# A Causal Mechanism of Policy Innovation: The Reform of Colombia's Oil-Rents Management System*

**Bayron Paz** – Guillaume Fontaine*****

Reception date: March 14, 2017 · Acceptance date: July 12, 2017 · Modification date: July 25, 2017
**DOI: https://dx.doi.org/10.7440/res63.2018.01**

**How to cite:** Paz, Bayron and Guillaume Fontaine. 2018. "A Causal Mechanism of Policy Innovation: The Reform of Colombia's Oil-Rents Management System." *Revista de Estudios Sociales* 63: 2-19. https://dx.doi.org/10.7440/res63.2018.01

ABSTRACT | This article analyzes policy innovation in Colombia, through the adoption of a new centralized oil-rent management system in 2011, after 20 years of decentralized policies. Using a policy-design framework, we identify a causal mechanism linking the opening of a policy window to policy change as a combination of the emergence of a new policy network, the adoption of a new policy paradigm and the selection of a new instruments mix. Drawing on Bayesian statistics, the 11 tests conducted on the causal mechanism show the importance of State resources of information, authority, treasury and organization to assess the outcome of policy change.

KEYWORDS | *Thesaurus*: Colombia; innovation. *Author*: policy design; process tracing; oil rents

## Un mecanismo causal de innovación de política pública: la reforma del sistema de gestión de rentas petroleras en Colombia

RESUMEN | Este artículo analiza la innovación de políticas públicas en Colombia, a través de la adopción de un nuevo sistema centralizado de manejo de rentas petroleras en 2011, tras veinte años de políticas descentralizadas de regalías. Aplicando un marco analítico basado en el diseño de políticas, identificamos un mecanismo causal relacionando la apertura de una ventana de oportunidad con un cambio de políticas públicas, como resultado de una combinación de la aparición de una nueva red de política, la adopción de un nuevo paradigma de políticas y la selección de una nueva mezcla de instrumentos. Con base en las estadísticas bayesianas, las once pruebas aplicadas al mecanismo causal muestran la importancia de los recursos estatales de información, poder, hacienda y organización para explicar el cambio de políticas públicas.

PALABRAS CLAVE | *Thesaurus*: Colombia; innovación. *Autor*: diseño de políticas; seguimiento de procesos; rentas petroleras

* This paper is a product of the research done on "Innovación en el diseño de políticas públicas: formalización bayesiana de process tracing aplicada a la política de regalías en Colombia," in partial fulfillment of the requirements for the Master's degree in Public Policy by Bayron Paz, under the supervision of Guillaume Fontaine. The project received a grant from the Academic Development Fund of FLACSO Ecuador, and discussed with the Research Group on Comparative Policy Analysis at FLACSO. The authors wish to thank Derek Beach, Edwin Alvarado and Iván Narváez for their comments. Any possible error remains the sole responsibility of the authors.

** Master's degree in Public Policy from FLACSO Ecuador. Degree in Economics from the Universidad de Nariño (Colombia). Professor, Researcher and Research Coordinator at the Universidad Mariana (Colombia). Research topics: policy change and innovation; extractive policy; fiscal policy; research methods. ✉ bpazn@umariana.edu.co

*** Doctorate in Sociology from Université Sorbonne Nouvelle-Paris 3. Post-Doctoral researcher at Sciences-Po (France). Senior researcher in the Department of Public Affairs of FLACSO Ecuador. Research topics: public accountability, policy design, resource nationalism. Latest books published: *El análisis de políticas públicas: Conceptos, teorías y métodos*. Quito: FLACSO – Anthropos, 2015; *Petropolítica: Una teoría de la gobernanza energética*. Quito: FLACSO – Instituto de Estudios Peruanos – Abya Yala, 2010. ✉ gfontaine@flacso.edu.ec

**Um mecanismo causal de inovação de política pública: a reforma do sistema de gestão de rendas petroleiras na Colômbia**

RESUMO | Este artigo analisa a inovação de políticas públicas na Colômbia por meio da adoção de um novo sistema centralizado de gestão de rendas petroleiras em 2011, após 20 anos de políticas descentralizadas de regalias. Aplicando um referencial analítico baseado no desenho de políticas, identificamos um mecanismo causal que relaciona a abertura de oportunidades com a mudança de política pública, enquanto combinação da emergência de uma nova rede de política, a adoção de um novo paradigma de política e a seleção de um novo mix de instrumentos. Com base nas estatísticas bayesianas, os 11 testes aplicados ao mecanismo causal mostram a importância dos recursos estatais de informação, autoridade, tesouro e organização para explicar a mudança de políticas públicas.

PALAVRAS-CHAVE | *Thesaurus*: Colômbia; inovação. *Autor*: desenho de políticas; seguimento de processos; rendas petroleiras

## Introduction: Innovation as Policy Change

Innovation is often related to competition and the private sector both in scholarship and in political discourse (OECD 2005; Reta 2016, 12), but governments are also interested in innovations that allow them to do more with less and, furthermore, in a transparent and consensual fashion (Considine, Lewis and Damon 2009, 25). Yet the public sector has been described as a rigid scenario, with low sensibility to contextual change and poor incentives (Altshuler 1997; Ballart 2001, 16), and a bureaucratic layout that tends to limit the emergence of new innovative proposals (Armstrong and Ford 2001, 18; Crozier, 1974; Kooiman 2003, 116).

A policy innovation is essentially a policy change, although not every policy change is an innovation (Howlett 2014, 397; Moore, Sparrow and Spelman 1997). So what special features should a particular policy change present in order to qualify as innovation? (Jordan and Huitema 2014a, 389). Policy innovations are not limited to postulating new and fascinating ideas but rather to implementing new practices implying a significant change of commonly accepted protocols (Altschuler and Zegans, 1997; Cejudo, Daussage and Michel 2016, 18). Thus, for a change to qualify as innovation, it ought to be relevant, general and sustainable for its effects on behaviors and operations to be noticeable (Moore, Sparrow and Spelman 1997, 276).

The definition of policy innovation can refer to invention (Bingham 1978, 180), diffusion (Rogers 1983, 11; Walker 1969, 881), or even to evaluation of subsequent effects (Daglio, Gerson and Kitchen 2015, 4; Hilden, Jordan and Rayner 2014; Jordan and Huitema 2014a; 2014b). The concept rests on its attributes as the introduction of disruptive, original, hitherto unseen and disturbing practices that permanently alter the fundamentals of a policy by moving away from its core *status quo ante* (Howlett 2014, 2; Lynn 1997; Polsby 1984). Furthermore, its implementation implies a certain risk (Altshuler 1997; Bingham 1978) due to its experimental nature, the uncertainty of its real impacts, and the possibility of producing unexpected effects. It is therefore relevant to ask at some point why and how innovations are adopted in a policy design.

This article aims at defining a theoretical causal mechanism to explain policy innovation with enough abstraction and parsimony to be tested empirically in case studies. Drawing on a critical realist ontology (Jackson 2016), the research was guided by a neo-institutional theory of policy change explaining the interactions between actors' interests and ideas with formal and informal institutions (Hall 1993). The argument is that policy innovation is a particular kind of policy change in which new objectives and means are combined in a hitherto unseen way.

The causal explanation is not oriented toward a predictive theory, which would derive from a probabilistic inference, but rather toward the identification of a context-dependent pattern defined as a causal mechanism linking institutional change to policy innovation. The case study is based on Colombia's policy of oil-rents management. This case presents an insightful experience of policy change, through the creation of a centralized General Royalties System (SGR for its Spanish acronym) in 2011 that reversed two decades of decentralized distribution of royalties.

After this introduction, the remainder of the article is organized as follows. The second section presents a theoretical discussion on policy innovation as an institutional problem. The third section describes the methods utilized to test this theory by drawing on process tracing and Bayesian formalization. The fourth section includes a description and a discussion of the results. The final section draws some conclusions regarding the theoretical and methodological implications of the case study.

## Theoretical Discussion

### Policy innovation as an institutional problem

Since Peter Hall's path-breaking work on policy change (1986), scholars have proposed many different explanations of policy change based on the interactions of interests, ideas and institutions (Fontaine, Narváez and Velasco 2017), which often overlook the need to define the causal mechanism through which these factors interplay to produce a particular outcome (Campbell 2002). Hall drew on Thomas Kuhn's theory of scientific revolutions (1971) to explain a paradigm change in the economic policy of the United Kingdom and France during the late 1970s and early 1980s.

In Hall's explanation, the trigger for a major policy change was a crisis revealing an accumulation of anomalies where real-life events could not be explained by the existing policy paradigm (Keynesian economics). To solve the crisis, the government was forced to adopt policy innovations going beyond short-term measures. This triggered the opening of a policy window for a shift in the locus of authority, with the advocates of the new paradigm (monetarist economics) proclaiming a new conception of policy problems and calling for new policy instruments. A policy network, linked to a specific profession in the State agencies in charge of a policy, gained increasing importance in the decision-making process. This theory depicts how policy innovation can result from stochastic events and resorts to the inventive and creative skills of political entrepreneurs who develop an acceptable solution to overcome the crisis under pressure. Hence new actors were able to diffuse new ideas, knowledge, habits and informal institutions through all the organizations where they were involved (Baumgartner 2013, 254; Powell and DiMaggio 1999, 109).

Admittedly, institutional inertia can prevent a crisis from altering behaviors and cause an imminent punctuated equilibrium (Mejía 2014, 182), as in the 2008 financial crisis, which presented the opportunity for a policy paradigm change (Béland and Cox 2013; Blyth 2013; Hall 2013) that has not occurred thus far. Policy innovation often goes hand in hand with institutional change, but institutional change is not immediate and can even take several decades before stabilizing the new interplays and normalizing the logics of action promoted by a new policy network. It is thus an incremental process by means of layering, displacement or shifting (Howlett and Cashore 2009; Thelen 2004).

Policy innovations are predictable inasmuch as they present solutions that anticipate problems that could affect the *status quo ante* in the future (Borins 2016, 118). As in the garbage can model (Cohen, Marsh and Olsen 1972), they are like existing solutions waiting for the opening of a policy window to foster the identification of a problem and the formulation of a new agenda (Kingdon 1995). Once a new policy network has gained enough legitimacy in a policy area, a strategy is developed to discredit the *status quo ante*, including the dramatization and diffusion of information on a policy's weaknesses or failures, the formulation of new programs of action, and the adaptation of discourse on the innovation to fit public sentiment (Campbell 1998). This strategy also includes a new policy design which ends up in a third-order change affecting various policy areas (Hall 1993). The innovation is secured when the new authority has won enough legitimacy to implement the new policy, even considering the high-risk scenario and without guaranteeing positive outcomes. The causal mechanism of a policy innovation is presented in Figure 1.

**Figure 1.** A causal mechanism of policy innovation

| Trigger | Entities engaging in activities | | | Outcome |
|---|---|---|---|---|
| Shift in the locus of authority (Institutional change + Opening of a policy window) | New policy network discredits the *status quo ante* | New policy network ideas overlook the policy area | Policy design expresses ideas of the new policy network | Policy innovation is implemented |

*Source:* Elaborated by the authors.

### A policy design framework

By explaining policy change and innovation at the level of instruments mix, this paper emphasizes the importance of the combination of instruments to achieve a government objective (Howlett and Rayner 2013, 172). A major problem for policy analysis is to find information that allows us to infer if a government is actually doing what they are saying they do, instead of performing a mere rhetorical exercise. Policy instruments are arguably the most objective materialization of the observed phenomenon, since they are the means available to a government to implement a policy (Howlett, Ramesh and Perl 2009; Lascoumes and Le Galès 2007; Linder and Peters 1998; Salamon 2010). Hence they offer a way to characterize the dependent variable and define elemental aspects such as the existence of a policy, or more complex ones such as the correspondence between aims and goals in a particular implementation style (Howlett 2009; Howlett and Cashore 2009).

The use of instruments mix as source of information allows us to evaluate the dynamics and complexity of innovation (Schaffrin *et al*. 2014, 861) through a typology of instruments by State resources available to promote or to restrict changes in their environment (Hood 1986; 2007; Howlett, Ramesh and Perl 2009). This typology includes instruments of nodality of information, authority or regulation, treasury and organization. Nodality instruments refer to a government's use of information in taking decisions. Authority instruments refer to legal mechanisms a government can create and use for coercion. Treasury instruments are those related to the financial capacity to define positive and negative incentives. Organization instruments are those related to the State's administrative capacity including ministries, bureaucratic agencies, etc.

This typology encompasses all the instrumental aspects of public policies. As such, the combination of nodality, authority, treasure and organization instruments is a necessary and sufficient condition for a policy to be effective, regardless of its inputs and outputs. In the first place, there is no policy without information, since one cannot take a decision without being informed about the problem to be treated; conversely, the State produces information on policies, since State actors transmit information through propaganda, awareness campaigns, accountability, etc. Second, there is no policy that is not regulated by legal rules, or that is not part of a legal framework: this can be either implicit (policy does not necessarily depend on a specific law) or explicit (law is the preferred policy instrument in certain regimes). Third, there is no policy without an economic counterpart: a policy without a budget is just a statement of principle. Finally, there is not a single policy whose formulation, implementation and monitoring do not depend on a team and a State agency: although many policies are not designed by those who implement them, no policy exists without an entity responsible for its implementation.

The interpretation of this information brings out a wealth of material for policy analysis and is particularly useful for comparison. These sources complement other qualitative techniques like interviews and direct observation. Organization charts are also a valuable input to identify qualified informants. Once information sources have been identified, policy instruments can be used for process tracing.

## Methods

### A theory-centered process tracing

Recent discussion on process tracing identifies three different applications of this method: for theory-building, theory-testing and explaining-outcome (Beach and Pedersen 2013, 13; George and Bennett 2005, 210). All three share a deterministic logic of inference and aim at opening the black box of causality. However, each version presents a different approach to causal mechanisms, defined as theoretical systems of interconnected parts that transmit a force from a trigger to an outcome (Beach and Pedersen 2013, 29). In theory-building process tracing, the inductive approach aims at identifying an unknown causal mechanism that is theorized in order to transcend a single case study. In theory-testing, the deductive approach aims at confirming or disconfirming the existence of an already known causal mechanism. Explaining-outcome process tracing aims at identifying or testing the existence of a causal mechanism, based on an outstanding historical event.

The research presented in this paper utilized a theory-testing process tracing method, in which a trigger and an outcome are linked by a mechanism that has already been conceptualized through a logical argument (Beach and Pedersen 2013, 14), based on the institutional theory of policy change. The aim is to increase confidence in the existence of a causal mechanism linking institutional change to policy innovation, by making empirical predictions about the evidence that should be present whether each part of the causal mechanism actually exists (Beach and Pedersen 2013, 95; Bennett and Checkel 2015, 30; Kay and Baker 2015, 12). These empirical predictions were formulated for each part of the causal mechanism of policy innovation and represent tests of inferential strength to assess their certainty and uniqueness (Bennett 2008; Collier 2011, 825; Van Evera 1997).[1] The causal inference in this sense proceeds through a "combination of affirmative evidence on some hypotheses

1 Increasing the degree of certainty means finding more evidence that a theory is true; increasing the degree of uniqueness means finding more evidence that rival theories can be discarded.

and eliminative induction of other hypothesized explanations that fail to fit the evidences" (Bennett 2008, 708).

The causal inference in this study follows a Bayesian logic, which aims at updating the degree of confidence in a theory. Bayesian logic particularly suits the logic of process tracing because it is the closest to the investigative reasoning common in daily life, medical experiments and trial procedures (O'Hagan and West 2010). Hence scholars wonder about the probability that a phenomenon would actually occur considering what is known about the context and what has been learnt from the theory and previous empirical studies. In its simplest expression, this process supports a narrative which is expected to be as consistent as possible. More formal process tracing aims at the modelization of a causal mechanism, based on the quantitative assessment of the available information.

The opportunity to engage in formal process tracing remains a moot point. Formalization and statistical models of a causal mechanism may create a false impression of precision, since the qualitative information does not necessarily fit the Bayesian categories of priors, hypotheses and evidence (Beach and Pedersen 2014, 10). However, advocates of formal process tracing argue that Bayesian statistics offers the best way to identify a specific locus of contention between researchers, which makes it an explanation process unlike over-indulgent congruence analysis (Checkel 2015). Most of all, formal process tracing leads to a higher commitment to transparent protocols and techniques that can be shared or discarded in the academic discussion (Bennett and Checkel 2015; Fairfield and Charman 2016; Humphreys and Jacobs 2015).

## The Bayesian logic underlying process tracing

That being said, Bayesian logic is based on a different conception of probability than quantitative methods like standard linear regression (Fairfield and Charman 2016, 2). While the latter seeks a regular causal relation based on a *ceteris paribus* assumption, the former aims at increasing the degree of confidence for each single causal relation, based on imperfect information. The question addressed at the formalization stage of a causal mechanism is how much does our initial confidence in the existence of each part of this mechanism increase (or decrease) when bringing out new evidence? (Befani and Mayne 2014, 23).

This process starts with Bayes' theorem, which measures how the examination of new evidence increases or decreases our degree of confidence in a theory or a hypothesis. In its simplest version the theorem requires three components to calculate the posterior probability of a theory (h) to be true, given the existence of an evidence (e): p(h/e). These components are: 1/ a prior probability of the theory to be true p(h) or false p(¬h), 2/ a true positive, or probability of the theory to be true in presence of evidence p(e/h), 3/ a false positive, or probability of the theory to be false in presence of evidence p(e/¬h) (Bennett 2015, 278). These components are presented in the following equation (taken from Bennett 2015, 281):

$$p(h/e) = \frac{p(h)\ p(e/h)}{p(h)p(e/h) + p(\neg h)\ p(e/\neg h)}$$

To test a causal mechanism implies identifying what evidence (comparable to clues in a criminal investigation) is expected to be found for each part of the mechanism, in the case of a true theory (confirming evidence) or a false theory (disconfirming evidence) (Beach and Pedersen 2013, 95; Bennett and Checkel 2015, 30; Kay and Baker 2015, 12). These empirical predictions can be identified by answering the following question: if each part of the mechanism exists, what empirical observations should be found in the case studied? (Beach and Pedersen 2014, 21). According to this statement, such empirical predictions present a specific combination of certainty and uniqueness and, taken as a whole, they define the inferential strength that allows us to move forward (Bennett 2008; Van Evera 1997).

The certainty of empirical predictions depends on the expected observations that are likely to be found in the real world (e), when the proven part of the mechanism actually exists ($h_n$). In Bayesian terms, certainty is a way to calculate the probability of true positives $p(e/h_n)$. Conversely, the uniqueness of empirical predictions depends on the expected observations that are likely to be found in the real world (e), when the proven part of the mechanism does not exist ($\neg h_n$). In Bayesian terms, uniqueness is a way to calculate the probability of false positives $p(e/\neg h_n)$ (Bennett 2015, 278). Consequently, a high degree of certainty is measured by a high value of $p(e/h_n)$; a high degree of uniqueness is measured by a low value of $p(e/\neg h_n)$.

## The problem of scaling

Formalization requires a punctual numerical value for each conditional probability mentioned above, so that these values can be included in Bayes' theorem. In other words, we must quantify how likely a piece of evidence is to be found whether the hypothesis is true ($p(e/h_n)$) or false ($p(e/\neg h_n)$).

The confidence in a theory can be ranked on a five-grade ordinal scale of values from "almost certainly disconfirmed" (if 0%<p(h/e)<9,99%), to "disconfirmed" (if 10%<p(h/e)<29,99%), to "somewhat disconfirmed" (if 30%<p(h/e)<49,99%), to "somewhat confirmed" (if 50%<p(h/e)<69,99%), to "almost certainly confirmed" (if 70%<p(h/e)<100%). Drawing on this

scale, we can rank the degrees of certainty and uniqueness in order to assign them a random value within a reasonable range. Hence certainty could be ranked from "not certain" (if 0%<p(e/h)<9,99%), to "almost not certain" (if 10%<p(e/h)<29,99%), to "somewhat not certain" (if 30%<p(e/h)<49,99%), to "somewhat certain" (if 50%<p(e/h)<69,99%), to "almost certain" (if 70%<p(e/h)<100%). Likewise, uniqueness can be ranked from "almost unique" (if 0%<p(e/¬h)<9,99%), to "somewhat unique" (if 10%<p(e/¬h)<29,99%), to "somewhat not unique" (if 30%<p(e/¬h)<49,99%), to "almost not unique" (if 50%<p(e/¬h)<69,99%), to "not unique" (if 70%<p(e/¬h)<100%) (adapted from Beach and Pedersen 2014; see also CIA 2005) (Table 1).

**Table 1.** Ordinal scale of uniqueness, certainty and confidence

| Range | Confidence | Certainty | Uniqueness |
|---|---|---|---|
| 0,01-0,09 | almost certainly disconfirmed | not certain | almost unique |
| 0,1-0,29 | disconfirmed | almost not certain | somewhat unique |
| 0,30-0,49 | somewhat disconfirmed | somewhat not certain | somewhat not unique |
| 0,50-0,69 | somewhat confirmed | somewhat certain | almost not unique |
| 0,70-0,99 | almost certainly confirmed | almost certain | not unique |

*Source:* Elaborated by the authors, adapted from Beach and Pedersen, 2014.

This proposal links each category describing informal Bayesian logic with a numerical rank, which offers a bridge to the implementation of formal protocols of Bayesian inference in process tracing. However, even though this scale means a significant improvement toward formal scaling of the conditional probabilities of each empirical prediction, these grades are separated by an infinite number of intermediate probabilities. Therefore, the model presented here assigns a random value to the intermediate grade linked to this category, in order to reduce the arbitrariness in defining this value, after selecting the ordinal category with which the conditional probability is described.

Before incorporating new contextual evidence, initial confidence in the existence of the causal mechanism and its parts in the case study depend on the theoretical debate and the available empirical analysis (Beach and Pedersen 2013, 84; 2014, 18; Bennett and Checkel 2015, 278; Collier 2011, 824). In a mechanism that specifies each part and expresses them in terms of hypotheses regarding their existence, it is possible either to assess the prior confidence for the whole mechanism as being the same for each part (Beach and Pedersen 2013, 98; Fairfield and Charman 2016, 37), or to define a different prior confidence for each part (Beach and Pedersen 2014, 17). For the sake of parsimony the present model has followed the former option, since the latter would imply a complex process of evaluating specialized information for each part of the mechanism, a topic that exceeds the scope of the present research.

## Empirical tests of certainty and uniqueness

Based on the degrees of certainty and uniqueness of empirical observations, four kinds of tests can be conducted: 1/ straw-in-the-wind tests, combining low certainty and low uniqueness; 2/ doubly-decisive tests, combining high certainty and high uniqueness; 3/ hoop tests, combining high certainty and low uniqueness; and 4/ smoking gun tests, combining low certainty and high uniqueness (Beach and Pedersen 2013; Bennett 2010; Bennett and Checkel 2015; Collier 2011).

Passing a test means that the evidence fits in with the empirical prediction; conversely, failing a test means that the evidence was not found. When passing a straw-in-the-wind test, the evidence neither confirms nor disconfirms a theory (h), and it hardly disconfirms alternative theories (¬h). Failing this test (in absence of evidence) neither disconfirms nor confirms h, but it slightly confirms ¬h. Conversely, when passing a doubly-decisive test, the evidence simultaneously confirms h and disconfirms ¬h. Moreover, failing this test eliminates h and significantly confirms ¬h. When passing a hoop test, the evidence does not confirm h, and it hardly disconfirms ¬h, but failing this test eliminates h and significantly confirms ¬h. Finally, when passing a smoking-gun test, an evidence confirms h and significantly disconfirms ¬h, but failing this test neither disconfirms h nor confirms ¬h.

Weakest tests are of little interest for a causal theory such as the one presented here, since they do not allow us to confirm nor to disconfirm a theory or its alternative. Strongest tests are difficult to perform straightforwardly in a single case study. Therefore, formal process tracing commonly combines various intermediate tests for each part of the causal mechanism. Hoop tests aim at discarding each hypothesis in absence of evidence; smoking-gun tests aim at confirming each hypothesis in presence of evidence. The combination of these tests can eventually provide a degree of confidence similar to that of a doubly-decisive test.

## Protocol of Bayesian formalization of process tracing

The formalization of the causal mechanism of policy innovation followed a four-step protocol that is described below and summarized in Annex 1.

The first step consisted of the theoretical conceptualization of a causal mechanism based on the existing theory. This mechanism can be described as the INUS condition for the policy innovation to be explained, which means each part of it is an Insufficient but Necessary part of an Unnecessary but Sufficient condition (Beach and Pedersen 2013, 30; Mackie 1965). A major consequence of this conceptualization is that the explanation switched from a probabilistic to a deterministic inference, meaning that only taken as a whole does the mechanism produce the outcome.

The second step consisted of the operationalization of the causal mechanism through a series of hypotheses, which required the identification of the different parts of the mechanism as entities engaging in activity and the contextual elements allowing the mechanism to work effectively (Beach and Pedersen 2013). This operationalization produced an empirical narrative consistent with the theoretical causal mechanism conceptualized in Step 1. Moreover, it considered the expression of each part of the mechanism as a hypothesis, which refers to the reasonable conjecture that each part of the mechanism actually exists (Beach and Pedersen 2013, 83). The hypothesized existence of a part of the causal mechanism —coded as $h_n$— is called "true theory", meaning that the part (n) actually exists empirically. The alternative hypothesis —coded as $\neg h_n$— is called "false theory," meaning the non-existence of a part of the causal mechanism.[2]

The third step consisted in the design of inferential strength tests or the definition of empirical predictions. Hoop tests and smoking-gun tests were performed for all but one entity of the causal mechanism. An additional doubly-decisive test was performed on the policy outcome. The empirical evidence came from the policy instruments of nodality, authority, treasury and organization, thus complementing the information gathered through interviews, speeches and press assessments reflecting the different stakeholders' positions regarding the policy innovation.

The fourth step consisted in the assessment of the degree of confidence for each part of the causal mechanism and for the mechanism as a whole. The prior level of confidence was defined according to a principle of indifference (Fairfield and Charman 2016, 6), so that the presence or the absence of the mechanism and each of its parts received even treatment previous to the information analysis. This implies that the prior value of each hypothesis is $p(h_n)=1/n$, where (n) is the number of defined hypotheses. In the present case, two exclusive hypotheses were formulated, so that $p(h_n)=p(\neg h_n)=0,5$.

The posterior level of confidence was then estimated through the incorporation of the values of empirical evidence into Bayes' theorem, after conducting a series of tests for each part of the mechanism. The results of the Bayes theorem can be: if $p(h/e)>p(h)$, then e supports h; if $p(h/e)< (h)$, then e disconfirms h; if $p(h/e)=p(h)$, then e is neutral towards h (Howson and Urbach 2006, 92, 98).

## Results and Discussion

### Historical background

Colombia has historically gone through successive commodity boom-and-bust cycles. Stochastic variations, either in prices, production, or both, have caused significant income instability and raised concerns regarding their wicked effects on macro-economic indicators. After the discovery of huge oil reserves in *Caño Limón* in July 1983, the country became a net exporter, and a new policy area emerged in the early 1990s to deal with a new public problem: how to allocate and invest oil rents[3] for development purposes.

As in many mineral-exporting countries, governments face the dilemma of taking advantage of a major revenue source and mitigating its negative economic, social and environmental effects. Their royalty policies provide a way to balance public spending between operative costs and investments in order to cope with this dilemma. In 2011 the royalty policy registered a major change with the adoption of new institutional rules affecting the relationships between the national and local governments. The change consisted of centralizing the SGR, thus reforming the decentralized system that had been in force since the early 1990s. This process qualifies as a policy innovation since it implied a series of original and previously unheard of disruptions in the policy area, through new practices that moved away from the *status quo ante*. Its implementation also included a significant level of risk and uncertainty regarding the expected results.

Although the legal concept of royalties dates back to the late 19th century (Crudotransparente 2015), it was not framed within a specific policy until the late 20th century, when it became the object of systematic decisions for planning by the government. The history of

2 Note that $(\neg h_n)$ has a proper content and is not the mere undetermined negation of $(h_n)$. Besides, $(h_n)$ and $(\neg h_n)$ form an exhaustive group of possibilities, since the process cannot generate any result other than to exist or not to exist. They are both mutually exclusive hypotheses, since one part cannot be simultaneously existing and non-existing.

3 Oil rents are the benefits the State receives from the exploitation of oil, either as royalties or as direct income from exports. These rents represent a significant share of the State's fiscal income, since oil activities were producing around 4,68% of GDP between 2000 and 2013 (Fedesarrollo 2015, 9) and 45% of exports since 2007 (OEC 2016).

the oil-rent management system since that time can be divided into three different periods.

From 1984 to 1991, the priority given to recovering energy self-sufficiency literally eclipsed any attempt to plan the investment and distribution of oil rents throughout the national territory. Still, this could hardly be considered a specific policy, for the lack of explicit objectives coupled with consistent actions and a consistent mix of instruments. For instance, the adoption of Law 75 in 1986 aimed at defining the government's share of the production, but it did not actually define the priorities or distribution criteria of mineral rents.

The second period started with the adoption of the new Political Constitution in 1991. This was an opportunity to legitimate a new mode of governance fostering decentralization and involving non-state actors in policy areas such as royalties. As for the exploitation of non-renewable natural resource, the Constitution enforced the political autonomy of local governments by creating administrative territorial entities with specific treasury instruments, in addition to traditional income from the national government. Therefore, although the Constitution did not question State ownership of underground non-renewable natural resources, it prioritized the right to participate in rents for the territorial entities where natural resources would be exploited, produced or transported. In practical terms, two distribution mechanisms were created: 80% of the royalties would be transferred directly to the departments involved in those activities, and 20% would be transferred to a National Royalties Fund to finance high priority projects at the national level (MHCP, MINMINAS and DNP 2011, 12).

Meanwhile, a specific institutional system of oil-rent management was designed that combined a series of rules, procedures, categories, and bureaucratic structures. Legitimate agencies were created to monitor the generation of income from oil production as well as its investment and evaluation. This shows that the government was pressed to design a policy that would barely be altered for two decades, partly due to the armed conflict and the veto power of legal and illegal local actors (Fontaine 2007).

During this second period, the royalty policy was framed within the paradigm of territorial decentralization and subsidies, promoted by local governments in oil-producing and oil-infrastructure areas. Departing from the general definition of royalties as compensation to the Colombian State for the right to exploit non-renewable natural resources which was granted by the Constitutional Court of Colombia in 1997, these actors advocated for more regional autonomy and defined royalties as a subsidy for the negative social and ecological impacts caused by oil activities. Hence departments and municipalities in oil-producing and -transportating areas were favored in the productive chain within that framework.

This policy faced three major criticisms. First, the distribution was considered unfair and uneven because of the concentration of direct transfers of royalties among seven departments representing only 14% of the Colombian population (Bonet and Urrego 2014, 2). Second, the system lacked effectiveness, since these royalties did not have the expected effect on local development (Benavides *et al*. 2000). Third, lawsuits were continually brought by control agencies and civil organizations for corruption, dubious investments and a low level of social participation in decisions on the use of oil rents (Universidad Nacional de Colombia 2013, 2.4).

In spite of these criticisms, former President Alvaro Uribe (2002-2010) failed to reform the system. He called for a referendum on October 25, 2003, which included a question about reforming Article 361 of the Constitution, so that indirect royalties would be used to address priority needs in terms of education, sanitation and potable water supply. Positive answers to that question did not pass the threshold of votes required for approval. President Uribe himself reckoned that he had failed to present the Congress with a formula for savings based on increased oil rents.

The third period began under the administration of President Juan Manuel Santos (2010 - 2018). Bill 2011-05 was passed in 2011, thereby reforming Articles 360 and 361 of the Constitution and ordering the creation of the SGR. The adoption of a new royalty policy required a change of logic in the use of non-renewable natural resources and oil rents. The policy would be oriented by a new paradigm, according to which the country's fiscal sustainability would become the guiding principle in the policy area. This paradigm reaffirmed State ownership of non-renewable natural resources, considering underground resources to be part of the nation's heritage, which resulted in all local governments having the same access to these incomes, regardless of their location on the mineral map (MHCP, MINMINAS and DNP 2011, 11). Hence the management of oil rents should benefit the nation as a whole, extending beyond compensation for local negative impacts (Echeverry, Alonso and García 2011, 3).

Departing from the former system, the SGR was divided into four financial funds: one for savings and stabilization, one for regional development, one for regional compensation, and another one for science, technology and innovation. Thereafter, decisions on financing would be made by collegial organizations, including representatives of both the central and the local governments. The allocation of resources was to be made on demand, and through competitive bids, instead of being automatically renewed through annual State-budget

allocation. Finally, the projects would be monitored by expert commissions during their implementation. The transition from the second to the third period is what is called policy innovation in this case. The systematic comparison of policy contents before and after innovation shows a dramatic change in policy aims and means derived from the shift from one paradigm to another. Table 2 presents a synthesis of these changes.

The next section presents the causal mechanism of this policy innovation.

**Table 2.** Comparative policy components of *status quo ante*/innovation

| Level of Abstraction | | General | Strategic | Operational |
|---|---|---|---|---|
| Policy aims | Status quo | Administrative and financial decentralization guiding royalty distribution | Compensation for the effects of oil exploitation at local level | Priority given to local governments in oil-producing and infrastructure areas |
| | Innovation | Economic and fiscal sustainability principles guiding royalties distribution | Savings and economic stability through even rent management | All governments (local and national) benefit evenly from royalties |
| Policy means | Status quo | Decentralization to local government | Resource management oriented by offer, proportional compensations and low accountability | Direct royalties, minimum definition of priorities for investment and coverage, irregular monitoring |
| | Innovation | Co-governance between central and local governments | Resource management oriented by demand, stabilization funds and high accountability | Indirect access to royalties, no pre-allocation, projects subjected to formal competitive requirements, constant monitoring |

*Source:* Elaborated by the authors, adapted from Howlett and Cashore, 2009.

## Conceptualization and hypotheses (steps 1 and 2)

The causal mechanism of policy innovation presented here explains how a shift in the locus of authority (T) leads to a policy innovation (O), through discrediting of the *status quo ante* (A), the adoption of a new policy paradigm (B), and the design of a new policy (C) (see Figure 1 and Annex 1).

Based on this conceptualization, the operational hypotheses state that the shift in the locus of authority (T) resulted from the combination of enhancement of new technical and economic agencies advocating for a centralized SGR ($h_{T1}$) and the election of a new president, opening a policy window for these agencies ($h_{T2}$). This allowed the diffusion of a discourse discrediting the *status quo ante* at the highest level of the State ($h_A$). This led the core ideas defended by the new policy network to spread out in the policy area ($h_B$). This caused the design of a new policy to be influences by these ideas ($h_C$). This resulted in the implementation of inedited, disruptive and uncertain reforms in the policy area ($h_O$).

The alternate hypotheses state that one or all of these hypotheses may not be true.[4] In other words, there was no shift in the locus of authority ($\neg h_T$), new technical and economic agencies advocating for a centralized SGR were not enhanced ($\neg h_{T1}$), and the election of a new president did not open a policy window for these agencies ($\neg h_{T2}$). The discourse discrediting the *status quo ante* was not diffused at the highest level of the State ($\neg h_A$). The core ideas defended by the new policy network did not spread out in the policy area ($\neg h_B$). The new policy designed was not influenced by these ideas ($\neg h_C$). The reform implemented in the policy area was not inedited, disruptive and uncertain ($h_O$).

## Empirical predictions and tests (step 3)

For each part of the causal mechanism, a series of tests were designed for empirical predictions. First, a series of hoop tests were performed on evidence regarding each part of the causal mechanism, to discard alternate hypotheses (see Annexes 1 and 2). While passing these tests would not significantly increase our degree of confidence in our theory, not passing them would almost disconfirm it.

As for the trigger, we expected local governments to condemn centralism ($P_{T1}1$), if new technical and economic agencies advocating for a centralized SGR were enhanced ($h_{T1}$), although these governments might also condemn centralism in absence of such enhancement ($\neg h_{T1}$). We also expected the new president's background in economics to be akin to the nature of these technical and economic agencies ($P_{T2}1$), if his election opened a policy window for them ($h_{T2}$), although such a background might also be akin to them in the case that his election did not open a policy window ($\neg h_{T2}$).

Regarding part A of the mechanism, we expected local governments to condemn a strategy of discredit by the national government ($P_A1$), if a discourse discrediting the *status quo ante* was diffused at the highest level of the State ($h_A$), even though the government might condemn this alleged strategy in absence of such a discourse ($\neg h_A$).

Regarding part B, we expected the new policy network to lead debates on the policy change ($P_B1$), if their core ideas were to spread out in the policy area ($h_B$), even though they might lead these debates without their ideas spreading out in the policy area ($\neg h_B$).

Regarding part C, we expected local governments to condemn the concentration of resources by the national government ($P_C1$), if the new policy design was influenced by the new policy network ($h_C$), although these governments might express the same criticism in absence of instruments reflecting the policy network's core ideas ($\neg h_C$).

As for the outcome of the process, we expected the enhanced technical and economic agencies to praise the new policy design for its novelty ($P_O1$), if the implemented reform in the policy area was inedited, disruptive and uncertain ($h_O$), even though they might have the same attitude in absence of such a reform ($\neg h_O$).

In order to complement these tests, we proceeded with a series of smoking-gun tests on evidence regarding the same entities, to confirm or disconfirm our hypotheses. While this evidence would not completely allow us to discard alternate hypotheses, it would significantly increase our degree of confidence in our theory.

As for the trigger, we expected policy instruments of organization to indicate the presence of new technical and economic agencies ($P_{T1}2$), if new technical and economic agencies advocating for a centralized SGR were enhanced ($h_{T1}$). Conversely, it would be highly unlikely to find such evidence if these agencies were not enhanced ($\neg h_{T1}$).

Regarding part A of the mechanism, we expected to find fingerprints of the discredit of the *status quo ante* in the policy instruments of nodality ($P_A2$), if a discourse discrediting the *status quo ante* was diffused at the highest level of the State ($h_A$). Again, it would be unlikely to find such evidence in absence of diffusion of this discourse by the government ($\neg h_A$).

4 The alternate hypotheses are formulated for each part of the mechanism. Since causal inferences in process tracing are asymmetric, any causal relationship between these hypotheses is irrelevant for the present study.

Regarding part B, we expected the expression of the ideas of the new policy network to be part of the State's highest values ($P_B2$), if their core ideas were to spread out in the policy area ($h_B$). We assume that it would be unlikely to find such evidence if these ideas had not spread out in the policy area ($\neg h_B$).

Regarding part C, we expected to find policy instruments of authority affirming the centralization of the SGR ($P_C2$), if the new policy design was influenced by the new policy network ($h_C$). We assume that the presence of such instruments would be highly unlikely if the policy network had no influence on the new policy design ($\neg h_C$).

Eventually, a doubly-decisive test was performed on the evidence regarding the outcome, in order to confirm our theory and to discard alternate theories simultaneously. We expected to find fingerprints of hitherto unseen changes in both the aims and means of the new policy implemented ($P_O2$), if the implemented reform in the policy area was inedited, disruptive and uncertain ($h_o$).

The results of the inferential tests for each part of the causal mechanism, based on the available empirical evidence, are presented in the following section and summarized in Annex 2.

## Bayesian formalization (step 4)

### Tests on T: Confirming the shift of authority

Evidence ($e_{T1}1$) for $P_{T1}1$ included newspapers, letters and debates, assessments by advocates (scholars, intellectuals and politics) of territorial autonomy against centralism. They indicate that governments from oil-producing and infrastructure areas are the main opposition challenging the policy change. They are organized in a block of political and specialized organizations that monitor the national government's decisions regarding oil rents. Nonetheless, they produce a general narrative according to which the policies implemented by the Santos administration contradict decentralization and regional autonomy. These complaints were even expressed during the Uribe administration when the decentralization was actually implemented. This evidence qualifies as almost certain and almost not unique, so that $p(e_{T1}1/h_{T1})=0,87$ and $p(e_{T1}1/\neg h_{T1})=0,57$. Therefore our posterior confidence in $h_{T1}$ is 0,60, which upgrades our prior by 0,10.

Evidence ($e_{T2}1$) for $P_{T2}1$ came from the Curriculums Vitae of President Santos and former President Uribe and official information on their careers. Both leaders were founders of the *Partido Social de Unidad Nacional* (Social National Unity Party), also known as *Partido de la U*, in 2005. Juan Manuel Santos was also Minister of Defense during the Uribe administration, so the 2010 election did not constitute any shift of political parties. The change from one president with a highly political profile to another with a predominantly technical and economic profile may have triggered the opening of a policy window. The fact that President Santos is an economist whereas ex-President Uribe is a lawyer, makes the former more receptive than the latter to the ideas expressed by technical and economic agencies. However, the new president may have other ideas and prefer to listen to other agencies. Furthermore, institutional friction and resistance among other actors both within and outside the State may have prevented the opening of a policy window. This is consistent with the idea of not assigning too great a capacity for fostering policy change to a single actor, even to the most legitimate and powerful one. This evidence qualifies as somewhat certain but somewhat not unique, so that $p(e_{T2}1/h_{T2})=0,60$ and $p(e_{T2}1/\neg h_{T2})=0,38$. Therefore our posterior confidence in $h_{T2}$ is 0,61, which upgrades our prior by 0,11.

Evidence ($e_{T1}2$) for $P_{T1}2$ came from incremental change in the instruments of organization reinforcing the National Planning Department during the 1991-2012 period. A long-run strategy of institutional strengthening leads to the multiplication of competences and to increasing the number of personnel, agencies and other organization instruments. However, other agencies (either new or older ones) become part of the State's organization during this institutional reform. Moreover, the agencies supporting the reform could influence the policy by other means, e.g. through public-private partnerships, private advising, etc. Conversely, finding information about the presence of the new policy network within the organization of the State would hardly be possible in absence of a strategy strengthening these agencies, although it does not prove that they are the only ones influencing the president. This evidence qualifies as somewhat not certain but somewhat unique, so that $p(e_{T1}2/h_{T1})=0,37$ and $p(e_{T1}2/\neg h_{T1})=0,13$. Therefore our posterior confidence in $h_{T1}$ is 0,74, which upgrades our prior by 0,24.

Eventually, our prior confidence (0,5) in the existence of a shift of authority (T) increased by 0,24 after performing $P_{T1}2$ on $h_{T1}$, but only by 0,11 after performing $P_{T2}1$ on $h_{T2}$, so that it remains somewhat confirmed.

### Tests on A: Confirming the diffusion of discourse discrediting the *status quo ante*

Evidence ($e_A1$) for $P_A1$ included public statements by local governments about what some of them called "a campaign of humiliation and disinformation" that accused them of corruption and inefficiency. Since such actors are supported by an active political block, we can logically surmise that, when feeling threatened by a strategy, they would protest against it with all available means. Conversely, it is highly unlikely that they would not react when feeling threatened, yet one may also assume that these actors would deploy a strategy of constant opposition to any government organization threatening

their resources. This evidence qualifies as almost certain but almost not unique, so that $p(e_A1/h_A)$=0,91 and $p(e_A1/\neg h_A)$=0,65. Therefore our posterior confidence in $h_A$ is 0,58, which upgrades our prior by 0,08.

Evidence ($e_A2$) for $P_A2$ included official publications by the Ministry of Finance condemning implementation gaps in the former policy and justifying the new reform. Discredit can also be done by non-official means that are hard to detect (for instance, by leaking information to the mass media), but the Finance Minister personally acknowledged having used new language based on metaphors and images that would help the diffusion of technical measures. In a strategy of discredit, a government would normally produce and analyze information to assess the strengths and weaknesses of former policies. Furthermore, it is almost impossible to find official information discrediting the *status quo ante* in absence of a strategy of discredit deployed by the policy network. This evidence qualifies as somewhat certain and almost unique, so that $p(e_A2/h_A)$=0,52 and $p(e_A2/\neg h_A)$=0,01. Therefore our posterior confidence in $h_A$ is 0,98, which upgrades our prior by 0,48.

Eventually, our prior confidence (0,50) in the diffusion of a discourse discrediting the *status quo ante* at the highest level of the State (A) was notoriously upgraded after performing $P_A2$ test on $h_A$, so that it was almost certainly confirmed.

### Tests on B: Confirming the spreading out of new ideas in the policy area

Evidence ($e_B1$) for $P_B1$ came from the legislative debates on Bill 2011-05 prepared by the national government to reform the Political Constitution on the matter of oil-rent management. This bill had to pass through eight debates, including commissions and plenary sessions of the Senate and the Chamber of Representatives, during the course of a single administration, and was finally approved by absolute majority in the second round. In a democratic system with high horizontal accountability, these ideas could easily be vetoed when they are not supported by a majority. Moreover, if the core ideas of the new policy network overlooked the policy area, it is highly probable that this network would have benefitted from majority support to do so. Conversely, there is little probability that this policy network would be able to lead the change in absence of a majority, yet it is possible to find a political majority supporting the new ideas, even when these ideas do not overlook the policy area. This results from the fact that great reforms supported by a political coalition can be vetoed by controlling and monitoring agencies (such as the Attorney General or the Constitutional Court). This evidence qualifies as almost certain but somewhat not unique, so that $p(e_B1/h_B)$=0,88 and $p(e_B1/\neg h_B)$=0,36. Therefore our posterior confidence in $h_B$ is 0,71, which upgrades our prior by 0,21.

Evidence ($e_B2$) for $P_B2$ came from the Political Constitution of 1991 and Bill 05 of July 3, 2011 which established the basis for the incorporation of economic and fiscal principles as the main criteria for the new policy. Formal referents of the core values of the State (such as the Political Constitution) are instruments dedicated to superior principles guiding the overall public policies of a government. As such, they leave little room for dispute among individuals and organizations supporting contradictory ideas. However, as the neo-institutional theory states, formal referents are but one specific kind of guide and even in their absence certain ideas might overlook a policy area. Furthermore, when observing that an idea has been raised to the highest level of normative references, it is almost impossible for the theory to be false. Hence the new ideas would arguably overlook the policy area. This evidence qualifies as somewhat not certain and almost unique, so that $p(e_B2/h_B)$=0,33 and $p(e_B2/\neg h_B)$=0,01. Therefore our posterior confidence in $h_B$ is 0,97, which upgrades our prior by 0,47.

Eventually our prior confidence (0,50) in the spreading out of the core ideas of the policy network in the policy area (B) was significantly upgraded after performing $P_B2$ test on $h_B$, so that it was almost certainly confirmed.

### Tests on C: Confirming the influence of the policy network on the new instruments

Evidence ($e_C1$) for $P_C1$ came from the debate on the SGR, during which advocates of the *status quo ante* would condemn the alleged overrating of the National Planning Department. The demands were received by the Attorney General's Office (*Contraloría*) who emitted a prior warning. We also considered the fact that local governments protested against the concentration of resources, responsibilities and leadership positions by the national government. Yet during the 1994 reform, local governments from oil-producing and infrastructure areas also protested against the concentration of resources and power by the national government, even though the national government shared leadership with them on royalty policy. This evidence qualifies as almost certain and somewhat not unique, so that $p(e_C1/h_C)$=0,74 and $p(e_C1/\neg h_C)$=0,42. Therefore our posterior confidence in $h_C$ is 0,64, which upgrades our prior by 0,14.

Evidence ($e_C2$) for $P_C2$ came from Law 1530-2012, which defined the many functions of the National Planning Department, including the management of a projects database, the technical secretariat of the responsible commission, and the system of monitoring and control. One significant advantage of designing a policy is that it helps to define the role played by other stakeholders and organizations during implementation of the policy. If technical and economic agencies were in a position of leadership in the design of the royalty policies, they would probably use instruments of authority to secure a large degree of control over the policy

orientation, even if it is possible to find technical and economic agencies conducting the design of a policy without any legal framework ordering it. Moreover, when finding evidence confirming that instruments of authority assign the technical and economic agencies a large degree of control over the policy orientation, there would be little probability that these agencies did not influence the policy design. This evidence qualifies as somewhat not certain but almost unique, so that $p(e_C2/h_C)$=0,33 and $p(e_C2/\neg h_C)$=0,01. Therefore our posterior confidence in $h_C$ is 0,97, which upgrades our prior by 0,47.

Eventually, our prior confidence (0,50) in the core ideas of the policy network influencing the design of the new policy (C) was significantly upgraded after performing $P_C2$ test on $h_C$, so that it was almost certainly confirmed.

### Tests on O: Confirming the implementation of inedited, disruptive and uncertain reforms

Evidence ($e_O1$) for $P_O1$ came from declarations by former Finance Minister Juan Carlos Echeverry about Colombia's royalty policy being an international innovation, and official documents on the royalty policy published by the Uribe and Santos administrations. The fact that technical and economic agencies involved in the new policy claim it to be novel is a basic signal to consider that a policy is actually innovative. If a policy is novel, it is highly probable that its promoters will claim that it is, even when there is a slight possibility that they could avoid doing so if they wish to avoid confrontation with their opponents. Nevertheless, the promoters of a policy may claim its novelty even if it is just a continuation of an older one. In other words, it is not sufficient to simply claim the novelty of a policy for it to actually be innovative. This evidence qualifies as almost certain and not unique, so that $p(e_O1/h_O)$=0,79 and $p(e_O1/\neg h_O)$=0,76. Therefore our posterior confidence in $h_O$ is 0,51, which upgrades our prior by 0,01.

Evidence ($e_O2$) for $P_O2$ includes a systematic comparison of the policy aims and means showing first-order and second-order change (see Table 1 above). On the one hand, the existence of objectives and means consistent with the new paradigm is highly certain evidence of the existence of a policy innovation. On the other hand, it is almost impossible to find hitherto unseen aims and means in absence of a policy innovation. This evidence qualifies as almost certain and almost unique, so that $p(e_O2/h_O)$=0,95 and $p(e_O2/\neg h_O)$=0,01. Therefore our posterior confidence in $h_O$ is 0,99, which upgrades our prior by 0,49.

Eventually, our prior confidence (0,50) in the implementation of inedited, disruptive and uncertain reforms in the policy area (O) was significantly upgraded after performing $P_O2$ test on $h_O$, so that it was almost certainly confirmed.

## Conclusions

This research utilized a neo-institutional framework to explain the process of policy innovation. In particular, the review of the literature on policy change offers an opportunity to fill in a conceptual gap in the study of innovations in the public sphere, which is valuable to overcome the dependence on theories of innovation in the private sphere (Ballart 2001; Lynn 1997; Setnikar and Petkovšek 2013). This analysis of change treats policy innovation as a dependent variable, to define the institutional determinants and to discriminate real innovations from symbolic or rhetorical ones.

The research contributed to understanding the causal mechanism of a policy innovation triggered by institutional change. A single case study, such as the 2011 reform of Colombia's royalty policy, could never establish that this mechanism is the only one possible. Yet our degree of confidence in the uniqueness of this theory was dramatically increased by performing empirical tests for each part of this causal mechanism. Even if the evolution of contextual factors —such as the decrease in oil prices since mid-2014— might put the innovation at stake, so far it appears to be stabilized and institutionalized in a way that has already altered the mode of governance and the relationships between the central and the local governments. Eventually, the causal mechanism tested in this study shows that even though policy innovations materialize at a specific moment, several years may pass before a particular event actually triggers them.

The key point in the methodological approach was the elaboration of a protocol that would help to achieve complementarity between a narrative process tracing and formal Bayesian logic. This would contribute to the systematic development and evaluation of the degrees of confidence with which the existence of a causal mechanism could be assessed. It would also confirm that a strict development of the different stages is necessary before engaging in modeling, from the operationalization of the mechanism to the definition of contextual factors, the elaboration of hypotheses, and the design of empirical predictions and tests of inferential strength guided by informal Bayesian logic.

The Bayesian formalization significantly increased the confidence in the existence of a causal mechanism of policy innovation in the case of Colombian oil-rent management since 2011. Moreover, this underlines the advantage of combining different empirical tests on the same hypothesis, in order to increase both the certainty and the uniqueness of the empirical evidence.

Another finding of importance for further policy analysis based on formal process tracing, is the added value provided by policy instruments as a major source of information. Inferences made from primary

or secondary sources of information involving what stakeholders say, instead of what they actually do, can easily be countered unless they are sustained by a complex and never-ending triangulation process. Policy instruments are highly relevant because they provide for clear information about what stakeholders actually do (or do not do) in a policy area. At the same time, the typology of State resources in terms of nodality, authority, treasury, and organization was extremely useful for parsimoniously describing the policy area and tracing the change through time.

Eventually, the use of policy instruments as a source of information goes beyond the mere description of the dependent variable (a policy change or innovation). Taken individually, each instrument qualifies as highly unique evidence; taken as a whole, a policy instrument mix qualifies as highly certain evidence. This means their presence or absence provides for strong discriminating evidence. This can be formalized by the likelihood ratio of empirical evidence LR(h)=p(e/h)/p(e/¬h), showing how the most convincing evidence is related to policy instruments. The likelihood ratio based on policy instruments ranges between 2,85 and 95, compared to a likelihood ratio ranging from 1,04 to 2,44 for other sources (see Annex 2), indicating clear superiority of the former set of evidence as compared to the latter.

## References


1. Acosta, Amylkar. 2012. *La reforma al régimen de regalías y su impacto en las finanzas territoriales.* Bogota: Impregon.
2. Altshuler, Alan. 1997. "Public Innovation and Political Incentives: The Innovations in American Government Program". Massachusetts: Ash Center for Democratic Governance and Innovation, Harvard Kennedy School. https://www.innovations.harvard.edu/sites/default/files/2595.pdf
3. Altshuler, Alan and Marc Zegans. 1997. "Innovation and Public Management: Notes from the State House and City Hall." In *Innovation in American Government: Challenges, Opportunities, and Dilemmas*, edited by Alan Altschuler and Robert Behn, 68-80 Washington: Brookings Institution Press.
4. Armstrong, Jim and Robin Ford. 2001. "Public Sector Innovations and Public Interest Issues." *The Innovation Journal: The Public Sector Innovation Journal* 6 (1): 1-22.
5. Ballart, Xavier. 2001. *Innovación en la gestión pública y en la empresa privada*. Madrid: Diaz de Santos.
6. Baumgartner, Frank. 2013. "Ideas and Policy Change." *Governance: An International Journal of Policy, Administration and Institutions* 26 (2): 239-258.
7. Baumgartner, Frank and Bryan Jones. 2002. *Policy Dynamics.* London: The University of Chicago Press.
8. Beach, Derek and Rasmus Brun Pedersen. 2013. *Process-Tracing Methods: Foundations and Guidelines.* Michigan: University of Michigan.
9. Beach, Derek and Rasmus Brun Pedersen. 2014. "Let the Evidence Speak: A Two-Stage Evidence Assessment Framework for Making Transparent the Process of Translating Empirical Material into Evidence." *American Political Science Association Annual Meeting.* Washington. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2451899
10. Béland, Daniel and Robert Henry Cox. 2013. "Introduction to Special Issue: The Politics of Policy Paradigms." *Governance* 26 (2): 193-195. http://dx.doi.org/10.1111/gove.12034
11. Befani, Barbara and John Mayne. 2014. "Process Tracing and Contribution Analysis: A Combined Approach to Generative Causal Inference for Impact Evaluation." *IDS Bulletin* 45 (6): 17-36. http://dx.doi.org/10.1111/1759-5436.12110
12. Benavides, Juan, Alberto Carrasquilla, Juan Gonzalo Zapata, Andrés Mauricio Velasco, Michelle Link. 2000. *Impacto de las regalías en la inversión de las entidades territoriales.* Bogota: Fedesarrollo. http://gidrot.com/materials/docs/est/est49.pdf
13. Bennett, Andrew and Jeffrey Checkel. 2015. *Process Tracing: From Metaphor to Analytic Tool.* Cambridge: Cambridge University Press.
14. Bingham, Richard. 1978. "Innovation, Bureaucracy and Public Policy: A Study of Innovation Adoption by Local Government." *The Western Political Quarterly* 31 (2):178-205. http://dx.doi.org/10.2307/447811
15. Bonet, Jaime and Joaquín Urrego. 2014. "El Sistema General de Regalías: ¿mejoró, empeoró o quedó igual?" *Documentos de trabajo sobre economía regional* 198. http://www.banrep.gov.co/es/dtser-198
16. Borins, Sandford. 2016. "La persistencia de la innovación en el gobierno: una guía para funcionarios públicos innovadores". En *La Innovación en el Sector Público: tendencias internacionales y experiencias mexicanas*, edited by Guillermo Cejudo, Mauricio Daussage and Cynthia Michel, 85-132. Mexico: CIDE-INAP.
17. Campbell, John. 1998. "Institutional Analysis and the Role of Ideas in Political Economy." *Theory and Society* 27: 377-409.
18. Campbell, John. 2002. "Ideas, Politics, and Public Policy." *Annual Review of Sociology* 28 (1): 21-38.
19. Cejudo, Guillermo, Mauricio Daussage and Cynthia Michel. 2016. *La Innovación en el Sector Público: tendencias internacionales y experiencias mexicanas.* Mexico: CIDE-INAP.
20. CIA. 2005. "Intelligence Report: Bayes' Theorem in the Korean War." Directorate of Intelligence. https://www.cia.gov/library/center-for-the-study-of-intelligence/kent-csi/vol16no2/html/v16i2a03p_0001.htm
21. Cohen, Michael, James March and Johan Olsen. 1972. "A Garbage Can Model of Organizational Choice." *Administrative Science Quarterly* 17 (1): 1-25. http://dx.doi.org/10.2307/2392088
22. Considine, Mark, Jenny Lewis and Alexander Damon. 2009. *Networks, Innovation and Public Policy.* London: Palgrave Macmillan.

23. Corte Constitucional de Colombia. 1997. Sentencia C-221/97. *Sentencias de la Corte Constitucional*. Bogotá: CCC.
24. Crudotransparente. 2015. "Informe especial: el reto de comprender el Sistema Nacional de Regalías." Retrieved 3 April 2016, from Crudotransparente.com: http://goo.gl/qtXbr6
25. Daglio, M., Gerson D., Kitchen H. 2015. "Building Organisational Capacity for Public Sector Innovation." Background Paper prepared for the OECD Conference *Innovating the Public Sector: from Ideas to Impact*, Paris, 12-13 November 2014. http://www.oecd.org/innovating-the-public-sector/Background-report.pdf
26. Echeverry, Juan Carlos, Gloria Alonso Masmela y Amparo García. 2011. ¿Por qué es necesaria la creación de un Sistema General de Regalías? Ministerios de Hacienda y Crédito Público, *Notas Fiscales* 2.
27. Fairfield, Tasha and Andrew Charman. 2016. "Formal Bayesian Process Tracing: Guidelines, Opportunities, and Caveats." *The London School of Economics and Political Science*. http://eprints.lse.ac.uk/62368/1/Fairfield_Formal%20Bayesian%20process%20tracing.pdf
28. Fedesarrollo. 2015. "Informe de coyuntura petrolera." Bogota: La imprenta editores S.A.
29. Fontaine, Guillaume 2007. *El precio del petróleo. Conflictos socio-ambientales y gobernabilidad en la región amazónica*. Quito: FLACSO – IFEA.
30. Fontaine, Guillaume, Iván Narvaez and Susan Velasco. 2017. "Explaining a Policy Paradigm Shift: A Comparison of Resource Nationalism in Bolivia and Peru." *Journal of Comparative Policy Analysis: Research and Practice*. Online version: 1-16 http://dx.doi.org/10.1080/13876988.2016.1272234
31. Fundación Foro Nacional por Colombia. 2013. *La normativa minera en países de América Latina: un estudio sobre Bolivia, Brasil, Chile, Colombia, Ecuador, México y Perú.* Bogota: Foro.
32. George, Alexander and Andrew Bennett. 2005. *Case Studies and Theory Development in the Social Sciences.* Massachusetts: Belfer Center Studies in International Security.
33. Hall, Peter. 1986. *Governing the Economy: The Politics of State Intervention in Britain and France*. Oxford: Oxford University Press.
34. Hall, Peter. 1993. "Policy Paradigms, Social Learning, and the State: The Case of Economic Policymaking in Britain." *Comparative Politics* 25 (3): 275-296.
35. Hall, Peter. 2003. "Aligning Ontology and Methodology in Comparative Politics." In *Comparative Historical Analysis in the Social Sciences*, edited by James Mahoney and Dietrich Rueschemeyer, 373-404. New York: Cambridge University Press.
36. Hall, Peter. 2013. "Brother, Can You Paradigm?" *Governance: An International Journal of Policy, Administration, and Institutions* 26 (2): 189-192. http://dx.doi.org/10.1111/gove.12031
37. Hartley, Jean. 2005. "Innovation in Governance and Public Services: Past and Present." *Public Money and Management* 25 (1) 27-34.
38. Hildén, Mikael, Andrew Jordan and Tim Rayner. 2014. "Climate Policy Innovation: Developing an Evaluation Perspective." *Environmental Politics* 23 (5): 884-905 : http://dx.doi.org/10.1080/09644016.2014.924205
39. Hood, Christopher. 1983. *The Tools of Government: Public Policy and Politics.* London: Palgrave Macmillan.
40. Howlett, Michael. 2009. "Governance Modes, Policy Regimes and Operational Plans: A Multi-Level Nested Model of Policy Instrument Choice and Policy Design." *Policy Sciences* 42 (1): 73-89.
41. Howlett, Michael. 2011. *Designing Public Policies: Principles and Instruments.* New York: Routledge.
42. Howlett, Michael. 2014. "Why Are Policy Innovations Rare and So Often Negative? Blame Avoidance and Problem Denial in Climate Change Policy-Making." *Global Environmental Change* 29: 395-403. https://doi.org/10.1016/j.gloenvcha.2013.12.009
43. Howlett, Michael and Benjamin Cashore. 2009. "The Dependent Variable Problem in the Study of Policy Change: Understanding Policy Change as a Methodological Problem." *Journal of Comparative Policy Analysis: Research and Practice* 11 (1): 33-46. http://dx.doi.org/10.1080/13876980802648144
44. Howlett, Michael and Rayner, Jeremy. 2013. "Patching Vs Packaging in Policy Formulation: Assessing Policy Portfolio Design." *Politics and Governance* 1 (2): 170-182.
45. Howlett, Michael, M. Ramesh and Anthony Perl. 2009. *Studying Public Policy: Policy Cycles and Subsystems.* Oxford: Oxford University Press.
46. Howson, Colin and Peter Urbach. 2006. *Scientific Reasoning: The Bayesian Approach.* Illinois: Carus Publishing Company.
47. Humphreys, Macartan and Alan Jacobs. 2015. "Mixing Methods: A Bayesian Approach." *American Political Science Review* 109 (4): 653-673. https://doi.org/10.1017/S0003055415000453
48. Jackson, Patrick Thaddeus. 2016. *The Conduct of Inquiry in International Relations: Philosophy of Science and Its Implications for the Study of World Politics*. London: Routledge.
49. Jordan, Andrew and Dave Huitema. 2014a. "Policy Innovation in a Changing Climate: Sources, Patterns and Effects." *Global Environmental Change* 29: 387- 394. https://doi.org/10.1016/j.gloenvcha.2014.09.005
50. Jordan, Andrew and Dave Huitema. 2014b. "Innovations in Climate Policy: Conclusions and New Directions." *Environmental Politics* 23 (5): 906- 925. http://dx.doi.org/10.1080/09644016.2014.924209
51. Kingdon, John. 1995. *Agendas, Alternatives, and Public Policies.* New York: Longman.
52. Kuhn, Thomas. 1971. *La estructura de las revoluciones científicas.* Mexico: Fondo de Cultura Económica.
53. Lascoumes, Pierre and Patrick Le Galès. 2007. "Introduction: Understanding Public Policy through Its Instruments from the Nature of Instruments to the Sociology of Public Policy Instruments." *International Journal of Policy, Administration, and Institutions* 20 (1): http://dx.doi.org/1-21. 10.1111/j.1468-0491.2007.00342.x
54. Lynn, Laurence. 1997. "Innovation and the Public Interest: Insights from the Private Sector." In *Innovation in*

*American Government: Challenges, Opportunities, and Dilemmas*, edited by Alan Altschuler and Robert Behn, 83-103. Washington: Brookings Institution Press.
55. Mackie, J. L. 1965. "Causes and Conditions." *American Philosophical Quarterly* 2 (4): 245–264.
56. Mahoney, James. 2000. "Path Dependence in Historical Sociology." *Theory and Society* 29 (4): 507-548.
57. Mejía, Luis Bernardo. 2014. *The Changing Role of the Central Planning Offices in Latin America: A Comparative Historical Analysis Perspective (1950-2013).* Maastricht: Boekenplan.
58. Moore, Mark, Malcom Sparrow and William Spelman. 1997. "Innovation in Policing: From Production Lines to Jobs Shops." In *Innovation in American Government: Challenges, Opportunities, and Dilemmas*, edited by Alan Altshuler and Robert Behn, 274-298. Washington: Brookings Institution Press.
59. OECD. 2005. *Oslo Manual: Guidelines for Collecting and Interpreting Innovation Data.* Paris: OECD Publishing.
60. OEC. 2016. "*Exportaciones Colombia*." http://goo.gl/OYR9AF
61. O'Hagan, Anthony and Mike West (eds.). 2010. *The Oxford Handbook of Applied Bayesian Analysis*. Oxford: Oxford University Press.
62. Paz, Bayron and Edwin Alvarado. 2016. "Documento de trabajo: Anotaciones para un protocolo de formalización bayesiana de inferencia causal en process tracing." *Grupo de investigación en políticas públicas comparadas. Facultad Latinoamericana de Ciencias Sociales-Ecuador*.
63. Powell, Walter and Paul Dimaggio. 1999. *El nuevo institucionalismo en el análisis organizacional.* Mexico: Fondo de Cultura Económica.
64. Polsby, Nelson. 1984. *Political Innovation in America*. New Haven: Yale University Press.
65. Rogers, Everett. 1983. *Diffusion of Innovations.* New York: Free Press.
66. Reta, Carlos. 2016. "Presentación." In *La Innovación en el Sector Público: tendencias internacionales y experiencias mexicanas*, edited by Guillermo Cejudo, Mauricio Dausssage and Cynthia Michel, 11-16. Mexico: CIDE – INAP.
67. Schaffrin, André, Sebastian Sewerin and Sibylle Seubert. 2014. "The Innovativeness of National Policy Portfolios – Climate Policy Change in Austria, Germany, and the UK." *Environmental Politics* 23 (5): 860-883.
68. Schumpeter, Joseph. 1997. *Teoría del desenvolvimiento económico.* México: Fondo de Cultura Económica. http://dx.doi.org/10.1080/09644016.2014.924206
69. Setnikar, Cankar and Veronika Petkovšek. 2013. "Private and Public Sector Innovation and the Importance of Cross-Sector Collaboration." *The Journal of Applied Business Research* 29 (6): 1597-1605.
70. Thelen, Kathleen Ann. 2004. *How Institutions Evolve: The Political Economy of Skills in Germany, Britain, the United States, and Japan.* Cambridge: Cambridge University Press.
71. True, James, Bryan Jones and Frank Baumgartner. 2007. "Punctuated-Equilibrium Theory: Explaining Stability and Change in Public Policymaking." In *Theories of the Policy Process*, edited by Paul Sabatier, 155-187. Cambridge: Westview Press.
72. Walker, Jack. 1969. "The Diffusion of Innovations among States." *The American Political Science Review* 63 (3): 880-899.

**Annex 1.** A causal mechanism of policy innovation

| Level | | Trigger | | Entities | | | Outcome |
|---|---|---|---|---|---|---|---|
| Conceptualization | | T=T1+T2 | | A | B | C | O |
| | | Shift in the locus of authority=New agencies+Policy window opening | | New policy network discrediting *status quo ante* | New paradigm assumed by policy network overlooking policy area | New policy network influencing the design of a new policy | Government implementing policy innovation |
| Operational hypotheses | | $h_{T1}$ | $h_{T2}$ | $h_A$ | $h_B$ | $h_C$ | $h_O$ |
| | | Enhancement of new technical and economic agencies advocating for a centralized SGR | Election of a new President opened a policy window for technical and economic agencies advocating for a centralized SGR | Diffusion of a discourse discrediting the *status quo ante* at the highest level of the State | Spreading out of the core ideas of new policy network in the policy area | New policy instruments influenced by core ideas of the new policy network | Implementation of inedited, disruptive and uncertain reforms in the policy area |
| Alternate hypotheses | | $\neg h_{T1}$ | $\neg h_{T2}$ | $\neg h_A$ | $\neg h_B$ | $\neg h_C$ | $\neg h_O$ |
| | | No enhancement of new technical and economic agencies advocating for a centralized SGR | Election of a new President did not open a policy window for technical and economic agencies advocating for a centralized SGR | No diffusion of a discourse discrediting the *status quo ante* at the highest level of the State | No spreading out of the core ideas of new policy network in the policy area | New policy instruments not influenced by core ideas of the new policy network | No implementation of inedited, disruptive and uncertain reforms in the policy area |
| Empirical prediction or tests of inferential strength | Hoop tests | $P_{T1}1$: Local governments condemn centralism | $P_{T2}1$: New President's economist background was akin to the nature of technical and economic agencies | $P_A1$: Local governments condemn strategy of discredit | $P_B1$: New policy network leads debates on policy change | $P_C1$: Local governments condemn concentration of resources by the government | $P_O1$: Technical agencies praise the designed policy for its novelty |
| | Smoking-gun tests | $P_{T1}2$: Policy instruments of organization show presence of new technical and economic agencies | | $P_A2$: Policy instruments of nodality show discredit of *status quo ante* | $P_B2$: Ideas of new agencies become part of State's highest values | $P_C2$: Policy instruments of authority affirm centralization of SGR | |
| | Doubly- decisive test | | | | | | $P_O2$: New policy implements hitherto unseen changes in aims and means |

*Source:* Elaborated by the authors.

**Annex 2.** Bayesian formalization of the causal mechanism of policy innovation

| Empirical tests | Hoop | Smoking gun | Hoop | Hoop | Smoking gun | Hoop | Smoking gun | Hoop | Smoking gun | Hoop | Doubly decisive |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | $PT_1 1$ | $PT_1 2$ | $PT_2 1$ | $P_A 1$ | $P_A 2$ | $P_B 1$ | $P_B 2$ | $P_C 1$ | $P_C 2$ | $P_O 1$ | $P_O 2$ |
| p(h) | 0,50 | 0,50 | 0,50 | 0,50 | 0,50 | 0,50 | 0,50 | 0,50 | 0,50 | 0,50 | 0,50 |
| p(¬h) | 0,50 | 0,50 | 0,50 | 0,50 | 0,50 | 0,50 | 0,50 | 0,50 | 0,50 | 0,50 | 0,50 |
| p(e/h) | 0,87 | 0,37 | 0,60 | 0,91 | 0,52 | 0,88 | 0,33 | 0,74 | 0,33 | 0,79 | 0,95 |
| p(¬e/h) | 0,13 | 0,63 | 0,40 | 0,09 | 0,48 | 0,12 | 0,67 | 0,26 | 0,67 | 0,21 | 0,05 |
| p(e/¬h) | 0,57 | 0,13 | 0,38 | 0,65 | 0,01 | 0,36 | 0,01 | 0,42 | 0,01 | 0,76 | 0,01 |
| p(¬e\|¬h) | 0,43 | 0,87 | 0,62 | 0,35 | 0,99 | 0,64 | 0,99 | 0,58 | 0,99 | 0,24 | 0,99 |
| p(h/e) | 0,60 | 0,74 | 0,61 | 0,58 | 0,98 | 0,71 | 0,97 | 0,64 | 0,97 | 0,51 | 0,99 |
| p(h\|¬e) | 0,23 | 0,42 | 0,39 | 0,20 | 0,33 | 0,16 | 0,40 | 0,31 | 0,40 | 0,47 | 0,05 |
| C=p(h\|e)-p(h) | 0,10 | 0,24 | 0,11 | 0,08 | 0,48 | 0,21 | 0,47 | 0,14 | 0,47 | 0,01 | 0,49 |
| LR(h)= p(e/h)/ p(e/¬h) | 1,53 | 2,85 | 1,58 | 1,40 | 52,00 | 2,44 | 33,00 | 1,76 | 33,00 | 1,04 | 95,00 |

*Source:* Elaborated by the authors, based on Fairfield and Charman, 2017; Bennett, 2015.