The Hotelling Rule in Non-Renewable Resource Economics: A Reassessment

Roberto P. Ferreira da Cunha, Antoine Missemer

THE HOTELLING RULE IN NON-RENEWABLE RESOURCE ECONOMICS: A REASSESSMENT

-

Roberto P. FERREIRA DA CUNHA[*]

Antoine MISSEMER[†]

-

The pagination of the published version is indicated in the margin.

-

Abstract

Harold Hotelling's 1931 contribution is known for providing a basic principle—the Hotelling rule—to the economics of non-renewable resources. Nearly 90 years later, empirical tests conclude the rule lacks empirical validity, requiring strong amendments to describe the long-term, aggregate behaviour of its target object. On the basis of Hotelling's unpublished archival material, this paper revisits the place given to the Hotelling rule in non-renewable resource economics. Our reconstruction shows that Hotelling's 1931 paper has been misinterpreted: from the outset, the Hotelling rule was not valid for mineral resources. In contrast, the consideration of two inherent geological constraints, alongside exhaustibility, offered the opportunity for an alternative basic framework, capable to generate bell-shaped and U-shaped equilibrium trajectories for supplies and prices, respectively. Inspired by this unknown aspect of Hotelling's work brought to light by our archival investigation, we sketch this alternative basic model, enabling non-renewable resource economics to circumvent the empirical shortfalls of the Hotelling rule.

[*] [In 2020] Berkeley Research Group, LLC. Rua da Quitanda 86, 2° andar – Centro, Rio de Janeiro, RJ, Brazil. E-mail: rcunha@thinkbrg.com
[†] [In 2020] CNRS, CIRED Paris – Centre international de recherche sur l'environnement et le développement (UMR 8568 CNRS/ENPC/EHESS/AgroParisTech-University-of-Paris-Saclay/CIRAD), 45 bis av. de la Belle Gabrielle, 94736 Nogent-sur-Marne Cedex, France. E-mail: missemer@centre-cired.fr

1. Introduction

"The Economics of Exhaustible Resources," published in April 1931, is considered a seminal contribution for non-renewable resource economics. Harold Hotelling proposed an extensive analysis of the intertemporal optimization of extraction paths under various competitive conditions. He laid the foundations for many research questions, including the role of taxation and public regulation in extractive sectors. Even if the Hotelling model is no longer taught to students in the original form, the lessons drawn from the 1931 article are still at the ground of contemporary research and teaching, in which the equation $p_t = p_0 \cdot e^{r \cdot t}$ plays the role of basic principle.

In 1974, Robert M. Solow paid tribute to Hotelling, labelling this equation as "the Hotelling rule, the fundamental principle of natural-resource economics" (Solow 1974, p. 12). As far as we know, that was the first occurrence of the expression "Hotelling rule," and Solow's impulse, within the favourable context of the oil price shock, reignited research on the subject of non-renewable resources (e.g., Levhari and Liviatan 1977, Dasgupta and Heal 1979, Devarajan and Fisher 1981, Farzin 1984, 1992). Since then, the literature retains *finiteness*, in the form of Hotelling's hard constraint $\int_0^\infty q \cdot dt \leq a$, as the central defining factor of the class of non-renewable natural resources.

Tests on the empirical validity of the Hotelling rule, by itself, have concluded it is unable to stick to reality, both in terms of assumptions and capacity to predict real market prices. Gaudet (2007), analyzing the price evolution of 10 mineral resources throughout the 20th century, concluded the rate of change in prices is approximately centred at zero, and not at the interest rate. Livernois (2009) reviewed 34 empirical studies on the Hotelling rule, concluding data does not provide overwhelming support for it. Slade and Thille (2009), reviewing roughly the same studies, concluded for the rejection of the basic model by data. Further empirical tests, of specific aspects of Hotelling's model, or of elementary extensions of the basic model, reached the same conclusion (e.g., Hart and Spiro 2011, Gaugler 2015, Atewamba and

Nkuiya 2017, Karp 2017).

The rejection of the basic model by data has been remediated by the consideration of additional real-world factors. The main ones capable of bringing the basic model closer to reality are technological progress (Slade 1982), degradation costs (Pindyck 1978), durability (Levhari and Pindyck 1981), market structure (Salant 1976) and uncertainty (Kemp 1976).[1] Each extension inaugurated a modern branch in non-renewable resource economics, expanding and detailing the level of sophistication of each factor within the original basic framework. Nonetheless, the implementation of different additional factors leads to questions about the robustness and readability of the basic theory: by varying the choice of factors to implement on the model, one can attain diverging results. By requiring strong amendments to represent its target object, its value as a theoretical starting-point comes into question.

This paper inquires whether a deeper understanding of Hotelling's 1931 contribution can help us circumvent the empirical invalidity of the Hotelling rule. We hypothesize the prevailing understanding stems from a problematic interpretation of the original contribution, inaugurated through Solow's 1974 labelling of Hotelling's basic result as a rule. To test this hypothesis, we investigated Hotelling's project on non-renewable resources through the history of economic thought methodology, allowing us to leverage unpublished materials towards demonstrating potential theoretical gaps left behind along the evolution of thought, possibly changing the way we understand and use theory going forward.[2]

We based our investigation on an unprecedented exploration of Hotelling's archives, stored at Columbia University (hereafter HHP).[3] These materials seem to have

---

[1] Many other extensions of Hotelling's model have been proposed in the literature over the past decades. For the most recent years, we can notably cite Holland (2008), Venables (2014), Okullo et al. (2015) and Anderson et al. (2018). This list is not exhaustive.

[2] Previous works in the field of history of economic thought referring to non-renewable resources include Robinson (1980, 1989), Kula (1998) and Missemer (2017, 2018).

[3] HHP: Harold Hotelling Papers, Rare Book and Manuscript Library, Columbia University, USA. Catalogue available at findingaids.cul.columbia.edu/ead/nnc-rb/ldpd_4078401.

been consulted only by a few other scholars for previous research (e.g., Crabbé 1986; Darnell 1988, 1990; Hands and Mirowski 1998), but none of these works offer a comprehensive view on Hotelling's project regarding the economics of non-renewable resources. The documents stored at Columbia University consist of several thousand sheets, filed in 58 boxes. Items related to exhaustible resources are scattered in different boxes. Thanks to a systematic review, we found about 20 to 30 significant unpublished drafts, preparatory documents and letters from 1924 to 1925 onwards worthy of interest for our inquiry.[4]

Our paper is organized as follows. In section 2, we investigate what sort of research Hotelling was conducting in the 1920s, showing his hesitation between a theoretical abstract ambition and empirical concerns for what he named at the time *exhaustible* resources.[5] This section allows a clearer picture on the area of validity, from the outset, for his basic result— which is different than what has been retained from his article. In section 3, we closely review Hotelling's work on geological constraints for his 1931 article. We reveal a neglected fundamental role for them in his analysis. Hotelling considered cumulative production $x(t)$ and the change in production $q'(t)$ as fundamental variables, and not additional factors, constituting a potential alternative basic model capable of attaining U-shaped trajectory for prices—the now empirically accepted stylized fact for non-renewable resources. In section 4 we provide our concluding remarks, reassessing the Hotelling rule as the starting point of mineral resource analysis.

2. Hotelling's research on exhaustible resources in the 1920s

2.1. From (natural) resources to assets, and vice versa

Until the submission of his paper to the *Journal of Political Economy*, most likely in the fall of 1930,[6] Hotelling's archives reveal, regarding exhaustible resources, that he actually worked along two approximately parallel lines of research: a generic project on the valuation of exhaustible *assets*, in the line of his theory of depreciation, and a specific project on exhaustible *natural resources* (minerals, fossil fuels), more empirically grounded. In his final article, Hotelling uses several terms to describe his research: "exhaustible (natural) resources," "exhaustible assets," "irreplaceable (natural) resources," "irreplaceable assets," "mineral content," "resources of the earth," etc. By definition, the set of "natural resources" here contains both fossil energy sources (coal, oil, gas) and minerals. The set of "assets" is broader, containing "natural resources" as a subset. A mineral stock is an asset. A financial bond is also an asset, but obviously not a natural resource. Focusing on "natural resources" suggests that the precise subject of research is related to concrete forms of natural items. On the opposite side, focusing on "assets" opens the reflection to other case studies, beyond natural items.[7]

Hotelling's drafts, correspondence and notes throughout the 1920s confirm the co-existence of two levels of analysis: a *generic* level, illustrated by the term "exhaustible assets," and a more

---

[4] This paper is part of the research project "Bifurcations in Natural Resource Economics" (BNRE), sponsored by the European Society for the History of Economic Thought (ESHET). The objective has been to explore the genesis of modern natural resource economics, notably through the creation of a digital database of Hotelling's economic archives (7000+ images taken of 3000+ documents). The project team mobilized researchers in natural resource economics and history of economic thought over 2.5 years, producing to date several coordinated articles and working papers on complementary aspects related to Hotelling's research (e.g., Franco et al. 2019, Gaspard and Missemer 2019, Missemer et al. 2020, Missemer and Nadaud 2020).

[5] What Hotelling reffered to as "exhaustible resources" are now called "non-renewable resources." In the following, we use the modern term, except when it comes to evoking and analyzing Hotelling's project in its historical context.

[6] Two letters from Hotelling to R. A. Fisher and Chamberlin (dated July 23, 1930, and Aug. 8, 1930, respectively) indicate that Hotelling was working on the final version of the paper in the summer of 1930 (RAFP, digitized version online; Guicherd 2017, p. 179). A letter from Mills (dated Jan. 17, 1931) shows that the final version of the paper was already circulating at the very end of 1930 (HHP, Box 1, Mills, Frederick C.). Arrow (1974, p. 1103; 1980; 1987, p. 670) reports that Hotelling (who was his PhD supervisor) submitted his paper to the *Economic Journal* before submission to the *Journal of Political Economy*. As shown in Gaspard and Missemer (2019), there is, however, no clear evidence that he did so.

[7] This statement is not retrospective. In the early 20th century, Irving Fisher (1906) had already discussed the *generic* nature of the words "assets" and (not necessarily natural) "resources."

specific level, with the term "exhaustible natural resources." We found Hotelling hesitated between which of the two to use in his final paper. Until only a few weeks before submission, the title of the paper was "The Economics of Exhaustible Assets," as indicated in a letter to the statistician Ronald A. Fisher:[8]

> The book [on statistics] is gradually progressing, though I am leaving it alone for a few weeks at present to finish a paper on "The Economics of Exhaustible assets" which I have been at for five years and for which various people have been calling.

Only in January 1931, four months before publication, had the final title been chosen, as another letter from Fredrick C. Mills (NBER) testifies to:[9]

> Schultz was kind enough to let me read a copy of your paper "The Economics of Exhaustible resources." I returned this to him. It seemed to me to be an extremely suggestive attack upon a problem which economists have largely neglected.

Hotelling therefore finally chose an intermediary title, aiming at covering *both* his generic, theoretical ambition and his specific, more concrete concerns. The archives further reveal that from 1924/1925 to the final version, Hotelling constantly changed the title of his research. A first draft from 1924 was entitled "Exhaustible Assets under Competition." [10] Another draft from 1925 was labelled "The Exploitation of Irreplaceable Resources." [11] In December 1929, a full abstract was named "The Economics of Exhaustible Assets."[12]

The contents of these drafts are not exactly the same. When assets are considered, the analysis makes a few mentions of mineral resources or concrete questions, instead focusing

on the use of advanced mathematics to deal with abstract principles.[13] When natural resources, or minerals, are scrutinized, references to real-world situations are more present, as shown by the 1925 draft entitled "The Exploitation of Irreplaceable Resources":[14]

> The complete monopoly and the perfectly fluid competition of traditional economics are extremes between which lie all the cases which actually occur in the world.

The generic subject (exhaustible assets) offered opportunities to draw theoretical principles, such as the equation $p_t = p_0 \cdot e^{r \cdot t}$ (i.e., the Hotelling rule). The specific subject (exhaustible natural resources), a particular case of exhaustible assets, required more attention to the real world.

The chronology of the archives helps us to reconstruct Hotelling's early treatment of exhaustible resources. The first draft on the subject dates back to 1924; it was written for a meeting of the American Mathematical Society in Chicago. These notes outline the close connection between Hotelling's research on depreciation and on exhaustible assets, as already mentioned. [15] The work on asset valuation was published in 1925 under the title "A General Mathematical Theory of Depreciation." As noted by Crabbé (1986), assuming Hotelling's asset valuation function from Hotelling (1925) is a negative exponential with a constant parameter—the rate of interest—yields the Hotelling rule published in 1931.

We found that it was in the winter of 1925 that Hotelling considered for the first time engaging with dedicated research on exhaustible assets. It is also when he decided to explore in more detail the calculus of variations to "simplify" the issue:[16]

---

[8] Letter from Hotelling to Fisher, July 23, 1930. Online digitized version. RAFP: Ronald A. Fisher Papers, Special Collections Library, University of Adelaide, Australia. Catalogue available at digital.library.adelaide.edu.au/dspace/handle/2440/3860.

[9] Letter from Mills to Hotelling, Jan. 17, 1931. HHP, Box 1, Mills, Frederick C.

[10] HHP, Box 10, AMS Reports and Correspondence (3).

[11] HHP, Box 42, Exploitation of Irreplaceable Assets.

[12] HHP, Box 42, Exploitation of Irreplaceable Assets.

[13] For instance, in the draft "Exhaustible Assets under Competition" from 1924, Hotelling states his problem as "applying Euler's equation." HHP, Box 10, AMS Reports and Correspondence (3).

[14] HHP, Box 42, Exploitation of Irreplaceable Assets.

[15] HHP, Box 10, AMS Reports and Correspondence. For a full inquiry into the intertwining of Hotelling's projects on depreciation and exhaustible assets, see Missemer et al. (2020).

[16] Monthly Report from Hotelling to the Directors of the FRI. Mar. 2, 1925. HHP, Box 41, Agriculture III.

The recent work on the economic of irreplaceable assets had during the past month been undergoing revision and extension. The revision is made necessary by a simplification of the Calculus of variations method which is found to be applicable to the problem, and also by the publication of a paper by G. C. Evans (Proc. Nat. Acad. Sci., Jan. 1925, pp. 90-95) touching upon the subject.

It appears that Hotelling clearly regarded his generic project—i.e., the theory of exhaustible *assets*—as an extension of his theory of depreciation. Interestingly, another draft from the same time, March 1925, indicates that he had meanwhile started to investigate the more concrete and real-world side of the problem of exhaustible resources, with special attention to minerals, oil and natural gas in the case of an oligopoly. This confirms not only the entanglement of the two research lines but also their synchrony:[17]

> The problem is inderteminate because after one mine is exhausted the other has a monopoly; and since we have assumed the demand inelastic the profits of the monopoly are unlimited.

In summary, Hotelling's economics of exhaustible resources in the 1920s was composed of two levels of analysis. Because the 1931 article mixed the two levels, it became difficult for readers to identify the purposes and limits of the different arguments, including the area of validity of the Hotelling rule. The archival material helps provide a clearer understanding: the apparent discrepancy between Hotelling's concrete subject (non-renewable resource depletion) and abstract proposals (e.g., the Hotelling rule) is due to the existence of two initially distinct lines of work developed in parallel, and later merged in the same contribution. For the treatment of real-world resources, Hotelling needed more information on the actual dynamics and issues pertaining to extractive sectors. Archival materials helped us identify where he found this information.

---

[17] HHP, Box 42, Exploitation of Irreplaceable Assets.

2.2. Conservation (or not) and energy policy

Gaudet (2007) states Hotelling wrote his 1931 paper in reaction "to the demand for regulation of the exploitation of exhaustible resources by the conservation movement, which had been particularly strong in the U.S. during the period 1890 to 1920" (p. 1034). Livernois (2009) asserts, "Hotelling's seminal article was a response to the conservation movement's claims that unregulated private markets would lead to the overexploitation of non-renewable resource stocks" (p. 23). Their conclusion is likely based on the opening paragraph of "The Economics of Exhaustible Resources," where Hotelling textually frames his research scope in relation to the "conservation movement" (p. 137). However, he mentions the movement only one more time (p. 143), in section 3.

The reference to conservation is intriguing because, in the 1920s, among economists, the movement was not as popular as in the 1900s and 1910s, when Lewis C. Gray (1913, 1914) and Richard T. Ely (1918) wrote on the subject.[18] Of course it was still an important trend in the reflections about natural resources (e.g., Ise 1925). However, the importance given by Hotelling to the conservation movement as a counterpoint to his analysis and as a source of information about on-going issues in extractive sectors is not clear at all and deserves to be examined in detail.

The archives reveal that early drafts make no mention of the conservation movement, even along the specific, more empirical research line. In a preliminary introduction of the paper, in December 1929, and which was very close to the final version, the catchphrase is already written (apart from the last addition on taxation), but the part on conservation is missing. Instead, Hotelling insists on competitive settings and on exhaustible assets as a specific case of a wider issue, in line with his previous drafts starting in

---

[18] On the economic thought of the first conservation movement, see for instance Smith (1982), Ramos Gorostiza (2003), Kula (1998) and Missemer (2017). Gray worked on conservation when he was at the University of Saskatchewan (Canada), initiating a long tradition of Canadian research in resource and environmental economics (see Copeland and Taylor 2017). On Gray's contributions, see Crabbé (1983).

1924/1925:[19]

> Contemplation of the world's disappearing supplies of minerals, forests, and other exhaustible assets has led to demands for regulation of their exploitation, *by taxation or otherwise. On the other hand the existence of monopolies and quasi-monopolies has raised the same problems of excessively curtailed production here as in other industries, but in a more complex form. The economics of the exploitation, that is, the study of the rates of exploitation which on the one hand tend to take place under commercial conditions, and of those which on the other hand ought in the public interest to take place, has however received no serious attention* [our emphasis].

This provisional introduction reveals that conservation was not so directly decisive in Hotelling's undertaking.

If not in the conservation literature (especially the economic literature), where did Hotelling find information about the concrete dynamics of extractive sectors (minerals, oil and gas)? Several allusions disseminated in the 1931 article find justification in issues of the 1920s not directly related to conservation:

> The prohibitions against oil and mineral development and cutting timber on certain government lands have this justification, as have also closed seasons for fish and game and statutes forbidding certain highly efficient means of catching fish (p. 137).

> On the other hand, certain technical conditions most pronounced in the oil industry lead to great wastes of material and to expensive competitive drilling, losses which may be reduced by systems of control which involve delay in production (p. 138).

> The government of the United States under the present administration has withdrawn oil lands from entry in order to conserve this asset, and has also taken steps toward prosecuting a group of California oil companies for conspiring to maintain unduly high prices, thus restricting production (p. 138).

Daniel Yergin's 1991 work on oil history

illuminates the events Hotelling refers to. The 1920s in the United States were marked by fast growing oil demand brought by gasoline-fuelled cars (Yergin 1991, p. 208). Meanwhile, very few new oil discoveries had been registered on the 1917–1920 period, leading oil prices to triple in the period. Within this context took place in the early 1920s the "Teapot Dome Scandal." This episode of corruption was about the exploitation of several oil fields in Wyoming initially preserved by the US government as naval oil reserves under the Taft and Wilson administrations (Yergin 1991, p. 211). In 1924, when Hotelling started to work on exhaustible resources, the presidential election was marked by the scandal. The mention of government ownership over lands and resources in the 1931 paper (above quotes) matches this episode.

Hotelling's allusion to technical conditions leading to waste can also be related to legal debates in the 1920s. The so-called "Rule of Capture" attributed ownership over oil resources to the owners of the land immediately above it. When several properties spanned a single reservoir, operators would rush to produce as rapidly as possible in order to avoid letting the neighbours drain the volumes under their properties through their wells. This ultimately led to an excessive number of wells and lowered the ultimate recovery of oil from the reservoir. Debates took place mostly under the Coolidge administration (1923–1929), leading to the establishment of the Federal Oil Conservation Board, in 1924 (Yergin 1991, pp. 220–23).

Therefore, the remarks on minerals, oil and gas disseminated by Hotelling in his final article, not fully related to the academic debates on conservation, actually correspond to various affairs linked to the US federal policy. His allusions to practical cases were not thought experiments, or pretexts to explore fictive situations; they were based on real political and economic experiences that marked American public life in the 1920s.

In this context, the Hotelling rule, although essential for the generic project on exhaustible *assets*, and which supposed perfect competition and no State intervention for achieving the social optimum, was clearly established as inadequate for the analysis of concrete, specific *natural resources* (minerals, oil and gas), as

---

[19] HHP, Box 42, Exploitation of Irreplaceable Assets.

demonstrated by the closing of section 2:

> ...[T]here are in extractive industries *discrepancies* from our assumed conditions leading to particularly wasteful forms of exploitation which might well be regulated in the public interest (pp. 143–44; our emphasis).

The word "discrepancies" is significant as a strong indicator of the inoperability of the basic model for describing concrete situations. As soon as these situations are involved, the "real economic world" (Hotelling 1931, p. 171) matters, up to assigning a role to regulators in the extractive sector, in contrast to what is usually retained from Hotelling's analysis in terms of State intervention.[20] Theoretically speaking, empirical realities require further developments beyond the basic result.

An insight emerges from the realization that Hotelling did not see what Solow (1974) labelled "the Hotelling rule" as a rule for non-renewable resource economics because it could not describe what took place in reality in extractive sectors, as Hotelling found in concrete experiences from the 1920s. The area of validity of the rule was limited to finitely available assets, in the *absence* of some realities he proceeded to analyze. Among these realities, apart from alternate market structures and taxation issues, we discovered Hotelling paid close attention to geological constraints, which we will show he concluded to be *necessary conditions* to depict natural resources.

3. Hotelling and geological contraints

3.1. Geological constraints in the 1931 article

"The Economics of Exhaustible Resources" is known for its basic model, which is described in sections 2 and 3 of the 1931 article. The final paper, nonetheless, contains 15 sections. Archive research revealed the basic result dates back to the first drafts, circa 1924/1925. Subsequent drafts deal with more substantial models, which one can find in the less-known sections of the

1931 article. Slade and Thille (2009) acknowledge that "Hotelling derived *several variants* of his model" (p. 241; our emphasis). These "variants" actually represented the core of Hotelling's work, at least as reflected in the archives.

From sections 8 to the end of his analysis, Hotelling deals with his specific research line on concrete mineral resources, directing his attention to the realities of mining extraction.[21] There he leaves aside the basic equation and turns to the issue of "cumulative production affecting price" (p. 152). The "net price" function becomes $p(x(t),q(t),t)$: cumulative production $x(t) = \int_0^t q(t)$ dt is now a variable, affecting the formation of net prices. Hotelling justifies the need for this variable on the cost side. It turns out resources are located under the ground, *in varying depths*. A first geological constraint (GC1) other than finiteness thus appears: "the cost of extraction increases as the mine goes deeper" (p. 152).

In section 12, "The Need for Steadiness in Production" (p. 162), Hotelling further expands his net price function to portray the rate of change in production $q'(t)$ as an additional variable. The cost of increasing production, Hotelling argues, depends on the magnitude of the increase. In developing a mine, producers have to make capital investments, which makes sudden increases in production more expensive than gradual ones. This is due to the *uneven distribution of concentrations of resources under the ground*—a second geological constraint (GC2) other than finiteness, which imposes a cost for drawing the labour force to or from the site of the mine, making large variations in output significantly costlier than small variations. Hotelling closes the section by noting:

> The cases considered in the earlier part of this paper all led to solutions in which the rate of production of a mine always decreases. By considering the influence of fixed investments

---

[20] Regarding Hotelling's conception of State intervention, some insight can be found in Franco et al. (2019).

[21] The word "mine(s)" is also employed a few times in the sections dedicated to the basic model, but to designate stylized mines, for which finiteness is the only characteristic considered. Hotelling makes clear from section 8 onwards that concrete mines do not fall under the same analysis.

and the cost of accelerating production at the beginning, *we may be led to production curves which rise continuously from zero to a maximum, and then fall more slowly as exhaustion approaches*. Certain production curves of this type have been found statistically to exist for whole industries of the extractive type, such as petroleum production (p. 164; our emphasis).

In section 15, the closing section of the paper, Hotelling comes back to the discussion between perfect competition and monopoly, presenting the case of duopoly, in which a few competing sellers can share the market's demand. In this model, Hotelling avoids the complex formulation of section 12, excluding for simplicity the influence of the rate of change in production as a determinant of the price. However, at this point, Hotelling makes reference to a third geological constraint (GC3), other than finiteness, namely the uncertainty on *the actual size of reserves* in the beginning and the need to acquire geological information through drillings:[22]

> The problems in exhaustible resources involve the time in another way besides bringing on exhaustion and higher prices, namely, as bringing increased information, both as to the physical extent and condition of the resource and as to the economic phenomena attending its extraction and sale (p. 174).

Therefore, aside from market structure and taxation, Hotelling's 1931 displays another, often overlooked, clear axis of progression from his basic result, towards the incorporation of GC1, GC2 and GC3, notably through $x(t)$ and $q'(t)$, as fundamental variables. While the number of firms that operate on a given market or its tax regime can in principle be affected by

policy, geological constraints are something no one can do anything about. Therefore, they should not be considered as "variants" or "possibilities," as mentioned by Slade and Thille (2009), nor as *additional factors*, but as fundamental inalterable characteristics with the same status as *finiteness*, affecting all mineral resources. In addition to the already mentioned closing of section 2, in the opening of section 8, Hotelling states the influence of $x(t)$ on the exploitation of resources (which results from GC1) is a *certainty*, not a mere *possibility*: "[t]he net price...*depends*...on past production,", i.e., it *depends* on it, not *may* depend on it (p. 152; our emphasis).[23]

Therefore, sections 8 to 15 of Hotelling's article discuss concrete real-world factors that define the category of *mineral resources* within the class of non-renewable resources, itself a subset of *non-renewable (exhaustible) assets*. Once more, in contrast to modern interpretations that keep the basic rule as a central principle to be sophisticated, it clearly emerges that this theoretical principle does not apply here. The rule governs only generic finite *assets* in the *absence* of geological constraints other than finiteness (GC1, GC2 and GC3). It therefore does not govern mineral resources (minerals, oil and gas), as presently overstated.

3.2. Hotelling's preparatory work in geology and engineering

By contrasting the archival materials with the final paper, we discovered the relative degree of effort Hotelling devoted to the topic of geological constraints. Two drafts from November 1928 are dedicated entirely to the issue of "strong variations," in which Hotelling includes for the first time cumulative production $x(t)$ as a fundamental variable, and devote significant effort to understanding the implications to the overall equilibrium.[24] Furthermore, in the draft abstract written in 1929, Hotelling does not present the basic model at all, starting with the model with $x(t)$, using

---

[22] In all mineral activities, the underground location of mines is initially unknown, and each mine or well drilled brings increased geological knowledge that facilitates the discovery of the next underground occurrence. Costs are not known with certainty from the start, which carries substantial investment risks. Only upon drilling wells and evaluating their output or how they respond after another well is drilled into the same oil reservoir, it it possible to infer engineering properties such as reservoir pressure and porosity and, as a result, determine the total number of wells needed to optimally exploit the resource.

[23] In other parts of the article (e.g., section 12), Hotelling uses less clear-cut vocabulary for the occurrence of geological factors, but he always argues they have to be of the utmost importance.

[24] HHP, Box 42, Exploitation of Irreplaceable Assets.

empirical examples to justify his formulation:[25]

> If the price obtainable varies with the amount which has been extracted, as with gold and diamonds, or if the cost of production increases as the shaft sinks deeper, the problem is to maximize $\int_0^T e^{(-\gamma t)} p(x,q,t) \, qdt$.

By 1930, Hotelling continued to focus on how geological constraints other than finiteness could affect the net price function. A noticeable preparatory work is his correspondence with Stanley C. Herold, professor of petroleum engineering at Stanford University—Hotelling was still at Stanford in 1930.[26] Herold wrote to Hotelling on July 7, 1930, in a friendly tone, providing empirical data on how, in the case of petroleum, costs increase as shafts sink deeper (GC1); see table 1.

TABLE 1
Increase in costs as shafts sink deeper

Depth	Cost ($)	Depth (ft.)	Cost ($)
3,000	10,000	6,000	100,000
4,000	30,000	7,000	150,000
5,000	60,000	8,000	210,000

SOURCE: Letter from Herold to Hotelling, July 7, 1930. HHP, Box 42, Exploitation of Irreplaceable Assets
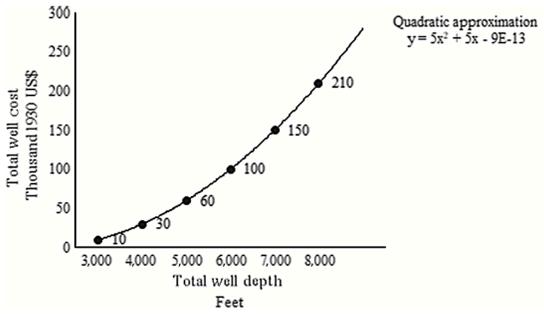

FIGURE 1   Our approximation of Herold's data sample on extractive costs

Table 1 shows a non-linear (quadratic) relationship between cost and depth; costs increase more than proportionately with well depth (figure 1).

GC2 and GC3 are also addressed in the letter, and some remarks can be traced to Hotelling's statements in his final article. For production speed at the beginning of the extraction, Herold writes:[27]

> The cost mounts more rapidly than the footage on account of the fact that deeper holes require heavier rigs and equipment, also more time is consumed in running in and out the holes.

Indeed, because different oil and gas reservoirs vary in depth and many other geological conditions, their unit production costs vary. Producers seek to developthe lowest cost ones first (Herfindahl 1967), although sometimes they are unable at first to discover the best reservoirs when only limited geological information is available (Uhler 1976). This leads to Hotelling's remark on GC2, when he mentions "the cost of accelerating production in the beginning" (1931, p. 164).

Still in the same letter, Herold describes the spatial element affecting costs, thus insisting on the heterogeneity of wells:[28]

> ...[H]oles at Kettleman Hills are costing a little more than shown here on account of transportation of equipment and general operating expenses, longer delays on replacement of equipment, etc.

This explains Hotelling's remarks in section 12 about costs related to the relocation of production factors, in particular the labour force:

> Under the term "capital" might possibly be included the costs, both to employers and to laborers, in drawing laborers to the mine from other places and occupations (Hotelling 1931, p. 163).

The example of Kettleman Hills is significant because it also highlights the variability of costs resulting from new discoveries: when promising wells are discovered—the Kettleman Hills oil fields, in particular around the North Dome, were a major discovery in California in the late 1920s—they attract capital and labour, with

---

[25] HHP, Box 42, Exploitation of Irreplaceable Assets.
[26] Hotelling becase professor of economics at Columbia University in 1931 after being approached by Wesley Mitchell (letter from Mitchell to Hotelling, Jan. 21, 1931. HHP, Box 50, Columbia Employment).

[27] Letter from Herold to Hotelling, July 7, 1930. HHP, Box 42, Exploitation of Irreplaceable Assets.
[28] Letter from Herold to Hotelling, July 7, 1930. HHP, Box 42, Exploitation of Irreplaceable Assets.

potentially abrupt variations in both production and costs.

Therefore, Hotelling's work between 1928 and 1930, as reflected in the archives, dealt almost exclusively with the impact of geological constraints on his basic model from 1925. The importance given to the incorporation of $x(t)$ as a variable is unrivalled, comparable only to the distinction between monopoly and perfect competition. In addition to Herold's 1930 letter to Hotelling, the 1928/1929 drafts, and how these documents were reflected in the final paper, demonstrate that geological constraints are a *necessary factor* to describe real-world minerals and fossil fuels. All such resources are subject to geological constraints.

3.3. Towards an alternative basic Hotelling model?

The constraints GC1 and GC2 are at the foundation of the technical understanding of what we see today as bell-shaped supply trajectories insofar as they explain the specific dynamics of extraction in relation to reserve degradation and to uncertainty about reserve size. We know that Hotelling was aware of C. E. Van Orstrand's 1925 article "On the Empirical Representation of Certain Production Curves" because Hotelling refers to it in his own 1931 article (p. 164). In his contribution, Van Orstrand pays particular attention to extractive sectors and depicts their dynamics by drawing three bell-shaped graphs (1925, pp. 26 and 32), including two with a long tale, depending on estimations. Hence, Hotelling knew of the existence of bell-shaped representations of the exploitation of exhaustible resources. Hotelling's description of extractive industries in his 1931 article corresponds to this kind of production trajectory, as shown in the earlier quotation:

> By considering the influence of fixed investments and the cost of accelerating production at the beginning, we may be led to production curves which rise continuously from zero to a maximum, and then fall more slowly as exhaustion approaches. Certain production curves of this type have been found statistically to exist for whole industries of the extractive type, such as petroleum production (p. 164).

More intriguing is precisely what Hotelling retained from his reading of Van Orstrand. In the archival material, we found a short memo, entitled "Mine Economics," that contained a full reference to Van Orstrand's paper as well as a handwritten graph, quickly drawn, representing a bell-shaped curve with a long tale (figure 2). Unlike Van Orstrand, who specified on the axes of his graphs the dates and volumes of production from real data on West Virginia, Ohio and Pennsylvania oil fields and so on, Hotelling sketches only the general slope, without the original figures, merely noting "production of petroleum" next to the vertical axis and "t" (for time) next to the horizontal axis.
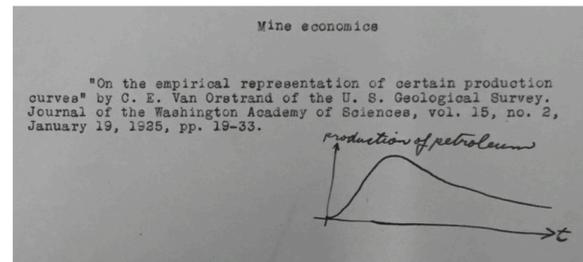

FIGURE 2 Hotelling's bell-shaped curve (ca. 1930), from HHP, Box 42, Exploitation of Irreplaceable Assets (fair use)

This is not anecdotal and suggests that his vision of the bell-shaped curve was generic and not contingent on this or that field. This type of approach—to transform specific quantified cases into generic stylized facts—reminds us of Cournot, who was the first in the history of economics (1838) to draw stylized curves from specific cases, in particular for demand functions. We know that Cournot played an important role in Hotelling's mathematical economics training (Hotelling 1929, p. 51; Darnell 1990, pp. 12 and 14), and we found references to Cournot in drafts and notes from the 1920s.[29] Hotelling's ambition with the bell-shaped curve therefore can be related to the description of a stylized fact for mineral resources, in Cournot's tradition.

The fact that "fixed investments and the cost of accelerating production at the beginning...

---

[29] "Talk on 'Mathematical Economics' before Math. Club, May 1928." HHP, Box 26, Mathematical Economics. We thank Marion Gaspard for drawing our attention to the connection between Cournot and Hotelling.

may...le[a]d [us]" to trajectories like the one sketched in Hotelling's hand-drawn graph suggests he knew that the incorporation of *q'* on the net price function would change the model towards a bell-shaped equilibrium. We were able to reproduce it ourselves through a simulation exercise, adjusting the linear net price function in section 8 to *p = a-bq-cq'-dq''*, where *q* is the cumulative production at time *t*, assuming Hotelling's original parameters and *d = 10.000* (see box 1 and figure 3).[30] Results indicate a bell-shaped equilibrium for production *q'* and an U-shaped equilibrium for market prices *p = a-cq'*.
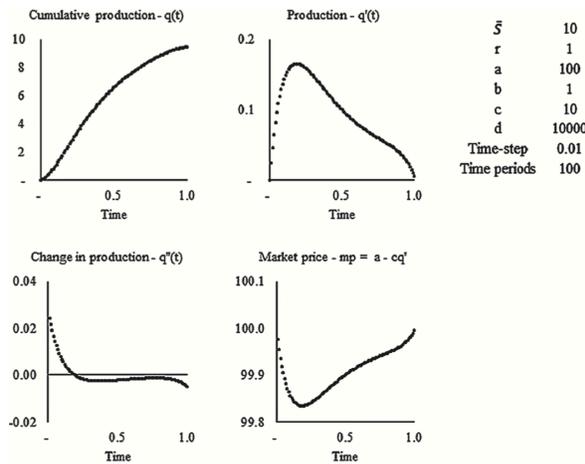

FIGURE 3 Numerical simulation of Hotelling's implied model in section 12

$$\max_q \int_0^\infty p(q,q',q'') \cdot q' \cdot e^{-rt} \cdot dt$$

subject to: $q \leq \bar{S}$, $q(0)=0$ and $q' \geq 0$, $q'(0)=0$

with: $p(q,q',q'')=a-bq-cq'-dq''$

where: $q(t)$ = cumulative production at time t
$q'(t)$ = production at time t
$q''(t)$ = change in production at time t
r = risk-free interest rate
$\bar{S}$ = total amount of recoverable resources
a,b,c,d = positive parameters

BOX 1 Mathematical formulation of Hotelling's implied model in section 12

In 1931, Hotelling did not have at his disposal mathematical tools such as optimal control theory (developed in the 1950s), allowing him to propose an alternative model to fit with this kind of observations. The hand-

---

[30] We make *g = 0* in order to avoid masking the U-shaped price path in equilibrium by an increasing demand as a function of time.

drawn graph, however, confirms his awareness of bell-shaped supply trajectories for the case of extractive industries, such as petroleum. His stylized-fact drawing, supplemented by his quotation, confirms that his basic model leading to the Hotelling rule was not relevant for concrete mineral resources. To the best of our knowledge, the suggested model in section 12, as described in box 1, has not been depicted in the literature to date. While its complete characterization might be insightful, it is beyond the scope of the present paper.

4. Conclusion

Based on unprecedented archival inquiry, this paper has reconstructed Hotelling's project on non-renewable resources in the 1920s leading to the publication of "The Economics of Exhaustible Resources" in 1931. We showed that Hotelling actually worked along two lines of research: a *generic* one on asset valuation, in the line of his theory of depreciation, and a *specific* one on concrete mineral natural resources, requiring more empirical insights from the real world. In contrast to what is often argued, his source of information for implementing concrete situations in his framework was not the conservation movement but rather specific experiences in US policy throughout the 1920s.

We found Hotelling carried out substantial preparatory work to circumscribe the geological conditions of mineral resource extraction. This is hardly noticeable to the naked eye in the 1931 article. However, after consulting the archives, we presented a new reading of the article where the importance of geological constraints for the treatment of minerals and fossil fuels is significantly increased. After considering the implication of their presence, Hotelling became aware of, notably, bell-shaped trajectories linked to U-shaped price trajectories.

In addition to presenting these novel findings in the history of economic thought, our investigation allowed us to derive theoretical implications. After nearly 90 years of developments since Hotelling's 1931 article, we still attribute the title "Hotelling rule" to his basic result $p_t = p_0 \cdot e^{r \cdot t}$ and consider it to be the starting-point principle of non-renewable

resource economics. This seems problematic given the rule lacks empirical validity and requires strong amendments in order to describe the long-term, aggregate behaviour of its target object. Our reconstruction of Hotelling's research project sheds light on this theoretical matter.

In particular, we established that Hotelling never saw his basic result $p_t = p_0 \bullet e^{r \bullet t}$ as a *rule* for non-renewable resource economics. Instead, he stated his basic result referred to a hypothetical stylized asset marked exclusively by finiteness, which is not the case of mineral resources. For this category, elaborating on his preparatory work in geology and engineering, Hotelling concluded it was necessary to consider as fundamental factors at least the influences of cumulative production $x(t)$, because costs increase as mines go deeper (GC1), and the rate of change in production $q'(t)$, because strong variations in output are costlier than small ones (GC2).

By sketching the implied model, we were able to confirm it is possible to obtain bell-shaped trajectories for supplies and U-shaped ones for prices, as described by Hotelling, at least for given values of the parameters. Future research should investigate the formal mathematical demonstration of this result. If confirmed, two inherent geological constraints (GC1 and GC2), in addition to exhaustibility, would be sufficient to reflect the long-term equilibrium of mineral resources. Hotelling's 1931 contribution would thus remain the starting point of non-renewable resource economics, albeit in a form different from that advocated to date.

-

Acknowledgements

-

Supporting information

Supporting information is available in the online version of this article (publisher's website).

-

References

Anderson, S.T., R. Kellogg and S. W. Salant (2018) "Hotelling under Pressure," *Journal of Political Economy* 126, 984–1026

Arrow, K.J. (1974) "In Memoriam: Harold Hotelling, 1895-1973," *American Economic Review* 64, 1102–3

——. (1980) "Review of 'Foundations: Essays in Philosophy, Logic, Mathematics and Economics', by F. P. Ramsey, edited by D. H. Mellor," *Journal of Political Economy* 88, 636–38

——. (1987) "Hotelling, Harold." In *The New Palgrave. A Dictionary of Economics, Vol.2*, pp. 670–71, eds. J. Eatwell, M. Milgate, and P. Newman. London, Basingstoke, New York & Tokyo: Macmillan

Atewamba, C., and B. Nkuiya (2017) "Testing the Assumptions and Predictions of the Hotelling Model," *Environmental and Resource Economics* 66, 169–203

Copeland, B.R., and S. T. Taylor (2017) "Environmental and Resource Economics: A

Canadian Restrospective," *Canadian Journal of Economics* 50, 1381–1413

Cournot, A.A. (1838) *Recherches Sur Les Principes Mathématiques de La Théorie Des Richesses*. Paris: Vrin & CNRS

Crabbé, P.J. (1983) "The Contribution of L. C. Gray to the Economic Theory of Exhaustible Natural Resources and Its Roots in the History of Economic Thought," *Journal of Environmental Economics and Management* 10, 195–220

——. (1986) "Unpublished Fragments on Natural Resources Economics and Excerpt from an Autobiography by Harold Hotelling," *Cahiers de Recherche Du GREEN* 8601,

Darnell, A.C. (1988) "Harold Hotelling 1895-1973," *Statistical Science* 3, 57–62

——. (1990) "The Life and Economic Thought of Harold Hotelling." In *The Collected Economics Articles of Harold Hotelling*, pp. 1–28. New York: Springer-Verlag

Dasgupta, P.S., and G. Heal (1979) *Economic Theory and Exhaustible Resources*. London: James Nisbet Co. and Cambridge University Press

Devarajan, S., and A. C. Fisher (1981) "Hotelling's 'Economics of Exhaustible Resources': Fifty Years Later," *Journal of Economic Literature* 19, 65–73

Ely, R.T. (1918) "Conservation and Economic Theory." In *The Foundations of National Prosperity. Studies in the Conservation of Permanent National Resources*, pp. 1–92, eds. R. T. Ely, R. H. Hess, C. K. Leith, and T. N. Carver. New York: Macmillan

Farzin, Y.H. (1984) "The Effect of the Discount Rate on Depletion of Exhaustible Resources," *Journal of Political Economy* 92, 841–51

——. (1992) "The Time Path of Scarcity Rent in the Theory of Exhaustible Resources," *Economic Journal* 102, 813–30

Fisher, I. (1906) *The Nature of Capital and Income*. London & New York: Macmillan & Co.

Franco, M.P.V., M. Gaspard and T. M. Mueller (2019) "Time Discounting in Harold Hotelling's Approach to Natural Resource Economics: The

Unsolved Ethical Question," *Ecological Economics* 163, 52–60

Gaspard, M., and A. Missemer (2019) "An Inquiry into the Ramsey-Hotelling Connection," *European Journal of the History of Economic Thought* 26, 352–79

Gaudet, G. (2007) "Natural Resource Economics under the Rule of Hotelling," *Canadian Journal of Economics* 40, 1033–59

Gaugler, T. (2015) "What Drives Resource Prices? A Qualitative Review with Recommendations for Further Development of the Hotelling Model," *Mineral Economics* 28, 37–51

Gray, L.C. (1913) "The Economic Possibilities of Conservation," *Quarterly Journal of Economics* 27, 497–519

——. (1914) "Rent under the Assumption of Exhaustibility," *Quarterly Journal of Economics* 28, 466–89

Guicherd, T. (2017) "Essai sur la genèse de la théorie de la concurrence monopolistique d'Edward H. Chamberlin - Entre histoire subjective et reformulation d'une théorie." PhD thesis, University of Lyon 2

Hands, D.W., and P. Mirowski (1998) "Harold Hotelling and the Neoclassical Dream." In *Economics and Methodology: Crossing Boundaries*, pp. 322–97. New York: St Martin's

Hart, R., and D. Spiro (2011) "The Elephant in Hotelling's Room," *Energy Policy* 39, 7834–38

Herfindahl, O.C. (1967) "Depletion and Economic Theory." In *Extractive Resources and Taxation*, pp. 63–90, ed. M. Gaffney. Madison, Milwaukee & London: University of Wisconsin Press

Holland, S.P. (2008) "Modeling Peak Oil," *Energy Journal* 29, 61–79

Hotelling, H. (1925) "A General Mathematical Theory of Depreciation," *Journal of the American Statistical Association* 20, 340–53

——. (1929) "Stability in Competition," *Economic Journal* 39, 41–57

——. (1931) "The Economics of Exhaustible

Resources," *Journal of Political Economy* 39, 137–75

Ise, J. (1925) "The Theory of Value as Applied to Natural Resources," *American Economic Review* 15, 284–91

Karp, L. (2017) *Natural Resources as Capital*. Cambridge (MA): MIT Press

Kemp, M.C. (1976) "How to Eat a Cake of Unknown Size." In *Three Topics in the Theory of International Trade*, pp. 297–308. Amsterdam: North-Holland Publishing Company

Kula, E. (1998) *History of Environmental Economic Thought*. London: Routledge

Levhari, D., and N. Liviatan (1977) "Notes on Hotelling's Economics of Exhaustible Resources," *Canadian Journal of Economics* 10, 177–92

Levhari, D., and R. S. Pindyck (1981) "The Pricing of Durable Exhaustible Resources," *Quarterly Journal of Economics* 96, 365–77

Livernois, J. (2009) "On the Empirical Significance of the Hotelling Rule," *Review of Environmental Economics and Policy* 3, 22–41

Missemer, A. (2017) *Les Économistes et La Fin Des Énergies Fossiles (1865-1931)*. Paris: Classiques Garnier

——. (2018) "Fossil Fuels in Economic Theory - Back to the 19th century British Debates," *Revue Française de Civilisation Britannique* XXIII

Missemer, A., M. Gaspard and R. P. Ferreira da Cunha (2020) "Depreciation and Exhaustible Resources, Exploring Harold Hotelling's Twin Projects" *#BNREproject Working Paper*. Paris

Missemer, A. and F. Nadaud (2020) "Energy as a Factor of Production: Historical Roots in the American Institutionalist Context," *Energy Economics* 86, 104706

Okullo, S.J., F. Reynès and M. W. Hofkes (2015) "Modeling Peak Oil and the Geological Constraints on Oil Production," *Resource and Energy Economics* 40, 36–56

Pindyck, R.S. (1978) "The Optimal Exploration and Production of Nonrenewable Resources," *Journal of Political Economy* 86, 841–61

Ramos Gorostiza, J.L. (2003) "Ethics and Economics: From the Conservation Problem to the Sustainability Debate," *History of Economic Ideas* 11, 31–52

Robinson, T.J.C. (1980) "Classical Foundations of the Contemporary Economic Theory of Non-Renewable Resources," *Resources Policy* 6, 278–89

——. (1989) *Economic Theories of Exhaustible Resources*. London & New York: Routledge

Salant, S.W. (1976) "Exhaustible Resources and Industrial Structure: A Nash-Cournot Approach to the World Oil Market," *Journal of Political Economy* 84, 1079–94

Slade, M.E. (1982) "Trends in Natural-Resource Commodity Prices: An Analysis of the Time Domain," *Journal of Environmental Economics and Management* 9, 122–37

Slade, M.E., and H. Thille (2009) "Whither Hotelling: Test of the Theory of Exhaustible Resources," *Annual Review of Resource Economics* 1, 239–60

Smith, G.A. (1982) "Natural Resource Economic Theory of the First Conservation Movement (1895-1927)," *History of Political Economy* 14, 483–95

Solow, R.M. (1974) "The Economics of Resources or the Resources of Economics," *American Economic Review* 64, 1–14

Uhler, R.S. (1976) "Costs and Supply in Petroleum Exploration: The Case of Alberta," *Canadian Journal of Economics* 9, 72–90

Van Orstrand, C.E. (1925) "On the Empirical Representation of Certain Production Curves," *Journal of the Washington Academy of Sciences* 15, 19–33

Venables, A.J. (2014) "Depletion and Development: Natural Resource Supply with Endogenous Field Opening," *Journal of the Association of Environmental and Resource Economists* 1, 313–36

Yergin, D. (1991) *The Prize: The Epic Quest for Oil, Money and Power*. New York: Simon & Schuster