functions; forecast error variance decomposition; supply-chain pressure.*

Table of Contents

Declaración de Uso de Herramientas de Inteligencia Artificial Generativa en Trabajos Fin de Grado

ADVERTENCIA: Desde la Universidad consideramos que ChatGPT u otras herramientas similares son herramientas muy útiles en la vida académica, aunque su uso queda siempre bajo la responsabilidad del alumno, puesto que las respuestas que proporciona pueden no ser veraces. En este sentido, NO está permitido su uso en la elaboración del Trabajo fin de Grado para generar código porque estas herramientas no son fiables en esa tarea. Aunque el código funcione, no hay garantías de que metodológicamente sea correcto, y es altamente probable que no lo sea.

Por la presente, yo, Manuel de Vega Moreno, estudiante del Doble Grado en Administración y Dirección de Empresas y Análisis de Negocios (E2+Analytics) de la Universidad Pontificia Comillas al presentar mi Trabajo Fin de Grado titulado *Coffee as a commodity, market strucutre and price drivers*, declaro que he utilizado la herramienta de Inteligencia Artificial Generativa ChatGPT u otras similares de IAG de código sólo en el contexto de las actividades descritas a continuación:

1. **Brainstorming de ideas de investigación:** Utilizado para idear y esbozar posibles áreas de investigación.
2. **Crítico:** Para encontrar contra-argumentos a una tesis específica que pretendo defender.
3. **Referencias:** Usado conjuntamente con otras herramientas, como Science, para identificar referencias preliminares que luego he contrastado y validado.
4. **Constructor de plantillas:** Para diseñar formatos específicos para secciones del trabajo.
5. **Corrector de estilo literario y de lenguaje:** Para mejorar la calidad lingüística y estilística del texto.
6. **Sintetizador y divulgador de libros complicados:** Para resumir y comprender literatura compleja.
7. **Revisor:** Para recibir sugerencias sobre cómo mejorar y perfeccionar el trabajo con diferentes niveles de exigencia.
8. **Traductor:** Para traducir textos de un lenguaje a otro.

Afirmo que toda la información y contenido presentados en este trabajo son producto de mi investigación y esfuerzo individual, excepto donde se ha indicado lo contrario y se han dado los créditos correspondientes (he incluido las referencias adecuadas en el TFG y he explicitado para que se ha usado ChatGPT u otras herramientas similares). Soy consciente de las implicaciones académicas y éticas de presentar un trabajo no original y acepto las consecuencias de cualquier violación a esta declaración.

Fecha: 1 de junio de 2026

Firma: ______________________________

1. Introduction

Coffee is one of the most widely traded agricultural commodities in the world, with more than two billion cups consumed every day across the globe (Krishnan et al., 2021). Beyond its cultural and social weight, it functions as an economic engine for dozens of tropical producing economies and as a key input for the food, beverage and retail sectors of consuming countries. As discussed in the descriptive study on which this work builds, Coffee as a Commodity: Market Structure and Price Drivers (de Vega Moreno, 2026), the coffee market combines slow biological cycles, geographically concentrated production, long international logistics chains and highly financialised futures markets in a way that makes its price formation unusually complex. That earlier analysis identified a broad set of variables (climatic, cost-related, macroeconomic and financial) that comove with the global price of coffee, and built the original database that underpins the present study.

Identifying which variables matter, however, is only a first step. The natural follow-up question is how those variables interact dynamically over time, and how a shock to any one of them propagates through the rest of the system and ultimately reaches the price. This is precisely the gap that the present paper sets out to fill. Where de Vega Moreno (2026) answered the question of what is correlated with coffee prices, this work addresses how much each factor moves the price, how quickly that effect builds up and dissipates, and how important each factor is relative to the others. The instrument used to answer these questions is a Vector Autoregressive (VAR) model that captures the joint dynamics of the price and its fundamentals.

The VAR framework is well suited to this purpose for three reasons. First, it treats every variable as endogenous, which is appropriate in a market where supply, demand, input costs and macroeconomic conditions all influence one another simultaneously. Second, through a Cholesky identification scheme it allows the recovery of orthogonal structural shocks and the estimation of impulse response functions, historical decompositions and forecast error variance decompositions, tools that translate the estimated coefficients into economically interpretable statements about the magnitude, timing and relative importance of each shock. Third, the VAR has a long and influential record in the analysis of commodity and oil markets (Sims, 1980; Kilian, 2009; Baumeister & Hamilton, 2019), which provides a solid methodological precedent and a natural benchmark against which the present results can be read.

Concretely, the thesis is built around three research questions that the descriptive study could not answer on its own. How large is the effect of each fundamental on the real coffee price, once the joint dynamics of the system are taken into account? How quickly does that effect build up, and how long does it persist? And how important is each fundamental relative to the others, and relative to the price's own internal dynamics, in accounting for the variability of coffee prices over different horizons? The impulse response functions speak to the first two questions, while the forecast error variance decomposition and the historical decomposition address the third, so that the three structural tools together provide a structured answer to the overarching question of how shocks propagate through the coffee market.

The contribution of this thesis is therefore twofold. Methodologically, it shows how a descriptive map of a commodity market can be converted into a dynamic, structural model of its price-formation mechanism using a transparent and replicable workflow. Substantively, it provides quantitative magnitudes for the channels that de Vega Moreno (2026) could only describe qualitatively, and it does so over a long monthly sample running from 1999 to 2023. As a preview of the main results, the analysis finds that the real coffee price is overwhelmingly driven by its own forward-looking shock, which explains roughly 97% of its impact forecast error variance and still around 63% at a two-year horizon, while the nine fundamentals jointly account for the remaining third of the long-run variance, led by global industrial activity, rainfall, labour costs and supply-chain pressure.

The remainder of the work is organised as follows. Section 2 reviews the literature underlying both the indicators selected for the analysis and the VAR methodology used to estimate their joint dynamics, paying particular attention to which variables from the descriptive study are carried over, which are set aside, and which are introduced for the first time. Section 3 describes the dataset: the original form of each series, the transformations applied to ensure stationarity, the sample period, and a set of descriptive statistics including the contemporaneous correlation of each variable with the coffee price. Section 4 presents the empirical results and constitutes the core of the thesis, organised around the model outputs, namely the lag-selection and diagnostic statistics, the impulse response functions, the forecast error variance decomposition and the historical decomposition. Section 5 concludes with a synthesis of the findings, their connection with the theoretical framework of de Vega Moreno (2026), and the main limitations and lines for future research.

2. Literature Review

This section is organised in two blocks. The first discusses the indicators that the thesis incorporates into the VAR model, drawing on the empirical and conceptual analysis developed in de Vega Moreno (2026) and making explicit which variables are retained, which are deliberately excluded, and which are introduced for the first time. The second block sets out how the Vector Autoregressive model works, the identification scheme adopted, and the tools used to interpret it in the empirical section.

2.1 Selection of indicators

The selection of variables for the VAR is grounded in the descriptive evidence presented in de Vega Moreno (2026), where each candidate indicator was contrasted with coffee futures prices and discussed in light of the agronomic, financial and macroeconomic literature on coffee. Three criteria guide the choice. The first is economic relevance: a variable must have a credible structural link to coffee supply or demand. The second is parsimony: because a VAR estimates a large number of parameters, only the most informative representative of each economic channel is retained, so as not to exhaust the available degrees of freedom. The third is identifiability: variables whose position in a recursive causal ordering cannot be defended on economic grounds are avoided, because they would undermine the structural interpretation of the model.

Climate is the natural starting point, since coffee is highly sensitive to weather, with rainfall and temperature being the two most relevant agronomic variables (DaMatta & Ramalho, 2006; Bunn et al., 2015; Camargo, 2010). Because the Coffea plant takes several years to mature and reacts to water stress with a lag, climatic shocks reach prices only after a delay. De Vega Moreno (2026) showed empirically that contemporaneous rainfall correlations with coffee prices are weak (around 6% for Brazil, 2% for Vietnam and 8% for Colombia) but rise sharply once a one-year lag is introduced, reaching roughly 41% for Vietnam, 33% for Ethiopia, 32% for Indonesia and 22% for Colombia, in line with the biological lag between rainfall, flowering and harvest. Brazil is the notable exception: its lagged rainfall correlation is close to zero, because Brazilian production is protected by better irrigation and is instead dominated by temperature risk, in particular extreme events such as frost (Camargo, 2010; Bunn et al., 2015; Volsi et al., 2019). For this reason the model incorporates two distinct climatic variables: an aggregate rainfall measure for the main non-Brazilian producers (Colombia, Indonesia, Ethiopia and Vietnam) and a temperature variable specific to the four main Brazilian coffee states (Minas Gerais, Espírito Santo, São Paulo and Bahia).

The El Niño-Southern Oscillation (ENSO), although examined at length in de Vega Moreno (2026) as a potential global climatic driver, is deliberately excluded from the VAR. ENSO is a large-scale ocean and atmosphere oscillation that alters temperature and precipitation across the tropics (Trenberth, 1997), and it might therefore be expected to influence coffee supply. Empirically, however, the correlation between ENSO indices and coffee prices remains below 10% even when several lags are considered (de Vega Moreno, 2026). The reason, consistent with the literature, is that the effect of ENSO on coffee is strongly region-specific and tends to offset across producers: an El Niño episode may bring drought to Southeast Asia while simultaneously increasing rainfall in parts of South America, so that the opposing regional impacts cancel out in aggregate price data (DaMatta & Ramalho, 2006; Bunn et al., 2015; Camargo, 2010). Including ENSO would therefore add a noisy variable with little explanatory power, which in a moderately sized VAR is an inefficient use of degrees of freedom; its supply-relevant information is, moreover, already captured more directly by the rainfall and temperature variables that are retained.

Input costs constitute the second channel. De Vega Moreno (2026) documented strong correlations between coffee futures and both fertiliser prices (around 50%, proxied by the FRED Producer Price Index for fertiliser materials) and energy prices (around 51%, proxied by the FRED Global Price of Energy Index), in line with the broader literature that identifies input costs as one of the most stable and economically meaningful drivers of agricultural commodity prices (Gilbert, 2010). Both variables are retained. The pesticide price index, although correlated with coffee at roughly 43% in the descriptive study, is excluded to avoid redundancy: it belongs to the same FRED producer-price family as fertilisers, shares common cost drivers with it, and the resulting near-collinearity would blur the identification of separate structural shocks (Staver et al., 2001).

Labour is the third channel, since coffee production remains highly labour-intensive, particularly in harvesting and post-harvest handling. De Vega Moreno (2026) showed that minimum wages expressed in US dollars are strongly correlated with coffee prices across the four main producing countries for which reliable data exist, namely Brazil, Colombia, Vietnam and Indonesia, with correlations ranging from about 51% to 65%, while Ethiopia is excluded for lack of a dependable series (Countryeconomy.com, n.d.). Rather than introduce four highly correlated wage series, the model aggregates them into a single average minimum-wage variable, which preserves the labour-cost signal, captures the common wage trend along the producing supply chain and avoids inflating the dimensionality of the system.

Logistics is where the present analysis departs most clearly from the descriptive study, and it introduces the one genuinely new indicator. Coffee is overwhelmingly traded internationally and shipped along long maritime routes (International Coffee Organization, 2023), and de Vega Moreno (2026) captured this dimension through two separate proxies: a deep-sea freight cost index (correlated with coffee at about 57%) and the Brent crude oil price (about 49%), which were themselves highly correlated with one another (around 68%). For the dynamic model, both are replaced by the Global Supply Chain Pressure Index (GSCPI) published by the Federal Reserve Bank of New York. The GSCPI is a single standardised index that integrates more than twenty-seven variables related to maritime and air freight costs, container shipping, delivery times, inventories and order backlogs across the seven largest economies (Benigno et al., 2022). This substitution is motivated by two considerations. First, the GSCPI measures supply-chain stress in a far broader and more synthetic way than freight or oil prices alone, which is precisely the kind of disruption to which a commodity dependent on long shipping routes is exposed. Second, replacing two highly correlated proxies with one parsimonious index reduces the parameter count and sharpens identification. The Brent oil price is dropped on the same grounds, because its core information for coffee, the cost of fuel for shipping, is already absorbed jointly by the retained energy index and the GSCPI. Supply-chain pressure, absent from the descriptive study, thus enters the analysis as a new and conceptually distinct driver.

Demand-side and macroeconomic forces form the fourth channel, although their explanatory power is more limited than that of supply variables (de Vega Moreno, 2026). The World Industrial Production Index (WIPI) compiled by Baumeister and Hamilton (2019), correlated with coffee at about 34% in year-on-year terms, is retained as a proxy for global activity and the demand pressure that the world economy exerts on commodity prices. Its inclusion is also methodologically valuable, because it is the standard global-activity measure in the structural VAR literature on commodities (Kilian, 2009; Baumeister & Hamilton, 2019) and allows the present results to be benchmarked against that tradition. By contrast, the Global and US Economic Policy Uncertainty indices, although correlated with coffee at roughly 55% and 47% respectively, are excluded: they operate primarily through expectations and hedging behaviour rather than through physical supply or demand, and it is difficult to defend a specific position for such expectation-driven variables within a recursive Cholesky ordering without additional identifying assumptions. World GDP per capita, explored in the descriptive study with a correlation of about 29%, is likewise excluded on parsimony grounds, since it is observed only annually, moves slowly, and is largely subsumed by the monthly and more commodity-relevant WIPI (World Bank, n.d.).

The financial channel is captured through exchange rates. Because Arabica futures are denominated in US dollars while production takes place in countries with their own currencies, exchange-rate movements affect both producers' incentives and the dollar price of coffee, with the pass-through being quality and origin dependent (Anders & Fedoseeva, 2017). De Vega Moreno (2026) documented correlations of about 42% between the Brazilian real and coffee and about 36% between the Colombian peso and coffee. Both exchange rates are retained as separate variables rather than aggregated, because Brazil and Colombia occupy distinct segments of the Arabica market, Brazil being the largest producer with roughly 35% of global supply and Colombia the leading mild-Arabica origin with about 8%, and because the two currencies are exposed to different macroeconomic and political shocks, so that modelling them separately lets the VAR distinguish Brazil-specific from Colombia-specific exchange-rate effects.

Finally, the dependent variable of interest is the real price of the Arabica coffee future (the ICE Coffee 'C' contract), obtained from FactSet at monthly frequency and deflated to constant 2026 US dollars using the US Consumer Price Index from FRED, so that the analysis abstracts from general inflation over a sample spanning more than two decades. Taken together, the final system of ten variables is designed to represent the four economic channels through which the descriptive study found that fundamentals reach coffee prices, namely climate, input and labour costs, logistics and global activity, and finance, while remaining parsimonious enough to support a credible recursive identification.

2.2 The vector autoregressive model

The dynamics of the coffee market are estimated with a Vector Autoregressive (VAR) model (Sims, 1980). A VAR explains each variable in the system by the recent past of every variable, itself included, so that $Y_t = c + A_1 Y_{t-1} + \dots + A_p Y_{t-p} + u_t$, where Y_t collects the ten series and u_t are the residuals. Because no variable is treated as exogenous, the model lets a disturbance to any one of them ripple through the whole system over the following months. This is precisely what the thesis requires, since its question is not whether the fundamentals correlate with the price, but how a shock to each of them is transmitted to the coffee price, how fast, and how strongly. The model is estimated equation by equation by ordinary least squares.

The number of past months included, the lag order p, is the main specification choice. Although the Akaike and Bayesian information criteria (Akaike, 1974; Schwarz, 1978) are both minimised at $p = 1$ (Section 4), the model uses $p = 12$. Coffee is a perennial crop whose harvest cycle, shipping logistics and cost pass-through unfold over a full calendar year, so twelve monthly lags are needed to capture the delayed transmission that the study is about; a one-lag model would suppress exactly those dynamics. The decisive check is statistical: at $p = 12$ the residuals show no serial correlation under the Ljung-Box test (Section 4). Stationarity of the inputs is ensured by the transformations of Section 3 and assessed with the Augmented Dickey-Fuller test reported there.

To turn the estimated relationships into economically meaningful shocks, the residuals are orthogonalised through a Cholesky decomposition, which imposes an order of contemporaneous causality: a shock to a variable may move the variables placed after it within the same month, but reaches those placed before it only with a lag. The ordering therefore has to be defended on economic grounds (Kilian & Lütkepohl, 2017), and the one used here runs from the most exogenous drivers to the most domestic. It places the two climate variables first, rainfall and then temperature, since weather is fundamentally exogenous to every economic variable in the system. The Global Supply Chain Pressure Index follows, treated as exogenous because it is constructed from forward-looking shipping and logistics data (Benigno et al., 2022). Next comes the World Industrial Production Index, the proxy for global economic activity and hence for commodity demand (Kilian, 2009; Baumeister & Hamilton, 2019), followed by energy prices, which at this position capture the energy shocks not already explained by global demand (Kilian, 2009), and then fertiliser prices, which capture fertiliser supply shocks not already explained by demand or by energy costs. The real coffee price is placed after this global block but before a final domestic block formed by the Brazilian real, the Colombian peso and minimum wages. Coffee is positioned ahead of the domestic block on the argument that, although these are important determinants, the coffee price is set in a global market, so it is hard to believe that country-specific exchange rates or wages move it within a single month; the domestic variables are ordered last, with wages after the two currencies, because a change in wages is unlikely to move the value of a currency within the month, and none of these country-specific variables is likely to move the larger global block contemporaneously.

Three tools then read the model and map onto the research questions. Impulse response functions trace how the coffee price reacts, month by month over a twenty-four month horizon, to a one-standard-deviation shock to each variable, answering how much and how fast; they are shown with ±1.645 standard-deviation bands, which correspond to an approximate 90% confidence level. The forecast error variance decomposition measures the share of the price's variability explained by each shock at different horizons, answering how important each driver is. The historical decomposition attributes the price's actual movements at each date to the shocks that occurred, $HD_{k,j}(t) = \Sigma_h \Psi_h[k, j] \cdot e_j(t - h)$ (Kilian & Lütkepohl, 2017), linking the model to real episodes in the coffee market. Together they convert the VAR into a complete account of how the price responds to its fundamentals.

3. Data Description

The empirical analysis relies on a balanced monthly panel covering January 1999 to December 2023, which yields 300 observations on each of the ten variables. The panel was constructed specifically for this two-part research project: the underlying raw series were collected and documented in de Vega Moreno (2026), and were then re-processed and transformed for the dynamic analysis carried out here. This section describes the original form of each series, the transformations applied to render them stationary and economically interpretable, the resulting descriptive statistics, and the sample period.

3.1 Original data, transformations and stationarity

In their original form, the ten series are heterogeneous in both frequency and units, which is one of the main practical challenges in modelling a commodity market that joins financial, agronomic and macroeconomic data. The coffee price is the monthly settlement price of the Arabica ICE Coffee 'C' future, expressed in US dollars and retrieved from FactSet; it is first deflated to constant 2026 dollars using the US Consumer Price Index from the Federal Reserve Economic Data (FRED) service, producing a real price level, and then expressed as a month-on-month change. Rainfall is originally a set of annual precipitation totals, measured in millimetres for each producing country and sourced from the Humanitarian Data Exchange; the totals for Colombia, Indonesia, Ethiopia and Vietnam are aggregated into a single previous-year sum and then expressed as a year-on-year change, with Brazil deliberately omitted. Temperature is originally the mean temperature, in degrees Celsius, of the four main Brazilian coffee-producing states, taken from Global Data Lab and averaged into a single regional series before being expressed as a year-on-year change.

The Global Supply Chain Pressure Index is already published by the Federal Reserve Bank of New York as a standardised monthly index with mean zero and unit variance, so it enters the model in its native, absolute form without further transformation (Benigno et al., 2022). The World Industrial Production Index of Baumeister and Hamilton (2019), the Global Price of Energy Index and the Fertilizer Materials producer price index, the latter two obtained from FRED, are all originally index levels and are expressed as month-on-month changes. The average minimum wage is built from the annual statutory minimum wages of Brazil, Colombia, Vietnam and Indonesia, converted to US dollars, averaged across the four countries (Countryeconomy.com, n.d.) and expressed as a year-on-year change. Finally, the Brazilian real and

Colombian peso exchange rates against the US dollar are originally monthly nominal levels from FactSet and are expressed as month-on-month changes.

Three transformation types are therefore used, each chosen to match the nature of the underlying series. Month-on-month changes are applied to the high-frequency financial and price series (the coffee price, the energy and fertiliser indices, the WIPI and the two exchange rates), which yields stationary series that retain the genuine monthly variation of interest. Year-on-year changes are applied to the series that are intrinsically annual (rainfall, temperature and minimum wages), because computing a twelve-month change captures their authentic variation while removing the strong seasonal pattern and the intra-annual constancy that would otherwise dominate. The GSCPI is left untransformed because it is, by construction, already a stationary standardised index. This treatment follows standard practice in time-series analysis, where estimating a VAR on non-stationary variables produces inconsistent estimates and spurious dynamics (Hamilton, 1994; Lütkepohl, 2005). Table 1 summarises the variables, their original form, their transformation and their source.

Table 1. Variables in the VAR model: original form, transformation and source.

Variable	Original form	Transformation	Source
Real coffee price	Monthly Arabica ICE 'C' settlement price (USD), deflated to 2026 USD	MoM change	FactSet; US CPI from FRED
Rainfall (CO+ID+ET+VN)	Annual precipitation totals (mm) by country	YoY of previous-year sum	Humanitarian Data Exchange
Temperature (Brazil)	Mean annual temperature (°C), four states	YoY change	Global Data Lab
GSCPI	Standardised monthly index (mean 0, sd 1)	Absolute (native)	Federal Reserve Bank of New York
WIPI	World industrial production index level	MoM change	Baumeister & Hamilton (2019)
Energy price	Global Price of Energy Index level	MoM change	FRED (PNRGINDEXM)
Fertiliser price	PPI fertiliser materials index level	MoM change	FRED
Avg. minimum wage	Annual statutory minimum wage (USD), four countries	YoY change	Countryeconomy.com
BRL/USD	Monthly nominal exchange-rate level	MoM change	FactSet
COP/USD	Monthly nominal exchange-rate level	MoM change	FactSet

Note. Own elaboration. The aggregate rainfall variable excludes Brazil, whose climatic risk is dominated by temperature. The pesticide price index, deep-sea freight index, Brent oil price, ENSO index, World GDP per capita and Economic Policy Uncertainty indices, all explored in de Vega Moreno (2026), are excluded for the reasons given in Section 2.1.

The stationarity of each transformed series is examined formally with the Augmented Dickey-Fuller test (Dickey & Fuller, 1979), including a constant and a lag length chosen by the Akaike criterion. The results are reported in Table 2. The unit-root null is rejected at the 5% level for nine of the ten series, frequently with very large test statistics. The single exception is the average minimum wage, for which the test does not reject at conventional levels ($p = 0.13$). This borderline outcome reflects the strong persistence induced by observing an annually-set figure at monthly frequency, which gives the year-on-year wage series a markedly slow-moving, step-like profile. Because a year-on-year growth rate is bounded and cannot contain a genuine unit root, and because the series clearly reverts over the sample (Figure 1), it is retained in the model; this point is nonetheless noted among the limitations in Section 5. The estimation can therefore proceed in the levels of the transformed series, without further differencing or cointegration analysis.

Table 2. Augmented Dickey-Fuller stationarity tests on the transformed series.

Variable	ADF statistic	p-value	Reject unit root (5%)
Real coffee price (MoM)	−14.48	<0.001	Yes
Rainfall (YoY)	−3.67	0.005	Yes
Temperature (YoY)	−3.63	0.005	Yes
GSCPI (absolute)	−3.57	0.006	Yes
WIPI (MoM)	−7.06	<0.001	Yes
Energy price (MoM)	−12.24	<0.001	Yes
Fertiliser price (MoM)	−8.34	<0.001	Yes
Avg. minimum wage (YoY)	−2.46	0.125	No
BRL/USD (MoM)	−12.29	<0.001	Yes
COP/USD (MoM)	−11.10	<0.001	Yes

Note. Own elaboration. Augmented Dickey-Fuller test with an intercept and an automatic lag length (Akaike criterion). The null hypothesis is the presence of a unit root; rejection indicates stationarity. The minimum-wage series is the only borderline case, for the reasons discussed in the text.

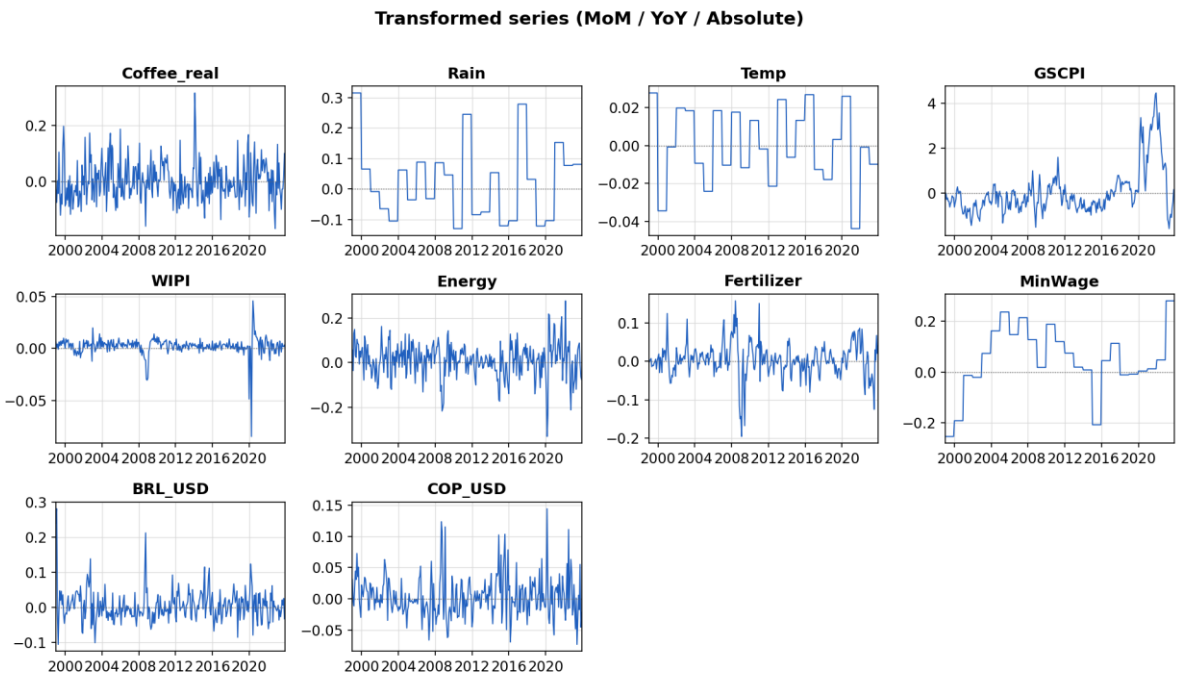

Figure 1. Transformed series entering the VAR (month-on-month, year-on-year, or native standardised form), January 1999 to December 2023. Own elaboration.

Figure 1 plots the ten transformed series. Several features are visually apparent and are confirmed by the formal stationarity tests: all series fluctuate around a stable level without obvious trends, which is the intended effect of the transformations. The month-on-month series (coffee, energy, fertiliser, WIPI and the two exchange rates) display the high-frequency, mean-reverting noise typical of differenced financial data, with occasional volatility clusters around well-known stress episodes. The year-on-year series (rainfall, temperature and minimum wages) are visibly smoother and show the characteristic step-like persistence induced by annual data observed monthly, which is most pronounced for the minimum-wage series and explains its borderline stationarity test. The GSCPI, left in its native form, is the only series whose level carries information directly, and its pronounced spike in 2021 corresponds to the post-pandemic supply-chain crisis.

3.2 Descriptive statistics and correlation with the coffee price

Table 3 reports the mean, standard deviation, minimum and maximum of each variable over the full sample, together with its contemporaneous correlation with the real coffee price. The averages are economically sensible: the real coffee price grows by about 0.19% per month on average, consistent with a moderate real appreciation of the

Arabica future over twenty-five years; the GSCPI has a near-zero mean and unit-scale dispersion, in line with its construction as a standardised index, and reaches a maximum of 4.47 standard deviations at the December 2021 peak of the supply-chain crisis (Benigno et al., 2022); and the average minimum wage in producing countries grows by about 4.75% per year, reflecting their domestic inflation dynamics.

The most instructive column of Table 3 is the last one. Once the series are transformed to ensure stationarity, the contemporaneous correlations between the fundamentals and the coffee price are uniformly small: the largest in absolute value are the Brazilian real (−0.25) and the Colombian peso (−0.17), followed by the GSCPI (0.12), the minimum wage (0.11) and the WIPI (0.09), while rainfall and temperature are essentially uncorrelated with the contemporaneous price. This stands in sharp contrast with de Vega Moreno (2026), which reported much higher correlations, of the order of 50% for energy, fertilisers and minimum wages, computed on the variables in levels. The contrast is not a contradiction but a direct consequence of the transformation: correlations in levels are dominated by shared long-run trends, so that any two trending series appear strongly related, whereas differencing removes those common trends and leaves only the high-frequency comovement, which is genuinely weak on a contemporaneous, month-to-month basis.

This observation is, in fact, the central motivation for the modelling strategy adopted in this thesis. If the fundamentals were strongly correlated with the price contemporaneously, a simple static regression would suffice; the fact that they are not implies that whatever relationship exists must operate through lags and through the joint dynamics of the system, exactly the kind of structure a VAR is designed to capture. The negative sign on the two exchange rates is itself informative and consistent with exchange-rate pass-through: a month in which the producing-country currency depreciates against the dollar (a rise in BRL/USD or COP/USD) tends to coincide with a softer dollar price of coffee, as cheaper local currency reduces the dollar cost of supply (Anders & Fedoseeva, 2017; Durevall, 2018). The remainder of the thesis therefore turns from contemporaneous correlation to dynamic, lagged transmission, which is where the economically meaningful relationships are expected to reside.

Table 3. Descriptive statistics and contemporaneous correlation with the real coffee price (January 1999 to December 2023, N = 300).

Variable	Mean	Std. dev.	Min	Max	Corr. coffee
Real coffee price (MoM)	0.0019	0.0705	−0.1665	0.3148	1.000
Rainfall (YoY)	0.0240	0.1242	−0.1295	0.3147	0.023
Temperature (YoY)	0.0001	0.0196	−0.0437	0.0277	0.019
GSCPI (absolute)	0.0344	1.0470	−1.5800	4.4700	0.124
WIPI (MoM)	0.0021	0.0087	−0.0849	0.0457	0.094
Energy price (MoM)	0.0089	0.0771	−0.3321	0.2759	0.061
Fertiliser price (MoM)	0.0045	0.0437	−0.1953	0.1575	0.077
Avg. minimum wage (YoY)	0.0475	0.1288	−0.2542	0.2808	0.110
BRL/USD (MoM)	0.0056	0.0445	−0.1053	0.2816	−0.245
COP/USD (MoM)	0.0037	0.0320	−0.0730	0.1444	−0.166

Note. Own elaboration on the sources listed in Table 1. Series in MoM or YoY are dimensionless growth rates; the GSCPI is expressed in standard deviations from its mean. The final column is the Pearson correlation of each transformed series with the transformed real coffee price.

3.3 Sample period

The choice of January 1999 as the start of the sample is dictated by data availability after the transformations. The GSCPI begins in January 1998, and the year-on-year computation for rainfall, temperature and minimum wages consumes a further twelve observations, which moves the first usable month to January 1999. The sample ends in December 2023, the most recent month for which all ten series were available at the time of writing. Deliberately, the sample stops short of the extraordinary coffee price increases observed in 2024 and 2025: incorporating that episode would mix a structural-analysis sample with an unusual, ongoing price spike whose dynamics may not be representative of the historical relationships the model is meant to estimate, and it is therefore left as a natural robustness extension for future research. The resulting window of 300 months is long enough to span several full climatic, macroeconomic and supply-chain cycles, which is important for the credible estimation of a richly lagged VAR.

4. Empirical Results

This section presents the empirical implementation of the VAR and is organised around the model outputs. Section 4.1 reports the specification, the lag-selection statistics and the residual diagnostics. Section 4.2 discusses the impulse response functions, which describe how a shock to each variable propagates to the coffee price. Section 4.3 reports the forecast error variance decomposition, which quantifies the relative importance of each shock at different horizons. Section 4.4 presents the historical decomposition, which attributes the realised movements of the coffee price to the estimated structural shocks. The model is estimated in MATLAB using the Econometrics Toolbox.

4.1 Specification, lag selection and diagnostics

The VAR is estimated on the ten transformed variables of Section 3 over the sample of 300 months, with the variables ordered as described in Section 2.2: rainfall, temperature, the GSCPI, the WIPI, energy, fertilisers, the real coffee price, the Brazilian real, the Colombian peso and minimum wages. With N = 10 variables and p = 12 lags, each equation contains 1 + 10 × 12 = 121 parameters, estimated from 288 effective observations after the twelve initial lags are used up. This is a demanding but feasible specification, and the estimated system has a log-likelihood of 6,289.0.

Although a system with 121 parameters per equation is demanding, the 288 effective observations leave a margin of roughly 167 degrees of freedom in each equation, which is sufficient for stable estimation of a model of this size. The estimated VAR is dynamically stable, in the sense that all the eigenvalues of its companion matrix lie inside the unit circle, so the impulse responses and decompositions reported below are well defined and converge over the chosen horizon.

Figure 2 reports the lag-selection exercise. Both the Akaike and the Bayesian information criteria are computed for every lag order from one to twelve, and both are monotonically increasing in the number of lags over this range, so that both are minimised at p = 1. Taken at face value, the information criteria therefore point to a first-order model. As explained in Section 2.2, the model is nonetheless estimated with twelve lags, for reasons that are economic, statistical and diagnostic. Monthly data on a perennial crop require a lag structure able to span the annual cycle and to represent the delayed transmission of weather, cost and supply-chain shocks that the study is designed to measure; the information criteria, which penalise the parameter count of a richly lagged system very heavily, favour a specification that is optimal for one-step forecasting but that would suppress precisely those dynamics; and, as the diagnostics below confirm, only the richer specification produces residuals that are free of serial correlation. The dashed line in Figure 2 marks the chosen order.

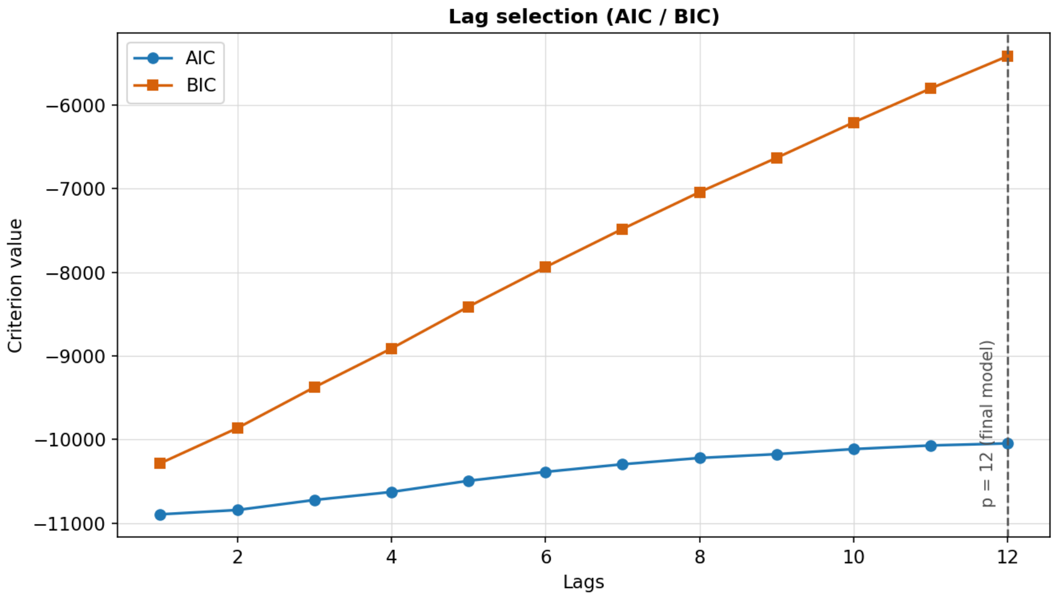

Figure 2. Akaike and Bayesian information criteria as a function of the lag order, p = 1 to 12. Both criteria are minimised at p = 1; the model is estimated at p = 12 (dashed line). Own elaboration.

The adequacy of the chosen lag order is verified through the Ljung-Box test for residual autocorrelation. Applied to the residuals of the coffee equation with twelve lags, the test yields a p-value of 0.999, so the null hypothesis of no serial correlation cannot be rejected at any conventional level. This is a strong result: it indicates that the twelve-lag specification has absorbed essentially all of the predictable time-series structure in the coffee residual, leaving white noise behind, and it directly validates the decision to override the information criteria. A first-order model, by contrast, would almost certainly have left seasonal and lagged structure in the residuals, invalidating the inference that follows. The combination of a clean residual diagnostic and an economically motivated lag length provides a sound basis for the structural analysis presented in the remainder of the section.

4.2 Impulse response functions

Figure 3 collects the impulse response functions, each panel showing the response of the real coffee price to a one-standard-deviation structural shock to one of the ten variables over a twenty-four month horizon, surrounded by ±1.645 standard-deviation bands (an approximate 90% confidence level). The single most prominent feature is the own price shock: a one-standard-deviation coffee shock raises the monthly real return by about five percentage points on impact and then collapses to essentially zero within two to three months. This pattern is exactly what financial theory predicts for a forward-looking, informationally efficient futures price. Because the dependent variable is a monthly return, a transitory response of the return corresponds to a permanent shift in the price level: the shock moves the price to a new level immediately and the price then behaves like a near-martingale, with no predictable further drift.

The recursive ordering shapes the contemporaneous responses in an informative way. The variables placed before coffee, namely the GSCPI, the WIPI, energy and fertilisers, are allowed to move the price within the same month, and they do: a supply-chain pressure shock and a global-activity shock each raise the coffee return by about 0.4 to 0.5 percentage points on impact, an energy shock by about 0.3 points, and a fertiliser shock lowers it by about 0.5 points on impact. The two climate variables are also ordered before coffee but produce a near-zero impact response, which is consistent with the very weak contemporaneous correlation between weather and the price documented in Section 3.2; their effect, as shown below, builds up only with a lag. By contrast, the domestic block placed after coffee, the two exchange rates and minimum wages, has an impact response of exactly zero by construction, reflecting the assumption that country-specific shocks do not move the globally-set price within the month; these variables reach the price only from the second month onward.

Beyond the impact period the responses to the fundamentals are an order of magnitude smaller than the own shock, peaking at between roughly 0.5 and 0.8 percentage points, and they oscillate around zero, with much of each path lying inside the wider ±1.645 standard-deviation bands. Individually, therefore, the dynamic effect of any single fundamental is modest and estimated with considerable uncertainty. Because the period-by-period responses oscillate, the economically meaningful information is most clearly read from the cumulative responses, which net out the oscillation and measure the total effect of a shock over the full two-year horizon. Read in this way, the signs align with economic priors for the most important drivers. A positive rainfall shock in the producing countries, a signal of better expected yields, generates the largest cumulative effect of any fundamental, lowering the real coffee price by about 3.4% over twenty-four months, consistent with the supply interpretation of the climatic channel and with the biological lag between rainfall and harvest. A positive minimum-wage shock raises the price cumulatively by about 2.9%, the clearest cost-push response in the system, and a positive supply-chain pressure shock raises it by about 1.7%. A positive global-activity shock raises the price modestly, consistent with a demand channel, while the fertiliser and energy responses are more oscillatory and the two exchange rates produce small cumulative effects whose sign is less precisely determined at this horizon.

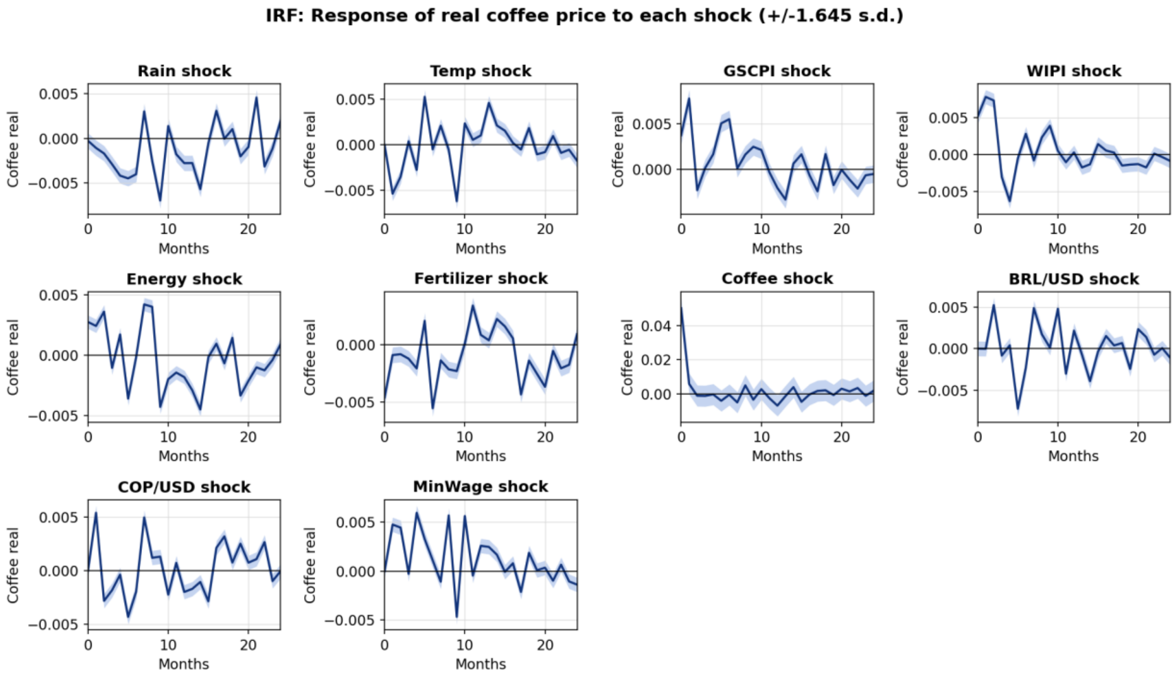

Figure 3. Impulse responses of the real coffee price to a one-standard-deviation structural shock to each variable, 24-month horizon, with ±1.645 standard-deviation bands. Panels follow the Cholesky ordering. Own elaboration.

4.3 Forecast error variance decomposition

The forecast error variance decomposition provides a compact summary of the relative importance of each shock. Table 4 reports the share of the forecast error variance of the real coffee price attributable to each structural shock at the impact horizon and at one, six, twelve and twenty-four months, and Figure 4 plots the contributions of the nine exogenous shocks across the full horizon. Because the coffee price is no longer ordered first, the impact horizon is itself informative: about 97% of the contemporaneous variance is attributed to the own coffee shock, while the global block ordered before it (the GSCPI, the WIPI, energy and fertilisers) already accounts for close to 3% within the month, and the climate and domestic variables contribute essentially nothing on impact.

Three patterns stand out. First, the own coffee shock dominates at every horizon, but its share declines steadily as the horizon lengthens: from about 97% on impact to 90% at one month, 76% at six months, 68% at twelve months and 63% at twenty-four months. This gradual erosion is the quantitative counterpart of the impulse-response evidence, with the price driven primarily by its own forward-looking innovations in the short run and the fundamentals making themselves felt only slowly. Second, the contribution of the fundamentals builds up gradually and stabilises around the second year, by which point the nine exogenous shocks jointly explain roughly 37% of the long-run forecast error variance. Third, the ranking of the fundamentals at the two-year horizon is economically coherent: global industrial activity and rainfall are the most important exogenous drivers, at about 5.3% each, followed by the average minimum wage (4.7%), supply-chain pressure (4.3%) and the Brazilian exchange rate (4.2%), then by temperature (3.8%), energy (3.5%), fertilisers (3.3%) and the Colombian exchange rate (3.2%).

This ranking is the quantitative validation of the qualitative conclusions of de Vega Moreno (2026). The four channels that the descriptive analysis identified as the principal routes through which fundamentals reach coffee prices, namely climate, global activity, labour costs and finance, re-emerge here at the top of the variance decomposition, but now with explicit magnitudes attached. The VAR thus does more than confirm that these factors matter; it measures how much each of them matters and at what horizon, converting the descriptive map of de Vega Moreno (2026) into a quantified ordering of drivers.

It is also instructive to compare these variance shares with the level correlations reported in the descriptive study. There, energy and fertilisers showed some of the highest raw correlations with the price, yet here they rank among the smaller contributors to the forecast error variance, while rainfall and labour costs, whose lagged rather than contemporaneous links were emphasised in the descriptive analysis, rise to the top. This re-ordering is exactly what one would expect once the analysis moves from static, level-based correlation to a dynamic decomposition that isolates orthogonal shocks: it confirms that the channels which matter for the coffee price are those that operate with a lag, through the slow-moving fundamentals of supply and labour, rather than those that merely share a common trend with the price.

Table 4. Forecast error variance decomposition of the real coffee price (%), by horizon.

Structural shock	Impact	1 month	6 months	12 months	24 months
Rainfall	0.0	0.0	2.0	3.6	5.3
Temperature	0.0	1.0	2.3	3.2	3.8
GSCPI	0.6	2.6	4.1	4.0	4.3
WIPI	1.0	3.0	5.8	5.5	5.3
Energy	0.3	0.5	1.3	2.6	3.5
Fertiliser	0.8	0.8	1.9	2.2	3.3
Coffee (own)	97.3	90.2	75.7	67.9	62.5
BRL/USD	0.0	0.0	2.5	3.8	4.2
COP/USD	0.0	1.0	1.8	2.5	3.2
Min. wage	0.0	0.8	2.7	4.6	4.7

Note. Own elaboration from the estimated VAR(12) identified by Cholesky factorisation with the ordering of Section 2.2. Columns sum to 100% up to rounding. At the impact horizon the domestic block (exchange rates and wages), ordered after coffee, contributes zero by construction, whereas the global block ordered before coffee already contributes within the month.

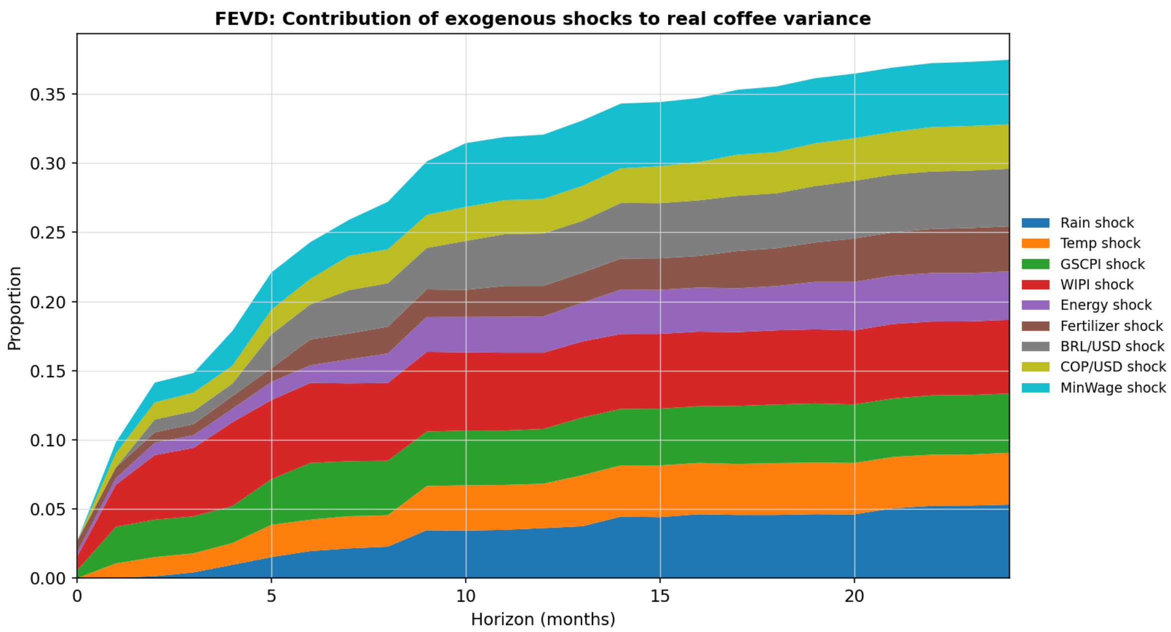

Figure 4. Contribution of the nine exogenous structural shocks to the forecast error variance of the real coffee price, by horizon (the residual up to one is the own coffee shock). Own elaboration.

Figure 4 makes the build-up visible. The exogenous contribution rises from about 3% on impact to roughly 37% by month twenty-four, with most of the increase concentrated in the first twelve months, after which the shares flatten. The figure also shows that no single fundamental dominates: the long-run variance is spread fairly evenly across global activity, rainfall, minimum wages, supply-chain pressure and the two exchange rates, which is consistent with the view of coffee as a commodity buffeted by several distinct forces of comparable size rather than by one overriding factor.

4.4 Historical decomposition

While the FEVD describes the average importance of each shock across all possible histories, the historical decomposition asks a complementary, date-specific question: how much of the realised movement of the coffee price in each month of the sample can be attributed to each structural shock. Because nine exogenous shocks are difficult to read in a single chart, Figure 5 groups them into four economically meaningful blocks, climate, global demand and supply chain, input costs, and domestic factors, and plots each group separately between 2002 and 2023.

Two messages emerge. The first is one of magnitude and is fully consistent with the variance decomposition: the contributions of every group are high-frequency and modest in size, individually within roughly ±0.05 in monthly real-return terms, confirming that month by month the fundamentals push and pull the price around a path that is largely set by its own dynamics. The second message concerns timing. The amplitude of the contributions is not constant but widens markedly around the global stress episodes that bracket the sample. The global demand and supply-chain block shows its largest swings in 2008 and 2009, during the global financial crisis, when activity collapsed across commodity markets (Kilian, 2009; Baumeister & Hamilton, 2019), and again in 2020 to 2022, when the pandemic and the subsequent supply-chain and energy disruptions associated with the war in Ukraine produced unusually large contributions. The climate block contributes a clear positive push around the 2014 Brazilian drought, illustrating the lagged transmission of weather stress to prices (Camargo, 2010; de Vega Moreno, 2026), while the input-cost and domestic blocks contribute steadier but smaller amounts throughout.

A closer look at which shocks contribute most over the sample reinforces the FEVD ranking. Measured by their average absolute contribution, rainfall and minimum-wage shocks are the most influential exogenous drivers historically, closely followed by global activity and the Brazilian exchange rate, the same set of variables that lead the long-run variance decomposition. The grouped historical decomposition therefore does not merely repeat the FEVD in a different format; it grounds it in identifiable episodes and tells a clearer story group by group, showing that the statistical importance of these channels corresponds to recognisable events in the recent history of the coffee market. Taken together, the impulse responses, the variance decomposition and the historical decomposition tell a single coherent story: a price dominated in the short run by its own forward-looking innovations, gradually shaped over the medium run by a balanced set of climatic, cost, activity and financial fundamentals, whose influence becomes most visible during periods of global disruption.

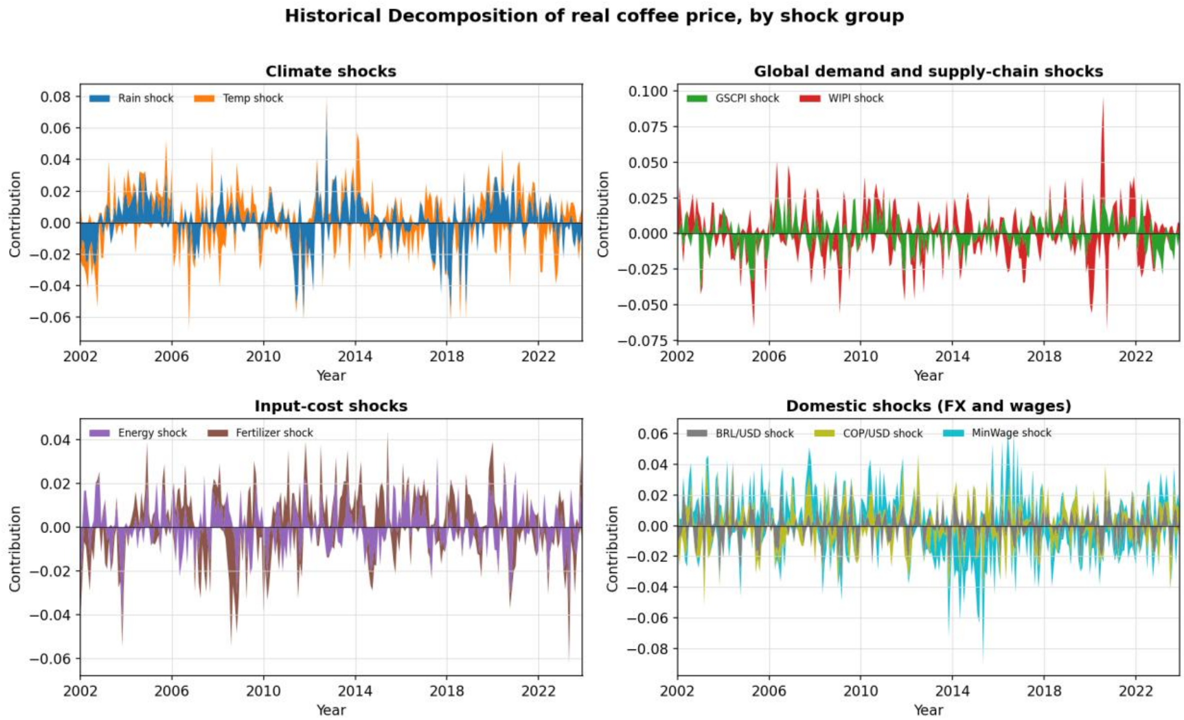

Figure 5. Historical decomposition of the real coffee price, with the nine exogenous shocks grouped into climate, global demand and supply chain, input costs, and domestic factors, 2002 to 2023. Own elaboration.

5. Conclusions

This paper has built a structural Vector Autoregressive model of the coffee market that quantifies, for the first time within this two-part research project, the dynamic interactions between the real price of coffee and nine fundamental drivers identified in de Vega Moreno (2026): rainfall in the main non-Brazilian producers, Brazilian temperature, the Global Supply Chain Pressure Index, the World Industrial Production Index, energy prices, fertiliser prices, the average minimum wage of producing countries, and the bilateral exchange rates of Brazil and Colombia against the US dollar. The model was estimated on monthly data for 1999 to 2023 and identified through a recursive Cholesky decomposition that runs from the most exogenous drivers (climate, supply-chain pressure, global activity and input costs) through the real coffee price to a final domestic block of exchange rates and wages.

Three main empirical findings stand out. First, the real coffee price displays a very high degree of self-persistence: its own shock explains about 97% of the forecast error variance on impact and still around 63% at twenty-four months, while the impulse responses show that the own shock moves the price to a new level almost instantly and then dies out. This is the empirical signature of an informationally efficient, forward-looking futures market (Nguyễn et al., 2020; Gomes et al., 2019). Second, once the fundamentals are allowed to operate, their effects build up gradually and become collectively significant by the second year, when the nine exogenous shocks jointly account for roughly 37% of the long-run variance, led by global activity and rainfall (about 5.3% each), minimum wages, supply-chain pressure and the Brazilian exchange rate (each above 4%). Third, the grouped historical decomposition links the estimated shocks to identifiable episodes, the 2008 to 2009 financial crisis, the 2014 Brazilian drought and the 2020 to 2022 pandemic and supply-chain crisis, providing a narrative that aligns with both the descriptive analysis of de Vega Moreno (2026) and the wider literature on commodity prices.

Methodologically, the work illustrates how a descriptive framework can be enriched with a structural econometric model. Where de Vega Moreno (2026) established which variables are associated with coffee prices, the present analysis establishes how much they matter, at what horizon and through which channel. A noteworthy methodological lesson concerns the role of the data transformation: the correlations between the fundamentals and the price, which appeared strong in levels in the descriptive study, fall close to zero once the series are differenced to achieve stationarity, precisely because differencing removes the shared trends that inflate level correlations. This is not a weakening of the earlier findings but a clarification of them: the genuine relationships are dynamic and lagged rather than contemporaneous, which is exactly why a VAR, and not a static regression, is the appropriate tool.

The results carry practical implications for the actors exposed to coffee-price risk. The finding that the price behaves, in the short run, like a near-martingale dominated by its own forward-looking shock implies that point forecasts of next month's price add little value beyond the current price, and that effort is better directed at managing risk than at predicting direction. The variance decomposition offers a concrete map for that task: because global activity, rainfall, labour costs, supply-chain pressure and the producing-country exchange rates each account for a measurable share of the medium-run variance, roasters, traders and producer cooperatives can monitor this compact set of indicators as early-warning signals and can hedge the corresponding exposures, namely currency risk through the Brazilian real and Colombian peso, cost risk through energy and fertiliser positions, and logistics risk through supply-chain indicators, rather than treating the coffee price as an undifferentiated black box.

Several limitations should be acknowledged. The Cholesky ordering, although defended on economic grounds, is itself an identifying restriction, and a fuller robustness analysis using alternative orderings, sign restrictions or external instruments would strengthen the structural interpretation. The impulse-response bands reported here are approximate analytical bands at ±1.645 standard deviations; a bootstrap or Monte Carlo procedure would deliver more reliable inference about the individual responses, several of which are imprecisely estimated. The lag length of twelve was chosen on economic and diagnostic grounds even though the information criteria pointed to a first-order model; while the clean Ljung-Box diagnostic supports that decision, the tension between the criteria and the chosen specification is real and would merit a formal sensitivity analysis across intermediate lag lengths. The average minimum wage is the one input whose stationarity is borderline, owing to the strong persistence of an annually-set series observed monthly, and a mixed-frequency or MIDAS specification could handle this and the other annual series more efficiently.

Finally, the sample ends in December 2023 and so excludes the dramatic price increases of 2024 and 2025, which constitute a natural and topical extension.

Beyond these specifics, the two-part research project on coffee illustrates a broader point about the value of combining descriptive and quantitative perspectives on commodity markets. The descriptive study provided the conceptual and empirical map of the coffee value chain; the present paper turned that map into a dynamic model of the price-formation mechanism. Together, the two works support the central argument that coffee is best understood not as a homogeneous good with a single price driver, but as the joint outcome of biological cycles, supply-chain logistics, macroeconomic conditions and financial markets, whose interactions can be measured, ranked and interpreted with the appropriate econometric tools.

References

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control, 19*(6), 716–723. https://doi.org/10.1109/TAC.1974.1100705

Anders, S., & Fedoseeva, S. (2017). Quality, sourcing, and asymmetric exchange-rate pass-through into U.S. coffee imports. *Journal of Agricultural and Resource Economics, 42*(3), 372–385.

Avelino, J., Cristancho, M., Georgiou, S., Imbach, P., Aguilar, L., Bornemann, G., & Morales, C. (2015). The coffee rust crises in Colombia and Central America (2008–2013): Impacts, plausible causes and proposed solutions. *Food Security, 7*, 303–321. https://doi.org/10.1007/s12571-015-0446-9

Baumeister, C., & Hamilton, J. D. (2019). Structural interpretation of vector autoregressions with incomplete identification: Revisiting the role of oil supply and demand shocks. *American Economic Review, 109*(5), 1873–1910. https://doi.org/10.1257/aer.20151569

Benigno, G., di Giovanni, J., Groen, J. J. J., & Noble, A. I. (2022). *A new barometer of global supply chain pressures* (Federal Reserve Bank of New York Staff Reports No. 1017).

Bunn, C., Läderach, P., Ovalle Rivera, O., & Kirschke, D. (2015). A bitter cup: Climate change profile of global production of Arabica and Robusta coffee. *Climatic Change, 129*(1–2), 89–101. https://doi.org/10.1007/s10584-014-1306-x

Camargo, M. B. P. de. (2010). The impact of climatic variability and climate change on Arabica coffee crop in Brazil. *Bragantia, 69*(1), 239–247. https://doi.org/10.1590/S0006-87052010000100030

Capps, O., Cheng, M., Kee, J., & Priestley, S. L. (2022). A cross-sectional analysis of the demand for coffee in the United States. *Agribusiness, 38*(3), 689–704. https://doi.org/10.1002/agr.21744

Countryeconomy.com. (n.d.). *National minimum wage* [Data set]. Retrieved 2026, from https://countryeconomy.com/national-minimum-wage

DaMatta, F. M., & Ramalho, J. D. C. (2006). Impacts of drought and temperature stress on coffee physiology and production: A review. *Brazilian Journal of Plant Physiology, 18*(1), 55–81. https://doi.org/10.1590/S1677-04202006000100006

de Vega Moreno, M. (2026). *Coffee as a commodity: Market structure and price drivers* [Bachelor's thesis, Universidad Pontificia Comillas (ICADE)]. Madrid.

Dickey, D. A., & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association, 74*(366a), 427–431. https://doi.org/10.1080/01621459.1979.10482531

Durevall, D. (2018). How do coffee prices pass through to consumers? *Applied Economics Letters, 25*(13), 887–890. https://doi.org/10.1080/13504851.2017.1380281

FactSet. (2026). *Arabica coffee futures prices and exchange rates* [Data set]. FactSet Research Systems.

Federal Reserve Bank of New York. (n.d.). *Global Supply Chain Pressure Index (GSCPI)* [Data set]. Retrieved 2026, from https://www.newyorkfed.org/research/policy/gscpi

Federal Reserve Bank of St. Louis. (n.d.). *Federal Reserve Economic Data (FRED)* [Data set]. Retrieved 2026, from https://fred.stlouisfed.org

Gilbert, C. L. (2010). How to understand high food prices. *Journal of Agricultural Economics, 61*(2), 398–425. https://doi.org/10.1111/j.1477-9552.2010.00248.x

Global Data Lab. (n.d.). *Subnational climate and temperature data* [Data set]. Retrieved 2026, from https://globaldatalab.org

Gomes, A. P., Roma, C. M. da S., Almeida, J. C. F. de, & Souza, F. A. de. (2019). The futures market and price discovery in the Brazilian coffee market. *Custos e Agronegócio On Line, 15*(4), 200–230.

Hamilton, J. D. (1994). *Time series analysis*. Princeton University Press.

Humanitarian Data Exchange. (n.d.). *Rainfall datasets for Brazil, Colombia, Ethiopia, Indonesia and Vietnam* [Data set]. Retrieved 2026, from https://data.humdata.org

International Coffee Organization. (2023). *Coffee market report and global consumption data*. International Coffee Organization.

Kilian, L. (2009). Not all oil price shocks are alike: Disentangling demand and supply shocks in the crude oil market. *American Economic Review, 99*(3), 1053–1069. https://doi.org/10.1257/aer.99.3.1053

Kilian, L., & Lütkepohl, H. (2017). *Structural vector autoregressive analysis*. Cambridge University Press. https://doi.org/10.1017/9781108164818

Krishnan, S., Matsumoto, T., Nagai, C., Falconer, J., Shriner, S., Long, J., Medrano, J., & Vega, F. E. (2021). Vulnerability of coffee genetic resources. *Genetic Resources and Crop Evolution, 68*, 2691–2710. https://doi.org/10.1007/s10722-021-01179-4

Ljung, G. M., & Box, G. E. P. (1978). On a measure of lack of fit in time series models. *Biometrika, 65*(2), 297–303. https://doi.org/10.1093/biomet/65.2.297

Lütkepohl, H. (2005). *New introduction to multiple time series analysis*. Springer. https://doi.org/10.1007/978-3-540-27752-1

Nguyễn, T. N., Ngan, N. N., Hong, T. T., & Cuong, N. D. (2020). Hedging with commodity futures: Evidence from the coffee market. *Investment*

Management and Financial Innovations, 17(4), 1–13. https://doi.org/10.21511/imfi.17(4).2020.01

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics, 6*(2), 461–464. https://doi.org/10.1214/aos/1176344136

Sims, C. A. (1980). Macroeconomics and reality. *Econometrica, 48*(1), 1–48. https://doi.org/10.2307/1912017

Staver, C., Guharay, F., Monterroso, D., & Muschler, R. G. (2001). Designing pest-suppressive multistrata perennial crop systems: Shade-grown coffee in Central America. *Agroforestry Systems, 53*(2), 151–170. https://doi.org/10.1023/A:1013372403359

Trenberth, K. E. (1997). The definition of El Niño. *Bulletin of the American Meteorological Society, 78*(12), 2771–2777.

Volsi, B., Telles, T. S., Caldarelli, C. E., & Câmara, M. R. G. (2019). The dynamics of coffee production in Brazil. *PLoS ONE, 14*(7), e0219742. https://doi.org/10.1371/journal.pone.0219742

World Bank. (n.d.). *World development indicators: GDP per capita* [Data set]. Retrieved 2026, from https://data.worldbank.org