PricePedia

Methodology

January 2025

StudiaBo srl
via Santo Stefano 57, 40125 Bologna (ITALY)
tel. +39 051 5870353
C.F e P.iva: 03087661207
www.studiabo.it

Table of Contents

Introduction

This document describes the methodology used by StudiaBo to process the data available on the **PricePedia** platform. The cornerstone of the PricePedia project is prices, which come from various sources and serve to extract valuable information for companies in the process of purchasing production inputs.

Information

Sources

The data used on the PricePedia platform[1] mainly refers to:

- customs flows (import and export) of European Union (EU) countries;
- customs flows from the U.S. Census Bureau;
- customs flows from General Administration of Customs of the People's Republic of China;
- official international quotations;
- periodic surveys conducted by national statistical offices.

The customs flows from EU countries are sourced from Eurostat, which periodically publishes the monthly trade declarations of Member States, with a delay of about 2-3 months (e.g., in March 2023, data was published for the declarations of December 2022 and January 2023).
For the PricePedia project, customs flow data was chosen for the following reasons:

- global sourcing processes have become increasingly international, making European data more robust and representative than national data;
- the 8-digit Combined Nomenclature classifies over 9000 products, allowing for a high level of detail;
- data published by Eurostat is free and public, ensuring transparency and accountability.

International quotations come from major commodity exchanges, including:

- London Metal Exchange (LME);
- Chicago Mercantile Exchange (CME);
- London Bullion Market (LBMA);
- Intercontinental Exchange (ICE);
- Shanghai Futures Exchange (SHFE);
- Dalian Commodity Exchange (DCE);
- Zhengzhou Commodity Exchange (DCE).

Despite their different sources, the data on the PricePedia platform is highly correlated.

Financial Prices

Financial data refers to the time series of daily prices of futures traded on major international markets.
Futures are standardized forward contracts traded on a stock exchange, where the underlying asset is the purchase of a certain quantity of a commodity at an agreed price at the time of the contract. The price is therefore set at the time of the agreement based on current information and future price expectations, even though the actual exchange happens later. The price of the same commodity, traded later, may change significantly.
The price considered refers to the *settlement* value, which is the final price set at the end of each trading day for futures contracts. Since these contracts are made with irregular frequency and different maturities, for a medium-to-long-term analysis, **continuous contracts** have been used (the methodology for building continuous contracts is detailed in the *Methodology* chapter).
In PricePedia, when available, both individual futures contracts and continuous contracts with at least two maturities are published: **a near-term continuous contract**, which can be as soon as

the next day (referred to as *day-ahead*), by the end of the month (spot), or within the following month (1-month delivery), and **one or more continuous contracts with longer-term maturities**, which can range from a minimum of two months to a maximum of 3 years.

The Role of Financial Prices

Monitoring financial dynamics is particularly important in the context of interconnected global commodity markets.
Market participants may have different purposes. Typically, buyers use financial contracts to:

- hedge against speculative risks;
- insure against possible supply shortages;
- increase their bargaining power with producers.

From the perspective of certain financial operators, such as hedge funds, geopolitical events can create opportunities for speculative actions, which often have a significant impact on real market participants. The interconnection between financial and real markets frequently enables the transmission of price increases or decreases driven by speculation.
In PricePedia, financial data serves a **dual function**. First, it provides insights into international commodity price trends. The differentiation between continuous contracts with 1-month (or spot) and long-term maturities also provides insight into "expert opinion". The price differences between the two contracts may reveal two types of market situations:

- **Contango**: a normal situation where short-term futures prices are lower than long-term ones, reflecting costs such as storage;
- **Backwardation**: an unusual situation where higher spot prices signal demand pressure or short-term supply shortages. Another interpretation could be that experts expect future prices to decline, believing the short-term price is unsustainable in the medium-to-long term.

Financial data on PricePedia is also used to update customs prices, specifically for real-time updates. The specific usage methods are detailed in the **nowcasting** section.

Customs Prices

Customs data results from processing European customs flows, in terms of quantities and values, both for imports and exports.
In PricePedia, various customs price measures are available to meet different objectives:

- provide users with an **objective measure** of European commodity prices, useful as an index in indexed-price contracts;
- provide users with **up-to-date information** for negotiation purposes.

Different objectives and needs are pursued by the following customs price measures:

- **History** (EU, USA, and China), referring to the average price derived from customs transactions. The History EU measure is calculated using customs declarations and is typically two months behind the current month. This aligns perfectly with customs declarations, providing maximum **objectivity**;
- **Last Price EU**, referring to the price of contracts made by operators in the same month as the customs clearance. The Last Price addresses the issue that customs declarations for a given month reflect contracts (and thus market conditions) not only from the current month but also from previous months. Considering all customs declarations for a given month, the resulting price reflects both current and past market

conditions, creating a "stickiness" effect on prices. The Last Price is calculated by considering only customs declarations based on a contract price agreed in the same month the goods passed through customs. This way, the Last Price reflects the market conditions of the month in question, providing users with the **most up-to-date information on market conditions** available for negotiations;

- **Last Price EU Non-revised**, like Last Price EU, provides the most up-to-date data. Unlike the Last Price, it incorporates only the latest available information at each update without any revisions to previous observations.
- **Dispersion EU**, referring to the price dispersion in the reference market for a given commodity. The Dispersion measures, *Minimum* (corresponding to the first quartile of the price distribution, i.e., the value below which the lowest 25% of prices are found) and *Maximum* (corresponding to the third quartile of the price distribution, i.e., the value above which the highest 25% of prices are found), allow for the evaluation of the price dispersion of a commodity over time and serve as an auxiliary measure compared to those previously listed.

StudiaBo Indices

To provide a broader understanding of the trends for products available on PricePedia, synthetic indices have been constructed, aggregated by product types. These indices offer a robust representation of ongoing phenomena within a particular product type. Since these are relative measures, they allow for a comparison of the trends of various indices, even though the products within each type may have very different price levels.

- an index for each product type (energy, ferrous, etc.);
- an index for each product family (thermoplastics, copper, etc.);
- a total index representing all product types (Total Commodity);
- a total index excluding energy and food (Industrial).

These indices are constructed by aggregating the prices of individual products within each type/family, weighted based on EU trade flows in a base year (2019)[2].

Notes

1. One of the main differences lies in the product detail: while quotations refer to commodities with standard form, size, and weight, customs prices refer to very specific commodities and semi-finished goods (8-digit Combined Nomenclature).
2. The indices are constructed by assigning weights to individual products based on EU imports (both intra-EU and extra-EU) in the base year (2019).

PricePedia Methodology

Calculation of EU Customs Price measure History EU

Below is the methodology used by PricePedia to produce the time series of customs prices for History EU data.
The EU countries report trade flow data in terms of values (in euros) and quantities (in kilograms). These flows are reported at the level of product (more than 9000 items, according to the Combined Nomenclature classification) declaring country and partner country (for example, for each product, the import flow of Italy from France or the export flow of France to China, etc., is reported).
PricePedia estimates potential missing Eurostat data for quantities and calculates the price for each customs flow as the ratio between value and quantity.
Once the missing quantities data is estimated, the price for each customs flow is calculated as the ratio between values and quantities. These prices are then subjected to an outlier detection procedure, which consists of:

- elimination of declarations that appear to be anomalies;
- elimination of declarations of flows that do not occur frequently enough;
- identification and replacement of outliers at the time series level for each flow.

After detecting and correcting potential outliers in trade flows and prices data, the price distribution is calculated for each period, weighing the prices based on quantities. The time series of the customs price is built by considering the average of the prices from the central part of the distribution, that is, all prices between the first and third quartile, excluding the lowest 25% and the highest 25% of prices.

Nowcasting

Nowcasting is a statistical technique used to estimate the trend of a variable in real time, or with a slight delay from the present, and is based on statistical models that utilize already available data, combined with high-frequency indicators (for example, daily or weekly data) such as financial prices or other related economic factors (benchmark variable). This way, forecasts can be produced that help make decisions without having to wait for official data, which often arrives with delays.
In the context of updating monthly customs price series, the methodology used can be divided into three phases:

- study of the relationship between customs prices and benchmark variables;
- ECM and calculation of the "forecast in sample";
- application of the "constant adjustment" model.

Study of the Relationship Between Prices

The first phase requires a preliminary study of the price series, which consists of graphical analysis and statistical tests such as the Engle and Granger cointegration test and the Dickey-Fuller test on the order of integration of the considered series. Both graphical analysis and tests provide the necessary information for studying the relationship between financial and customs price series.

The ECM and in-sample forecast

After the preliminary analysis, the error correction model (ECM) is specified. This is one of the most robust models for studying the relationship between time series. The ECM is able to capture both short-term and long-term effects, allowing for the calculation of the so-called "adjustment speed," which determines how many periods are needed for the model to reach long-term equilibrium (target equilibrium).

Since prices are monetary measures, to reduce any scaling effects, variables transformed into natural logarithms are used for estimation.

After estimating the model with the OLS method, the "**in-sample forecast**" is carried out. In this phase, the value of the dependent variable (customs price) is reconstructed over the entire time span of the explanatory variables, based on the estimated short-term and long-term coefficients. The result is what can be defined as the in-sample forecast fit: the evolution of the "expected" price series based on the estimated relationship between the customs price and the benchmark variables.

Constant Adjustment

The percentage changes in the in-sample forecast fit are calculated for the periods between the last historical observation of the customs price and the current month (last 2 months). These changes are then applied to the "missing" periods of the customs price time series to update it to the current month.

Construction of continuous contracts for financial prices

Continuous contracts are constructed by merging data from individual futures contracts to generate a single continuous time series. These are based on **settlement** prices, which are the final prices set at the end of each trading day for futures contracts. Although the definition may vary from exchange to exchange, generally, the **settlement price** represents the weighted average value of trades made in the final part of the market session and is used by exchanges as the official reference for closing open positions and determining margins. When constructing continuous contracts, these settlement prices for each futures contract at expiration are used to ensure a smooth and accurate transition from one contract to the next during the rollover process.

The rollover process

When a futures contract approaches expiration, a rollover occurs, which is the transition from the expiring contract to the next contract with the subsequent expiration date. There are various methods to determine the rollover timing. In PricePedia, *temporal rollover* is used, which involves switching to the new contract on a predetermined date (usually the last trading day of the considered contract).

Some examples of continuous contract construction

- **1-month continuous contract**: A 1-month rolled continuous contract is based on monthly expiring futures contracts. For example, if we consider wheat and the contract expiring in November 2024, at the time of the rollover (usually just before the November expiration), the next contract, expiring in December 2024, becomes the new basis for calculation. At each monthly expiration, the next contract is taken as a reference, maintaining a continuous price series based on the settlement of each contract.
- **12-month continuous contract**: A 12-month continuous contract, on the other hand, is constructed using futures with annual or

quarterly expirations, depending on the commodity. In the case of oil, for instance, if starting from the contract expiring in December 2024, at the time of the rollover, the contract expiring in December 2025 will be used, and so on for the following years. This way, a time series is obtained that captures the evolution of the commodity price over long periods, useful for long-term analyses.

EU Producer Prices

The "EU Producer Prices" section collects the total industrial producer price indices (PPIs) elaborated by Eurostat (original data available at the *Eurostat webpage*) and updated on a monthly basis.
Industrial Producer Price Index (PPI) is an output index which measures price changes from the seller's perspective and are calculated as a weighted average of the relevant products for a specific economic activity, identified by the NACE classification.
PPIs are calculated by Eurostat for the domestic and the non-domestic markets. The combination of the two measures corresponds to the total PPI. Furthermore, PPIs take into account the following characteristics:

- quantity of units sold;
- transport provided;
- rebates;
- service conditions;
- guarantee conditions;
- destination;
- subsidies on products received by the producer, if there are any.

PPIs do not take into account:

- VAT and similar deductible taxes which are directly linked to turnover,
- duties and taxes on the goods and services invoiced

For further information please see the *Eurostat glossary on Producer Price Index*.

US Producer Prices

The Producer Price Index (PPI) for the United States measures wholesale inflation in the country. Published monthly by the US Bureau of Labor Statistics (BLS), the index tracks the average change over time in selling prices reported by American producers of goods and services.
In PricePedia, the US PPI data is organized into two classifications developed by the BLS based on information collected directly from American producers. These are:

- **Industry-based classification**, which organizes over 300 PPI time series by industrial sector, identified by the North American Industry Classification System (NAICS) code. These PPIs measure price changes reported by producers in a given industrial sector for goods sold outside the sector itself;
- **Commodity-based classification**, which groups over two thousand products based on their similarity or material composition, regardless of the producing company's industrial classification. This classification does not align with any other standard coding structure.

For further information please see the *BLS webpage on Producer Price Index*.

PricePedia Scenario

The PricePedia Scenario contains 24-month forecasts of purchasing material prices and key economic indicators.
The PricePedia forecast scenarios are based on **dynamic specification models**, which combine a foundation of economic theory and sector-specific expertise with statistical tools. The for-

mer helps understand the long-term dynamics between the forecasted price and its determinants, while the latter accounts for short-term fluctuations.

In addition to the models, **exogenous variables**, which are determined outside the model, are at the core of the scenario. Specifically, the different scenarios for exogenous variables are derived by PricePedia by referencing forecasts prepared by specialized institutions such as the *World Bank*, the *European Central Bank*, the *Energy Information Administration (EIA)*, and others. The main exogenous variables include the price of oil and the prices of the most significant commodities traded on financial markets, such as copper and aluminum. Exogenous variables also include exchange rates and inflation rates of the world's leading economies. The global industrial cycle also plays a crucial role in shaping future price scenarios. This variable is directly produced by PricePedia and is forecasted by referencing the "consensus" on the economic cycle of the world's major economies.

Once the reference scenario, determined by exogenous variables, is built, forecasting the price of a product requires identifying the econometric model, i.e., determining which variables explain its variations and estimating the values of the related parameters.

The degree of uncertainty in a scenario depends on two different sources. The first concerns the realization of the forecast scenario for the exogenous variables. The second relates to the quality of the estimated structural model. This generates two types of risks: one linked to exogenous variables and the other to the model.

In the PricePedia forecast scenario, to minimize the "exogenous" risk as much as possible, the baseline scenario is considered, i.e., the scenario that the various analyzed sources consider most likely. This means that the forecasted price of a single commodity represents the value with the highest probability of realization, conditional on the realization of the scenario for the exogenous variables.

Data Update

Time series

The information available on PricePedia is updated based on the type of source:

- The update of **financial source** data, typically daily, occurs every day, with the latest available data referring to the previous day;
- The update of **customs source** data takes place by the end of each month (or at most within the first 5 days of the following month). In the case of the *Last Price EU* and *Last Price EU Non-revised* measures, the latest available data refers to the month of the update, while for the *History* measure, the latest available data refers to the month of the last customs data declaration (corresponding to one or two months prior, depending on the source).

PricePedia Scenario

The PricePedia Scenario is updated monthly, typically during the week following the publication of the EU Customs section update.
Such an update frequency allows for the immediate detection of turning points occurring in commodity markets and the continuous refinement of structural models.