Economics of Metals in the Long Run: A Short Overview of the Academic Literature

BY MAYLIS PEYRET AND FRÉDÉRIC GONAND

This article provides a concise overview of economic analysis in metallic raw material production. It examines the key concerns of economists over the past century, their relevance in light of available data, and recent developments over the last two decades. The subject's relevance for economic policy is significant, particularly in understanding a market with volatile demand, sluggish supply, and instable prices, to which capital-intensive business models add further complexity. Understanding primary metal production patterns is crucial for assessing current metal demand sustainability amid the low-carbon transition and digital economy.

The article exclusively focuses on the economic analysis approach, excluding geopolitical and ESR considerations. It also focuses on the optimal extraction of scarce resources under maximized intertemporal utility, leaving the minority branch of research dealing with intergenerational equity unaddressed.

A long-term macroeconomic analysis framework is applied, considering short-term metal price variations to have, on average, subdued impacts on long-term trajectories according to available studies (e.g., Ulloa 2015).

A chronological structure is followed, covering Hotelling's model and its theoretical importance (1), taking stock of its empirical limitations (2), highlighting the restricted utility of peak models (3), exploring recent Cumulative Availability Curve approach developments (4), and providing insights on market models focused on short-term demand fluctuations (5).

1. Hotelling's analysis: a rich theoretical framework...

Hotelling (1931) studies the optimal behavior of a raw materials producer. He likens a natural resource production site to an asset whose yield has to correspond to that of the financial markets. Compared with a financial asset, however, a raw material deposit is unique in that it pays neither interest nor dividends. Consequently, its return can only be linked to an increase in the price of raw material extracted. For Hotelling, this price depends on the supply behavior of the producer, who chooses between producing today at the current price, or producing tomorrow at a higher price. Hotelling thus analyzes the supply of raw materials within an intertemporal framework.

The dynamic framework of Hotelling's model leads to consider the notion of opportunity cost. In standard economic theory, a company in a competitive market produces output until the marginal cost of production is equal to the market price. In the case of the extractive industries, producing an additional unit today reduces the available reserves of non-renewable resources for the future, and therefore the future production of raw materials.

As a result, if owners of non-renewable resources follow Hotelling's rule, i.e., they extract and sell these resources over time to maximize their net present value with respect to the interest rate, then they will extract the resource faster when the price rises due to its scarcity or a deterioration in the quality of future reserves and leave less resource for the future.

Therefore, on the optimal production path, and if opportunity and extraction costs are constant, the producer will only extract ore if the market price increases at a rate at least equal to the interest rate. This is Hotelling's rule in its simplest version, known in the literature as the r-percent rule (where the private discount rate is assimilated to the long-term interest rate r).

The intuition is that the discounted profit of a unit of resource extracted from the soil must be the same in all periods, there is therefore no gain in shifting extraction from one period to another. For the present value of the price (net of the extraction cost) to be the same in all periods, the undiscounted value must grow precisely at a rate equal to the interest rate. In this framework, if ex-ante demand is stable from one year to the next, production declines monotonically over time.

Because of the existence of this opportunity cost, which the market price must cover, the price of the raw material will always be higher than the marginal cost of extraction. Hotelling concludes that there is no risk of overexploitation of mining resources: a price higher than the marginal cost of production implies lower demand than in a standard market, where equilibrium is reached for a price equal to the marginal cost of production alone.

In terms of production profile, Hotelling predicts an asymmetrical bell-shaped trajectory, with an acceleration of production to a rapidly reached maximum, followed by a decrease in production rate.

2. ... though its empirical validity is often questionable

In the wake of Solow's (1974) remarkable article on Hotelling, numerous contributions appeared in the years that followed (e.g., Levhari and Liviatan (1977), Dasgupta and Heal (1979), Devarajan and Fisher (1981)). Some of them introduced extensions to the basic model, mainly along three themes: the dynamics of extraction costs (Herfindahl (1967), Heal (1976), Solow et Wan (1976), Weitzman (1976), Hartwick (1978), Slade (1982)), uncertainty (Stiglitz (1975), Gilbert (1979), Louy...
(1978), Pindyck (1979, 1980), and the consideration of risk (Copeland et al., 2005, Young et Ryan (1996)).

At this point, the reader has probably already understood that, if Hotelling's basic assumptions are lifted (fixed reserves, absence of technical progress, no uncertainty...), then the model's empirical predictions for price dynamics become heterogeneous.

In fact, empirical studies testing Hotelling's rule on real data have so far failed to produce a consensus. Lee et al. (2006) describe the price trajectory of non-renewable raw materials over the 1870-1990 period as "stationary around a deterministic trend with structural breaks." Farrow (1985), Heal and Barrow (1981), Tilton (1999) and Cuddington (2000) also fail to confirm the hypothesis of increasing resource prices underlying Hotelling's model. The applicability of Hotelling's model to real data overall raises significant difficulties (cf. Svedberg and Tilton (2006)).

3. Peak models, a more empirical approach with no theoretical basis or predictive gain

3.1. Hubbert's approach (1956)

King Hubbert, a Shell geologist in the 1950s, wrote a paper for a conference in Texas entitled "Nuclear energy and the fossil fuels", which concluded that only nuclear power could ensure the sustainability of the world's energy demand, and that it should therefore be substituted for fossil fuels.

This paper, often quoted but rarely read, has no theoretical basis, which is not a criticism but an observation. Hubbert notes that between the mid-nineteenth century and the mid-twentieth century, the growth rate of coal and oil production in the United States tended to decline over time. More specifically, he notes that the long-period profile of crude oil production in Ohio and Illinois exhibits roughly a bell-shaped profile over this period, with a production peak followed by a subsequent rapid slowdown.

Hubbert, who systematically assumed the stability of available resources, generalized, and considered the bell-shaped profile to be a natural feature of mining.

As a result, his work consists exclusively of estimating peak production and, more importantly, the associated depletion date for mineral resources, based on the current rate of production and the estimated size of reserves. The focus is exclusively on estimating available reserves, such that the cost of extraction, price, risk, rock quality, and technical progress are left unconsidered.

While Hotelling's model was not lacking in rich theoretical intuitions (but suffered from an inconclusive confrontation with real data), Hubbert's approach relies on a few empirical cases, a somewhat dubious generalization, and a proven lack of theoretical construction.

3.2. Peak models

The peak models developed in the wake of Hubbert’s work have enjoyed relative success in the literature. In these models, there is only one input that defines peak production: the “Ultimate Recoverable Resources” (URR) that define the total supply over time. URR is an assumed estimate of the total mineral resources an economy can recover from mineral deposits, now and in the future (Prior et al., 2012).

Peak models explicitly assume that other determinants of supply (price, technology, exploration, or production costs) are irrelevant for studying the long-term depletion of non-renewable resources (Tilton, 2018). The quantity demanded in peak models is not a relevant variable if it is greater than or equal to the production of the peak function. This demand condition is implicitly guaranteed by non-decreasing per capita demand. All these assumptions seem very strong, and rather unreasonable.

Peak model calibrations consider different URR scenarios, but changing the URR does not lead to major changes in the peak year (Northey et al. (2014), Sverdrup et al. (2014)), which may provide an impression of robustness. In the case of peak models applied to copper, the literature of the last fifteen years has agreed on a shortage over the next 20 to 30 years (Bardi and Pagani (2007), Prior et al. (2012), Laherrère (2010), Northey et al. (2014), Sverdrup et al. (2014)) across heterogenous URR assumptions.

3.3. Serious criticism

Criticisms of bell-shaped models have been widely debated:

- These models often confuse geological availability with economic availability. The uncertainty of economically available geological stocks is a fact, yet it does not affect the behavior of agents in peak models that consider reserves and resources as fixed stocks (May et al. (2012), Meinert et al. (2016), Wellmer and Scholz (2018)).
- Furthermore, peak models do not consider the effect of technology, which increases the economic availability of reserves, resources, and undiscovered deposits (Kharitonova et al., 2013).
- Peak models often fail to consider the fact that the intensity of use of metallic materials declines as countries develop (Criqui (2013), Crowson (2011), Ericsson and Söderholm (2013)).

All things considered, it is possible to fear that the assumptions of peak models are highly questionable and undoubtedly biased in favor of a pessimistic forecast of the depletion of metallic mineral resources.

In the case of copper, for example, it is a constant that the resource is abundant and that the reserve is being maintained. In 2018, the US Geological Survey (USGS) inventoried the Reserves/Production ratio expressed in number of years of consumption (since “reserves” are a priori a stock while production is an annual flow) as calculated since the beginning of the 20th century. In 120 years of statistics, this ratio has always been relatively constant, fluctuating around 40 years of consumption.
4. A new paradigm? The Cumulative Availability Curve approach

Between Hotelling-style thinking, which employs concepts from economic analysis but suffers from a clear lack of confirmation in the data, and a highly empirical Hubbert-style approach that has no theoretical foundation and no gain in predictive power, is a third way possible?

The Cumulative Availability Curve (CAC) of an exhaustible natural resource is the graph of the function that relates a given price of this resource to the total world stock economically exploitable at this price. This CAC differs from the traditional supply curve in economics textbooks, which describes the flow of goods offered on the market for a given period (usually one year) as a function of price. The CAC corresponds not to a flow over a given period, but to a global stock available for the future. It shows the total quantity of natural resource recoverable in the economic sense of the term as a function of the price level (Tilton and Lagos (2007), Tilton et al. (2018)). However, like the traditional supply curve, the cumulative availability curve (CAC) assumes that, apart from price, all other determinants of metal availability are fixed (exploration and production costs, technological level).

The CAC approach is interesting for prospective exercises on the sustainability of metal demand. Indeed, the shape of the curve depends on geological factors that have occurred in the past, and on events that may or may not occur in the future: it can therefore be traced relatively objectively.

The combined calculations of CAC and global demand trends have led to the reasonable conclusion— with all due caution when it comes to projections—that global lithium demand should remain sustainable over the century, even with optimistic demand and conservative supply assumptions (Yaksic and Tilton, 2009). Once again, caution is called for in this kind of exercise, as geology and extraction techniques can sometimes lead to major surprises.

However, the CAC paradigm for assessing the sustainability of global demand for metals does not enjoy complete consensus on how to assess mineral resource depletion.

For some, the ability of markets to provide the necessary signals to compensate for resource depletion is not assured. High external social and environmental costs of mining are not internalized by markets (Segura-Salazar and Tavares, 2018). Price trends do not appear to signal mineral resource depletion, as price trajectories do not clearly differ between geologically abundant and scarcer minerals (Henckens et al., 2016).

Other critics argue that the opportunity cost paradigm may overestimate the role of technology in offsetting depletion (Gordon et al., 2007; Humphreys, 2013).

On a more fundamental aspect, we find two methodological limitations to the CAC approach. Firstly, the CAC is a purely accounting method - not an economic one, i.e., it does not include maximization behavior like Hotelling’s model. Secondly, the CAC approach is a partial equilibrium analysis, not a general equilibrium one.

The gradual depletion of mineral resources is assumed to drive up prices, curb demand, increase substitution, promote recycling, and encourage new sources of supply made possible by technology (carbon nanotubes, etc.). The CAC method does not include any price loop effect, where demand growth would be held back by soaring prices. This is a potentially important channel for analyzing the sustainability of global demand for a metal.

5. Market models and short-term price variations

The models of Hotelling, Hubbert, and their heirs generally did not consider metal demand as an explanatory factor for the price profile of the resource. Thanks to new econometric and statistical tools, the correlation between short-term phenomena, often but not exclusively linked to demand shocks, and long-term dynamics has enjoyed renewed interest in the literature since the 2000s, in the wake of the significant rebound in commodity prices observed at the turn of the century.

The first branch of this literature studies price cycles by breaking them down into transitory and permanent components. In general, this literature confirms the existence of price cycles affecting all commodities, while transitory shocks affect different commodities differently. Metal prices in particular are significantly influenced by short-term cyclical shocks.

The second branch focuses on the drivers of commodity prices, breaking down price changes into aggregate demand, commodity-specific demand, and commodity-specific supply shocks. Most of these studies concern oil prices. The literature on the drivers of metal prices is less abundant, but there is greater agreement that aggregate demand is the main determinant of short-term metal price shocks.

5.1. Price cycle models

Research into the existence of price cycles common to several commodity groups only really developed in the early 2000s, in the wake of the 60% surge in energy commodity prices between 1998 and 2001. This literature generally breaks down price movements into transitory and permanent components. This includes short-term cycles (business cycles), medium-term cycles (8 to 20 years) and possible “supercycles”, which concern many commodities and last several decades. Short- and medium-term cycles are fueled by transitory shocks that can have several origins: recessions (e.g., the global financial crisis of 2007-2009), accidents (e.g., Vale’s accident in Brazil in 2019, which disrupted iron ore supplies), conflicts or terrorist attacks.

For metals, the cyclical component of shocks accounts for a much larger share of their volatility than for other commodities: the variance is twice as high for metal prices as for those of energy and agricultural goods (Baffes and Kabundi, 2023).

5.2. Drivers of prices

The literature studying the drivers of commodity price shocks generally relies on the seminal study by
Kilian (2009) and his Structural Vector Autoregressive (SVAR) econometric model with sign restrictions to identify the relative importance of different shock drivers. Using data on commodity prices, demand and supply, price shocks are decomposed into aggregate demand shocks, commodity-specific supply shocks and commodity-specific demand shocks.

Global shocks to world demand include worldwide recessions (such as that associated with the 2008-09 global financial crisis) or pronounced expansions linked, for example, to industrialization or urbanization (e.g., China in the years 2000-2010). Commodity-specific supply shocks include accidents, strikes, conflicts, cartel production decisions, government policies and weather events.

Commodity-specific demand shocks are generally considered as a residual component of the SVAR model and reflect the influence of inventories (resulting from government stockpiling, producer stocks and market purchases), technological changes, shifts in consumer preferences, and government policies (e.g., carbon tax).

Stuermer (2018) and Jacks and Stuermer (2020) suggest that, in the case of metals and unlike hydrocarbons, aggregate demand shocks and commodity-specific demand shocks play a more sensitive role than supply shocks, and that their impact has increased over time.

Beyond the VAR approach, recent literature confirms that, on average, demand shocks have relatively little effect on long-term price trends. Thus, Ulloa (2015) shows through unit root tests conducted on numerous time series that, for copper, demand shocks affect only short-term price movements. Similarly, Wets and Rios (2015) model copper prices using a structural model that separates short- and medium-long-term dynamics and conclude by mentioning that their approach “should be applicable to a wide range of commodities”. However, since 2015, no studies applying the Wets and Rios (2015) model to other metals have emerged, probably due to a lack of data, either in terms of price or production time series length, or reliability.

** **

Today, there are two main ways of studying the depletion of mineral resources and the sustainability of global demand for metals. The fixed-stock paradigm used by peak models assumes that the supply of metals is predefined and intangible: from this, the life of reserves is deduced according to future demand scenarios. This seemingly logical approach runs into serious methodological difficulties. It ignores prices and costs, technical progress, and recycling, and fails to consider that physical reserves that are available may not be effectively exploitable in economic terms.

The other approach to the sustainability of world demand for metals takes a more economic approach, with prices playing a central role the so-called CAC approach. This approach studies changes over time in what a company is prepared to pay for an additional ton of metal, depending on the geological resource and the economic conditions under which it can be mined.

Market models are used to study, often econometrically, short-term variations in metal prices. Theoretical and statistical approaches suggest that their effects on medium- to long-term prices remain to be proven.

The future in this field will probably continue to reflect on the one hand the effects of depletion of mineral reserves, which influence the shape of the CAC curve and the speed at which the world economy moves along it; and, on the other, the effects of technological progress, which reduce extraction costs. In this respect, Hotelling had the right intuitions, but had not necessarily modeled them in the most effective way to study the sustainability of metal demand.

**Bibliography**


Footnotes

1 Pindyck (1978b) argues in favor of replacing the word “exhaustible” with “non-renewable”, since the concepts of reserves and their depletion are ultimately economic rather than geological or physical notions. This is where a strong tension arises between economists and geologists in their mode of reasoning, which we shall return to later: the former are more likely to consider that exploitable reserves of primary metal are not so much fixed by nature as variable according to various economic parameters.

2 Demand is not considered in Hotelling’s intertemporal modeling: the producer observes a price based on market conditions (raw material stock and discount rate) and adjusts his extraction rate based on these parameters alone. This approach is justified by the assumption that short-term market fluctuations (linked to the interaction between supply and demand) do not significantly affect the net value of the resource over the long term (see section 5).

3 Aluminum, coal, copper, iron, lead, natural gas, nickel, oil, silver, tin, and zinc.

4 The CAC gives no indication of the speed with which the global economy is consuming available stocks to the point of exhaustion.