



**HAL**  
open science

# A market regulation bilevel problem: a case study of the Mexican petrochemical industry

Héctor Maravillo, José-Fernando Camacho-Vallejo, Justo Puerto, Martine Labbé

## ► To cite this version:

Héctor Maravillo, José-Fernando Camacho-Vallejo, Justo Puerto, Martine Labbé. A market regulation bilevel problem: a case study of the Mexican petrochemical industry. *Omega*, In press, 10.1016/j.omega.2019.08.012 . hal-01922518

**HAL Id: hal-01922518**

**<https://inria.hal.science/hal-01922518>**

Submitted on 14 Nov 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# A market regulation bilevel problem: a case study of the Mexican petrochemical industry

Héctor Maravillo<sup>a</sup>, José-Fernando Camacho-Vallejo<sup>a,\*</sup>, Justo Puerto<sup>b</sup>, Martine Labbé<sup>c</sup>

<sup>a</sup>*Facultad de Ciencias Físico-Matemáticas, Universidad Autónoma de Nuevo León, Monterrey, Mexico*

<sup>b</sup>*Department of Statistics and Operational Research & Instituto de Matemáticas Universidad de Sevilla, Universidad de Sevilla, Seville, Spain*

<sup>c</sup>*Département d'Informatique, Université Libre de Bruxelles, Brussels, Belgium and INRIA, Lille, France*

---

## Abstract

In this paper, a bilevel programming model is proposed to study a problem of market regulation through government intervention. One of the main characteristics of the problem herein analyzed is that the government monopolizes the raw material in one industry, and competes in another industry with private firms for the production of commodities. Under this scheme, the government controls a state-owned firm to balance the market; that is, to minimize the difference between the produced and demanded commodities. On the other hand, a regulatory organism that coordinates private firms aims to maximize the total profit by deciding the amount of raw material bought from the state-owned firm. Two equivalent single-level reformulations are proposed to solve the problem. The first reformulation is based on the strong duality condition of the lower level and results in a continuous non-linear model. The second reformulation resorts to the complementarity slackness optimality constraints yielding a mixed-integer linear model. Additionally, three heuristic algorithms are designed to obtain good-quality solutions with low computational effort. In this problem, the feasible region of the dual problem associated to the follower is independent from the leader's decision. Therefore, the proposed heuristics exploit this particular characteristic of the bilevel model. Moreover, the third heuristic hybridizes the other two algorithms to enhance its performance. Extensive computational experimentation is carried out to measure the efficiency of the proposed solution methodologies. A case study based on the Mexican petrochemical industry is presented. Additional instances generated from the case study are considered to validate the robustness of the proposed heuristic algorithms. Numerical results indicate that the hybrid algorithm outperforms the other two heuristics. However, all of them demonstrate to be good alternatives for solving the problem. Additionally, optimal solutions of all the instances are obtained by using good quality solutions (given by the hybrid algorithm) as initial solutions when solving the second reformulation via a general purpose solver.

*Keywords:* Market regulation, Bilevel programming, Petrochemical industry

---

## 1. Introduction

Market regulation through government intervention has appeared in different ways, such as, fiscal policies (taxes and subsidies), anti-monopoly legislation, price control, quantity control, and nationalization of firms [1]. Two of the main motivations for regulating the economy is to seek social equity and to handle market failures [2]. A specific manner of government intervention is through the participation of state-owned firms, which are under the government control and compete with the private firms. The goal to regulate the market by participation has been very noticeable in developing countries, mainly in raw materials industries [3]. For instance, a list of countries in which market regulation by participation exists in oil industry is presented in [4], [5], [6]. At this point we would like to emphasize that this situation occurred in Mexico in the petrochemical industry from 1958 until 2014.

The first paper that considers the idea of regulate the market by including a state-owned firm is [7]. In that paper, a short-term analysis that concerns with the entrance of a state-owned firm into a three-firms oligopoly market is done. They conclude that the existence of a state-owned firm may improve the performance of the market, which it is also shown in [8] in the aerospace market. Following up with the existence of state-owned firms, [9] considers a situation in which the difference between the production and the production level is made up by the government. The novelty in that paper is to realize that state-owned firm acts as the dominant decision-maker, that is, it has complete information of the market and announces its decision. Hence, each private firm reacts to this decision by establishing its output level so that its marginal cost equals the price. Furthermore, [10] explicitly considers the state-owned firm as a leader in a Stackelberg game. Later, in [11], a mixed oligopoly model that helps to compare the differences among the nationalization of a state-owned firm and the entrance of a new state-owned firm is studied. This last paper also analyzes the cost effectiveness of the state-owned firms in the long-term and moreover, the authors consider for the first time a budgetary constraint associated with the state-owned firm to guarantee a minimum profit.

Literature reviews regarding mixed oligopolies can be found in [12] and [13]. In those reviews the common characteristics of these models are identified. Particularly, the context of the problem, the cases when the government provides complete information, the goal of the state-owned firm, the technology assumptions, the costs structure, and the hierarchy among the firms (in case of Stackelberg games). Another literature review that deals with some foundations for a theory of mixed oligopoly markets is presented in [14]. It is important to remark that all the above mentioned papers coincide in that there is a lack of an unified and accepted general mixed oligopolies modeling framework because each model pursues its own characteristics and set different basic hypothesis. At this point, it is also important to

---

\*I am corresponding author

*Email addresses:* hmaravillo@hotmail.com (Héctor Maravillo), jose.camachov1@uanl.edu.mx (José-Fernando Camacho-Vallejo ), puerto@us.es (Justo Puerto), mlabbe@ulb.ac.be (Martine Labbé)

39 emphasize that, to the best of our knowledge and after an intensive literature review, the  
40 intervention of the government via state-owned firms to regulate an industry conformed by  
41 two interrelated markets has not been studied before. However, these characteristics appear  
42 in real situations, as it will be described in the case study herein considered.

43  
44 The problem studied in this paper is inspired by a situation where a state-owned firm  
45 controls the monopoly of a market but this is interrelated with a second market in which  
46 state-owned and private firms compete. The problem considers two interdependent indus-  
47 tries, in which the first one produces the necessary supplies for the second one. The gov-  
48 ernment has the monopoly of the first industry, while in the second industry a state-owned  
49 firm competes with remaining private firms in the production of commodities. Private firms  
50 aim to maximize their profit. Government aims as a general social welfare to balance the  
51 market. The latter is achieved by minimizing the difference between the supply and demand  
52 of the final commodities.

53  
54 Given the privileged position that the government occupies in this market as the leader  
55 decision-maker, the government decides the amount of commodities that will be produced at  
56 the state-owned firms and the amount of supplies that will be offered to the private firms. By  
57 doing this, the government will regulate indirectly the production. One can easily appreciate  
58 a hierarchical relationship between both economic agents, in which the government makes a  
59 decision and the private firms reacts to it affecting the balance in the market. Furthermore,  
60 technical and technological issues in the raw material market limit the production, for ex-  
61 ample, the scarcity of natural resources and the existing production capacity of the firms.  
62 The hierarchy and inter-relationship among the actors of this situation allow to formulate a  
63 bilevel programming model, in which the government will act as the leader and the private  
64 firms as the follower. Other bilevel models in which the government intervenes to regulate  
65 social aspects are [15] and [16]. In the former, the government employs an intervention pol-  
66 icy based on subsidies in the automotive market, while in the latter, the government apply  
67 taxes via an agro-environmental policy imposed to the agriculturists. Additionally, the gen-  
68 eral model proposed by us perfectly fits to the petrochemical industry in several developing  
69 countries in the last decades and it was the case of Mexico for almost 56 years, where it was  
70 apparently run without a theoretically recognizable model.

71  
72 The main contributions of this paper can be listed as follows: a novel mathematical  
73 bilevel programming model to study a mixed-oligopoly market, three heuristic algorithms  
74 based on an iterative exploration of vertices in the inducible region of the bilevel problem, an  
75 extensive computational study analyzing the performance of our exact reformulations and  
76 the proposed heuristic algorithms, and a case-study based on the Mexican petrochemical  
77 industry.

78  
79 The remainder of the paper is organized as follows. Section 2 defines the mathematical  
80 programming bilevel model and sets the notation. Two reformulations of the bilevel model  
81 that reduce it to a single-level one are presented in section 3. The first one is a continuous

82 non-linear problem, and the second one reduces to a mixed integer linear program. In  
83 both cases, the resulting programs are hard to solve for medium to large sizes. Then,  
84 section 4 describes three proposed heuristic algorithms to provide good quality solutions  
85 for the bilevel model. Section 5 reports the numerical results according to the case study  
86 under consideration, and compares the results obtained through the reformulations and  
87 the algorithms. Also, the results obtained from additional experimentation with random  
88 instances are summarized. Conclusions and recommendations for future research are given  
89 in section 6.

## 90 **2. A market regulation bilevel problem**

### 91 *2.1. Problem's description*

92 The problem herein studied considers an industry conformed by two economic markets,  
93 one of them associated to raw material and the other one of the final commodities. The  
94 supplies (raw material) of the second economic market are produced in the first one. In  
95 this industry, there is a state-owned firm vertically integrated, that is, that produces in  
96 both economic markets. The state-owned firm monopolizes the production of the supplies  
97 in the first economic market, but in the other one, this firm competes against the private  
98 firms. All the firms have a maximum production capacity and the state-owned firm requires  
99 to obtain a minimum profit (there is a lower bound over its net profit). The objective of  
100 this latter firm is to balance the market by its intervention. To achieve this goal, it mini-  
101 mizes the lack and surplus of the offered commodities with respect to the demand. Hence,  
102 this firm determines its production of commodities and the amount of supplies to be offered  
103 to the private firms. On the other hand, private firms' goal will be to maximize their benefit.  
104

105 The decisions taken by the state-owned firm limit the admissible production by the pri-  
106 vate firms, and the decisions of the private firms affect the achievement of the government.  
107 As it is mentioned before, the government has the monopoly of the supplies production, and  
108 it is assumed that there exists a centralized organism that regulates the supplies demand.  
109 This organism determines the amount of supplies to each private firm, which is common  
110 in economic markets with government intervention. A typical example occurs when the  
111 government creates organisms to regulate the competition among firms and establishes par-  
112 ticular contracts with each private firm fixing the supplies to be sold to each one. Another  
113 case is when the private firms get together and create a centralized mechanism to which  
114 the responsibility of distributing supplies among them is delegated. In the latter case, that  
115 mechanism seeks for a global welfare of the market; this is the case of common lands coop-  
116 erative organization in agriculture. Under this scheme, the government will be the leader  
117 and the centralized mechanism the follower.

### 118 *2.2. Mathematical formulation*

119 In this section, the mathematical formulation of the problem described in the previous  
120 section is formally introduced. Let  $I$  be the set of commodities and let  $J$  be the set of  
121 private firms. Each commodity  $i \in I$  is sold at a price  $p_i$ , and it has a demand  $d_i$  in the

122 market. To produce a commodity  $i \in I$ , it costs  $c_i^G$  to the state-owned firm and  $c_{ij}^E$  to a  
123 private firm  $j \in J$ . The government fixes a minimum profit  $t$  to be obtained by the state-  
124 owned firm. The amount of primary resources that a private firm  $j \in J$  needs to produce  
125 a unit of commodity  $i \in I$  is denoted by  $b_{ij}$  and the amount of supply required to produce  
126 a unit of commodity  $i \in I$  is denoted by  $a_{ij}$ . Both types of firms have a limited production  
127 capacity. The state-owned firm has a maximum production capacity  $q_i^B$  to obtain supplies  
128 of the commodity  $i \in I$  and a maximum amount  $q_i^A$  to produce for each commodity  $i \in I$ .  
129 Also, each private firm  $j \in J$  has a maximum production capacity  $m_j$ .

130  
131 In order to present our mathematical programming formulation, we will use the following  
132 decision variables. The leader's decision variables are:  
133  $x_i$ , which denotes the production of the state-owned firm for each commodity  $i \in I$ .  
134  $z_i$ , which denotes the supply offered by the government to the private firms to produce a  
135 commodity  $in \in I$ .

136  
137 The follower's decision variables are:  
138  $y_{ij}$ , which denotes the amount produced by private firm  $j \in J$  of a commodity  $i \in I$ .

139  
140 In our model, non-negative auxiliary variables are introduced to express the leader's ob-  
141 jective function. Let  $r_i$  be the shortage of commodity  $i \in I$  and let  $s_i$  be the corresponding  
142 surplus.

143  
144 With the above elements, the proposed bilevel programming formulation to model the  
145 regulation of the economic market described above results as follows:

$$\min_{x,z,r,s,y} \sum_{i \in I} (r_i + s_i) \quad (1)$$

$$\text{s.t.} \frac{\sum_{j \in J} y_{ij} + x_i}{d_i} + r_i - s_i = 1 \quad \forall i \in I \quad (2)$$

$$\sum_{i \in I} (p_i - c_i^G) x_i \geq t \quad (3)$$

$$0 \leq x_i \leq q_i^A \quad \forall i \in I \quad (4)$$

$$0 \leq z_i \leq q_i^B \quad \forall i \in I \quad (5)$$

$$r_i \geq 0 \quad \forall i \in I \quad (6)$$

$$s_i \geq 0 \quad \forall i \in I \quad (7)$$

$$y \in \operatorname{argmax} \sum_{j \in J} \sum_{i \in I} (p_i - c_{ij}^E) \bar{y}_{ij} \quad (8)$$

$$\text{s.t.} \sum_{j \in J} a_{ij} \bar{y}_{ij} \leq z_i \quad \forall i \in I \quad (9)$$

$$\sum_{i \in I} b_{ij} \bar{y}_{ij} \leq m_j \quad \forall j \in J \quad (10)$$

$$\bar{y}_{ij} \geq 0 \quad \forall i \in I, \quad \forall j \in J \quad (11)$$

146 The leader's problem is defined by (1)-(8), in which (1) represents the leader's objective  
 147 function that aims to minimize the inefficiency of the market, namely the overall shortage  
 148 and surplus production with respect to the total demand. Constraint (2) seeks to balance the  
 149 demand of each commodity with the corresponding production of the state-owned and pri-  
 150 vate firms. In (3), it is ensured that the state-owned firm obtains a minimum required profit  
 151 from its commodities production. Constraints (4) and (5) establish the production capacity  
 152 associated with the commodities and offered supplies, respectively. The non-negativity of  
 153 the auxiliary variables is expressed in (6) and (7). Constraint (8) plays a key role in this  
 154 model, due to the fact that it requires that the value of the follower's variables is given by the  
 155 optimal solution of another mathematical programming problem. The follower's problem  
 156 is defined by equations (8)-(11), which aims to maximize the total profit of all the private  
 157 firms. The production of the private firms is limited by the amount of supplies provided by  
 158 the government, this is enforced by (9). Constraint (10) imposes that each private firm's  
 159 production cannot exceed its maximum capacity. Finally, (11) expresses the non-negativity  
 160 of the follower's variables.

161  
 162 In order to have a well-defined formulation for the proposed bilevel model, it is neces-  
 163 sary to make an assumption regarding multiple optimal solutions that may appear in the  
 164 follower's problem. We assume the optimistic approach defined in [17]. In other words,  
 165 among all the optimal follower's solutions, the one that minimizes the leader's objective  
 166 function is selected. The optimistic approach is suitable for the situation under study since  
 167 it may express the existence of a certain cooperation degree among the government and the  
 168 mechanism (committee) that represents the private firms.

### 169 3. Exact solution methodologies

170 A common methodology that applies, at times, to solve a bilevel programming model is  
 171 to transform it to a single-level reformulation. In order to achieve this goal, the character-  
 172 istics of the follower's problem have to be exploited. Once a leader's solution is established,  
 173 the leader's variables are fixed in the follower's problem. Particularly,  $z$  is going to be con-  
 174 sidered as a parameter in the problem defined by (8)-(11). Therefore, the follower's problem  
 175 results in a linear programming problem that has a corresponding dual, in which  $\alpha$  and  $\beta$   
 176 correspond to the associated dual variables. The dual problem associated with the follower's  
 177 problem is as follows:

$$\min_{\alpha, \beta} \sum_{i \in I} \alpha_i z_i + \sum_{j \in J} \beta_j m_j \quad (12)$$

$$\text{s.t.} \quad a_{ij} \alpha_i + b_{ij} \beta_j \geq (p_i - c_{ij}^E) \quad \forall i, \quad \forall j \quad (13)$$

$$\alpha_i \geq 0 \quad \forall i \quad (14)$$

$$\beta_j \geq 0 \quad \forall j \quad (15)$$

179 For both, the primal and dual of the follower's problem, the space of feasible solutions  
180 is bounded. Hence, the fundamental duality theorem states that both problems have op-  
181 timal solutions and their objective function values are equal [18]. Hence, two single-level  
182 reformulations that are equal to the bilevel model are presented in this section. The first  
183 reformulation ensures that the optimal solution of the reformulated model are in the in-  
184 ducible region of the bilevel problem by using a corollary of the strong duality theorem.  
185 The second reformulation substitutes a non-linear constraint of the first reformulation via  
186 the complementarity slackness constraint. The first reformulation results in a continuous  
187 non-linear model and the second one is a mixed-integer linear model.

### 188 3.1. Non-linear reformulation based on the strong duality condition

The first reformulation consists in adding the constraints associated with the dual of the follower's problem, that is, (12-15), and to use the necessary and sufficient optimality conditions to ensure that the follower's decision belongs to the set of rational reactions. Therefore, a constraint that equals the follower's primal and dual objective functions is added. For a predefined leader's fixed value of  $z$  the resulting constraint is as follows:

$$\sum_{j \in J} \sum_{i \in I} (p_i - c_{ij}^E) y_{ij} = \sum_{i \in I} \alpha_i z_i + \sum_{j \in J} \beta_j m_j \quad (16)$$

The first proposed reformulation consists in the following non-linear single-level problem:

$$\min_{x, z, r, s, y, \alpha, \beta} \sum_{i \in I} (r_i + s_i) \quad (\text{Ref.1})$$

$$\text{s.t. } \frac{\sum_{j \in J} y_{ij} + x_i}{d_i} + r_i - s_i = 1 \quad \forall i \in I \quad (17)$$

$$\sum_{i \in I} (p_i - c_i^G) x_i \geq t \quad (18)$$

$$0 \leq x_i \leq q_i^A \quad \forall i \in I \quad (19)$$

$$0 \leq z_i \leq q_i^B \quad \forall i \in I \quad (20)$$

$$r_i \geq 0 \quad \forall i \in I \quad (21)$$

$$s_i \geq 0 \quad \forall i \in I \quad (22)$$

$$\sum_{j \in J} a_{ij} y_{ij} \leq z_i \quad \forall i \in I \quad (23)$$

$$\sum_{i \in I} b_{ij} y_{ij} \leq m_j \quad \forall j \in J \quad (24)$$

$$y_{ij} \geq 0 \quad \forall i \in I, j \in J \quad (25)$$

$$a_{ij} \alpha_i + b_{ij} \beta_j \geq (p_i - c_{ij}^E) \quad \forall i \in I, j \in J \quad (26)$$

$$\alpha_i \geq 0 \quad \forall i \in I \quad (27)$$

$$\beta_j \geq 0 \quad \forall j \in J \quad (28)$$



$$\sum_{j \in J} \sum_{i \in I} (p_i - c_{ij}^E) y_{ij} = \sum_{i \in I} \alpha_i z_i + \sum_{j \in J} \beta_j m_j \quad (29)$$

189 Note that the linearity of the model is lost in (29). Further, it can be represented in the  
190 following manner since weak duality implies that the other inequality is always satisfied.

$$-\sum_{j \in J} \sum_{i \in I} (p_i - c_{ij}^E) y_{ij} + \sum_{i \in I} \alpha_i z_i + \sum_{j \in J} \beta_j m_j \leq 0 \quad (30)$$

### 191 3.2. Mixed-integer linear reformulation based on complementarity slackness

192 The second reformulation is based on the complementarity slackness constraints that  
193 ensures the optimality of the follower's problem. A single-level reformulation is recovered  
194 by introducing the following constraints:

$$\alpha_i \left( z_i - \sum_{j \in J} a_{ij} y_{ij} \right) = 0 \quad \forall i \in I \quad (31)$$

$$\beta_j \left( m_j - \sum_{i \in I} b_{ij} y_{ij} \right) = 0 \quad \forall j \in J \quad (32)$$

$$y_{ij} (a_{ij} \alpha_i + b_{ij} \beta_j - (p_i - c_{ij}^E)) = 0 \quad \forall i \in I, j \in J \quad (33)$$

195 Constraints (31), (32), and (33) are non-linear, nevertheless they can be linearized in a  
196 straightforward manner by introducing positive big- $M$  constants  $M_i^1$ ,  $M_i^2$ , and  $M_i^3$  for all  
197  $i \in I$ . Also, binary variables  $\gamma_i$ ,  $\delta_j$ , and  $\epsilon_{ij}$  for all  $i \in I$ ,  $j \in J$  are included. Constraints  
198 (31)-(33) are replaced by the following constraints:

$$\alpha_i \leq M_i^1 \gamma_i \quad \forall i \in I \quad (34)$$

$$z_i - \sum_{j \in J} a_{ij} y_{ij} \leq M_i^1 (1 - \gamma_i) \quad \forall i \in I \quad (35)$$

$$\beta_j \leq M_i^2 \delta_j \quad \forall i \in I, j \in J \quad (36)$$

$$m_j - \sum_{i \in I} b_{ij} y_{ij} \leq M_i^2 (1 - \delta_j) \quad \forall i \in I, j \in J \quad (37)$$

$$y_{ij} \leq M_i^3 \epsilon_{ij} \quad \forall i \in I, \forall j \in J \quad (38)$$

$$a_{ij} \alpha_i + b_{ij} \beta_j - (p_i - c_{ij}^E) \leq M_i^3 (1 - \epsilon_{ij}) \quad \forall i \in I, j \in J \quad (39)$$

Therefore, this reformulation results in the following single-level linear mixed-integer programming model.

$$\min_{x, z, r, s, y, \alpha, \beta, \gamma, \delta, \epsilon} \sum_{i \in I} (r_i + s_i) \quad (\text{Ref.2})$$

$$\text{s.t. } \frac{\sum_{j \in J} y_{ij} + x_i}{d_i} + r_i - s_i = 1 \quad \forall i \in I \quad (40)$$

$$\sum_{i \in I} (p_i - c_i^G) x_i \geq t \quad (41)$$

$$0 \leq x_i \leq q_i^A \quad \forall i \in I \quad (42)$$

$$0 \leq z_i \leq q_i^B \quad \forall i \in I \quad (43)$$

$$r_i \geq 0 \quad \forall i \in I \quad (44)$$

$$s_i \geq 0 \quad \forall i \in I \quad (45)$$

$$\sum_{j \in J} a_{ij} y_{ij} \leq z_i \quad \forall i \in I \quad (46)$$

$$\sum_{i \in I} b_{ij} y_{ij} \leq m_j \quad \forall j \in J \quad (47)$$

$$y_{ij} \geq 0 \quad \forall i, j \in J \quad (48)$$

$$a_{ij} \alpha_i + b_{ij} \beta_j \geq (p_i - c_{ij}^E) \quad \forall i \in I, j \in J \quad (49)$$

$$\alpha_i \geq 0 \quad \forall i \in I \quad (50)$$

$$\beta_j \geq 0 \quad \forall j \in J \quad (51)$$

$$\alpha_i \leq M_i^1 \gamma_i \quad \forall i \in I \quad (52)$$

$$z_i - \sum_{j \in J} a_{ij} y_{ij} \leq M_i^2 (1 - \gamma_i) \quad \forall i \in I \quad (53)$$

$$\beta_j \leq M_j^1 \delta_j \quad \forall j \in J \quad (54)$$

$$m_j - \sum_{i \in I} b_{ij} y_{ij} \leq M_j^2 (1 - \delta_j) \quad \forall j \in J \quad (55)$$

$$y_{ij} \leq M_{ij}^1 \epsilon_{ij} \quad \forall i \in I, j \in J \quad (56)$$

$$a_{ij} \alpha_i + b_{ij} \beta_j - (p_i - c_{ij}^E) \leq M_{ij}^2 (1 - \epsilon_{ij}) \quad \forall i \in I, j \in J \quad (57)$$

$$\gamma_i, \delta_j, \epsilon_{ij} \in \{0, 1\} \quad \forall i \in I, j \in J \quad (58)$$

### 199 3.3. Adjusting the value of the big $M$ 's

200 In the problem under study, parameters  $a$  and  $b$  are always positives, which means that  
 201 more commodity production exists, more supplies are required and more production capacity  
 202 is used in the private firms. Moreover, if term  $p_i - c_{ij}^E$  is negative, private firms are having  
 203 losses, which would lead to their exit from the economic market.

204 **Definition.** Define the upper bound  $UB(\cdot) \in \mathbb{R}$  of a vector as a real number that is  
 205 greater than or equal to all of the components of that vector.

206 To adjust the value of  $M_i^1$  for all  $i \in I$ , first we seek for an upper bound for  $\alpha_i$ , such that  
 207  $\alpha$  belongs to an optimal solution of the follower's dual problem. The worst scenario occurs  
 208 when  $\beta_j = 0 \forall j \in J$ , then  $\alpha_i \geq \frac{p_i - c_{ij}^E}{a_{ij}}$  for all  $i \in I, \forall j \in J$  must hold due to (49). Since (12)  
 209 aims to minimize, the optimal solution occurs at equality. To obtain the upper bound for  
 210  $M_i^2$  for all  $i \in I$ , constraint (53) is considered. Hence, it is possible to state the following  
 211 proposition.

212 **Proposition 1.**

213 1. The upper bound for any optimal value of  $\alpha_i$  for all  $i \in I$  is as follows:

214

$$215 \quad UB(\alpha_i) = \max_{j \in J} \left\{ \frac{p_i - c_{ij}^E}{a_{ij}} \right\}$$

216 2. The upper bound for constraint (53) is as follows:

217

$$218 \quad UB \left( z_i - \sum_{j \in J} a_{ij} y_{ij} \right) = q_i^B.$$

219 Analogously that for  $M_i^k$ , the upper bounds for  $M_j^k$   $i = 1, 2$  and  $k = 1, 2$  are computed.

220 **Proposition 2.**

221 1. The upper bound for any optimal value of  $\beta_j$  for all  $j \in J$  is:

222

$$222 \quad UB(\beta_j) = \max_{i \in I} \left\{ \frac{p_i - c_{ij}^E}{b_{ij}} \right\}.$$

223 2. The upper bound for constraint (55) is:

224

$$224 \quad UB \left( m_j - \sum_{i \in I} b_{ij} y_{ij} \right) = m_j.$$

225 To bound the values of  $M_{ij}^k$  for all  $i \in I, j \in J$  and  $k = 1, 2$ , the following proposition is  
226 stated.

227 **Proposition 3.**

228 1. The upper bounds for  $y_{ij}$  for all  $i \in I, j \in J$  are:

229

$$230 \quad UB(y_{ij}) = \frac{m_j}{b_{ij}},$$

231 *i.e.*  $y_{ij} \leq \frac{m_j}{b_{ij}}$  for all  $i \in I, j \in J$ .

232 2. The upper bound of constraint (57) is:

233

$$234 \quad UB(a_{ij}\alpha_i + b_{ij}\beta_j - (p_i - c_{ij}^E)) = a_{ij}UB(\alpha_i) + b_{ij}UB(\beta_j).$$

235 Therefore, the adjusted values for  $M_i^k, M_j^k$ , and  $M_{ij}^k$  for all  $i \in I, j \in J, k = 1, 2$  based  
236 on the computed upper bounds are as follows:

$$M_i^1 = \max_{j \in J} \left\{ \frac{p_i - c_{ij}^E}{a_{ij}} \right\}, \quad \forall i \in I \quad (59)$$

$$M_i^2 = q_i^B, \quad \forall i \in I \quad (60)$$

$$M_j^1 = \max_{i \in I} \left\{ \frac{p_i - c_{ij}^E}{b_{ij}} \right\}, \quad \forall j \in J \quad (61)$$

$$M_j^2 = m_j, \quad \forall j \in J \quad (62)$$

$$M_{ij}^1 = \frac{m_j}{b_{ij}}, \quad \forall i \in I, j \in J \quad (63)$$

$$M_{ij}^2 = a_{ij}UB(\alpha_i) + b_{ij}UB(\beta_j), \quad \forall i \in I, j \in J. \quad (64)$$

237 **4. Heuristic solution methodologies**

238 Usually, the reformulations introduced above present computational limitations for large-  
 239 size instances. Therefore, alternative approaches are required to solve the bilevel problem.  
 240 One common approach is to design a heuristic algorithm to obtain good quality feasible  
 241 solutions with lower computational burden. In this section, three tailor-made heuristic al-  
 242 gorithms are proposed to solve the problem under study. The first two heuristics exploit the  
 243 particular structure of the bilevel problem, while the third one hybridizes the other two. The  
 244 hybrid heuristic contains key aspects from the previously proposed heuristics to enhance its  
 245 performance.

246  
 247 Note that it has been already emphasized that the polyhedron that defines the feasible  
 248 region of the dual problem associated with the follower is independent from the leader's  
 249 variables. Therefore, if all the vertices of this polyhedron were known, equality (16) and  
 250 the dual constraints of the first reformulation described in section 3.1 could be substituted  
 251 by a set of constraints guaranteeing that the equality (16) would be achieved in one vertex.  
 252 Following the latter idea, the dual variables are replaced by parameters that represents those  
 253 known vertices. This is the approach exploited in the three proposed heuristic algorithms.

254  
 255 *4.1. Extreme points iterated algorithm (EPIA)*

256 In this algorithm, vertices of the dual polyhedron are iteratively generated, and a mixed-  
 257 integer programming problem named the Master Problem (MP) is solved for each new  
 258 vertex. The optimal solution of the MP ensures a feasible solution of the bilevel problem.  
 259 The algorithm stops when no improvement in the leader's objective function is obtained.

260  
 261 Let  $P$  be the polyhedron associated with the dual constraints of problem (13)-(15). Con-  
 262 sider  $v_k = (\alpha^k, \beta^k) \in P$ ,  $k = 1, \dots, |P|$  as the vertices of  $P$ , in which  $|P|$  define the number  
 263 of vertices in that polyhedron with  $\alpha^k = (\alpha_1^k, \dots, \alpha_{|I|}^k)$  and  $\beta^k = (\beta_1^k, \dots, \alpha_{|J|}^k)$ . Let  $F$  be  
 264 the leader's objective function (1). Also, let  $f$  and  $g$  be the follower primal (8) and dual  
 265 (12) objective functions, respectively. Consider  $\chi = (x, z, r, s)$  as a vector that groups the  
 266 leader's variables and  $\Upsilon = (y_1, \dots, y_{|J|})$  as a vector that groups the follower's variables. De-  
 267 fine  $\varphi$  as the set of vertices in  $P$  that has been already explored. At each iteration  $K = |\varphi|$   
 268 is updated using  $k \in K$  as the index of the explored vertices.  $UB(z_i)$  represents an upper  
 269 bound of variable  $z_i$ , which can be naturally fixed as  $q_i^B$ . Hence, the MP is defined as follows:  
 270

$$\min_{x, z, r, s, y, \lambda} \sum_{i \in I} (r_i + s_i) \quad (65)$$

$$\text{s.t.} \quad \frac{\sum_{j \in J} y_{ij} + x_i}{d_i} + r_i - s_i = 1 \quad \forall i \in I \quad (66)$$

$$\sum_{i \in I} (p_i - c_i^G) x_i \geq t \quad (67)$$

$$0 \leq x_i \leq q_i^A \quad \forall i \in I \quad (68)$$

$$0 \leq z_i \leq q_i^B \quad \forall i \in I \quad (69)$$

$$r_i \geq 0 \quad \forall i \in I \quad (70)$$

$$s_i \geq 0 \quad \forall i \in I \quad (71)$$

$$\sum_{j \in J} a_{ij} y_{ij} \leq z_i \quad \forall i \in I \quad (72)$$

$$\sum_{i \in I} b_{ij} y_{ij} \leq m_j \quad \forall j \in J \quad (73)$$

$$y_{ij} \geq 0 \quad \forall i \in I, j \in J \quad (74)$$

$$\sum_{i \in I} \alpha_i^k z_i + \sum_{j \in J} \beta_j^k m_j - (1 - \lambda_k) \left( \sum_{i \in I} \alpha_i^k UB(z_i) + \sum_{j \in J} \beta_j^k m_j \right) \leq \sum_{j \in J} \sum_{i \in I} (p_i - c_{ij}^E) y_{ij} \quad \forall k \in K \quad (75)$$

$$\sum_{k \in K} \lambda_k = 1 \quad (76)$$

$$\lambda_k \in \{0, 1\} \quad \forall k \in K \quad (77)$$

271 Constraints (75)-(77) guarantee that  $f(\Upsilon) = g(\chi, v_k)$  holds only for a single vertex, in  
 272 which  $\lambda$  is an auxiliary binary variable. The obtained solution  $\Upsilon^*$  belongs to the rational  
 273 reaction set. Hence, the optimal solution of MP is a feasible solution of the bilevel problem.  
 274

275 The process in which the dual problem (12)-(15) is solved for a fixed vector  $\mathbf{z}$  obtaining  
 276 an optimal solution  $v = (\alpha, \beta)$  is represented as  $D(z) \rightarrow v$ . Analogously, solve MP for a set  
 277 of vertices  $\varphi$  is denoted as  $\text{MP}(\varphi) \rightarrow (\chi, \Upsilon)$ . The result is a pair of solutions  $(\chi, \Upsilon)$ . The  
 278 pseudocode of the EPIA is presented in algorithm 4.1.  
 279

280

---

**Algorithm 4.1** Extreme points iterated algorithm

---

**Step 1.** Initialization:  $\varphi = \emptyset$ ,  $\rho = \infty$ ;

**Step 2.** Find the initial vertices  $D(\mathbf{0}) \rightarrow v_1$ .  $D(q_i^B) \rightarrow v_2$ .  $\varphi = \varphi \cup \{v_1, v_2\}$ ;

**Step 3.** Solve the master problem:  $\text{MP}(\varphi) \rightarrow (\chi^k, \Upsilon^k)$ .  $\pi = F(\chi^k, \Upsilon^k)$ ;

**Step 4.** Find a vertex  $k$ :  $D(z^k) \rightarrow v_k$ ;

**Step 5.** Evaluate leader's objective function.

- **If**  $\rho > \pi$  **then**  $\varphi = \varphi \cup \{v_k\}$ ,  $\rho = \pi$ . Return to **Step 3**;
- **If**  $\rho \leq \pi$  **then** stops.

**Output:**  $\pi, (\chi^k, \Upsilon^k)$

---

281 Note that optimality of the bilevel problem is not guaranteed by this algorithm. The  
 282 main reason is that during the process, the algorithm may not improve the leader's objective  
 283 function. This may occurs when the value  $z^k$  obtained at iteration  $k$  produces a vertex  $v_k$   
 284 that is already in set  $\varphi$ , or simply when  $F(\chi^k, \Upsilon^k) \geq F(\chi^{k-1}, \Upsilon^{k-1})$ .

285 *4.2. Follower's gap penalization algorithm (FGPA)*

286 The general idea of FGPA is to find a feasible solution by using vertices of the follower's  
 287 dual problem, but the dual admits infeasible solutions. The algorithm consists on itera-

288 tively solving a mixed-integer programming problem, named MMP, that is a modification  
 289 of the MP described in previous section. The MMP permits the existence of a gap be-  
 290 tween follower's primal objective function and the dual objective function value obtained by  
 291 evaluating the considered vertices. However, this gap is penalized in the leader's objective  
 292 function aiming to obtain a good feasible solution.

293  
 Since the primal and dual objective values of the follower can be different, the algorithm  
 may explore other dual vertices using solutions that are in the constraint region of the bilevel  
 problem, but not necessarily in the inducible region. Hence, to ensure feasibility, at the end  
 of the iterations another problem named Resulting problem (RP) must be solved. RP is a  
 linear single-level problem defined by constraints (1)-(7) of the leader, constraints (9)-(11) of  
 the follower, and a constraint that equals  $f(\Upsilon)$  with the value of the dual objective function  
 $g$  evaluated with the last obtained value of  $\mathbf{z}$  ( $\hat{z} = z^k$ ). Therefore,  $\psi = g(z^k)$  is a parameter  
 used in:

$$\sum_{j \in J} \sum_{i \in I} (p_i - c_{ij}^E) y_{ij} = \psi \quad (78)$$

294 In MMP, the gap is normalized by a constant  $M$  and multiplied by a coefficient  $\mu$  to  
 295 regulate the impact in the leader's objective function, that is, to balance the supply and  
 296 demand. The value of  $M$  can be bounded by the maximum among all the optimal values of  
 297 the dual problem obtained at each iteration. The modified master problem (MMP) is follows:

$$\min_{x, z, r, s, y, \lambda, \varepsilon} \sum_{i \in I} (r_i + s_i) + \frac{\varepsilon}{M} \mu \quad (79)$$

$$\text{s.t.} \frac{\sum_{j \in J} y_{ij} + x_i}{d_i} + r_i - s_i = 1 \quad \forall i \in I \quad (80)$$

$$\sum_{i \in I} (p_i - c_i^G) x_i \geq t \quad (81)$$

$$0 \leq x_i \leq q_i^A \quad \forall i \in I \quad (82)$$

$$0 \leq z_i \leq q_i^B \quad \forall i \in I \quad (83)$$

$$r_i \geq 0 \quad \forall i \in I \quad (84)$$

$$s_i \geq 0 \quad \forall i \in I \quad (85)$$

$$\sum_{j \in J} a_{ij} y_{ij} \leq z_i \quad \forall i \in I \quad (86)$$

$$\sum_{i \in I} b_{ij} y_{ij} \leq m_j \quad \forall j \in J \quad (87)$$

$$y_{ij} \geq 0 \quad \forall i \in I, j \in J \quad (88)$$

$$\sum_{k \in K} \lambda_k = 1 \quad (89)$$

$$\lambda_k \in \{0, 1\} \quad \forall k \in K \quad (90)$$

$$\varepsilon^k + \sum_{j \in J} \sum_{i \in I} (p_i - c_{ij}^E) y_{ij} = \sum_{i \in I} \alpha_i^k z_i + \sum_{j \in J} \beta_j^k m_j \quad \forall k \in K \quad (91)$$

$$\varepsilon^k \geq 0 \quad \forall k \in K \quad (92)$$

$$\varepsilon \geq 0 \quad (93)$$

$$\varepsilon \geq \varepsilon^k - (1 - \lambda_k) M \quad \forall k \in K \quad (94)$$

Therefore, the RP that is used to obtain a solution in the inducible region is:

$$\min_{x, r, s, y} \sum_{i \in I} (r_i + s_i) \quad (95)$$

$$\text{s.t.} \quad \frac{\sum_{j \in J} y_{ij} + x_i}{d_i} + r_i - s_i = 1 \quad \forall i \in I \quad (96)$$

$$\sum_{i \in I} (p_i - c_i^G) x_i \geq t \quad (97)$$

$$0 \leq x_i \leq q_i^A \quad \forall i \in I \quad (98)$$

$$r_i \geq 0 \quad \forall i \in I \quad (99)$$

$$s_i \geq 0 \quad \forall i \in I \quad (100)$$

$$\sum_{j \in J} a_{ij} y_{ij} \leq \hat{z}_i \quad \forall i \in I \quad (101)$$

$$\sum_{i \in I} b_{ij} y_{ij} \leq m_j \quad \forall j \in J \quad (102)$$

$$y_{ij} \geq 0 \quad \forall i, j \in J \quad (103)$$

$$\sum_{j \in J} \sum_{i \in I} (p_i - c_{ij}^E) y_{ij} = \psi \quad (104)$$

To refer to the MMP, define  $\text{MMP}(\varphi, M) \rightarrow (\chi, \Upsilon)$  as in the previous algorithm. The leader's objective function (79) is represented by  $F^*$ . Also,  $\text{RP}(\varphi, \psi, \hat{z}) \rightarrow (\chi, \Upsilon)$  denotes the process of solving RP using the set of vectors  $\varphi$  and the parameters  $\psi$  and  $\hat{z}$ . The pseudocode of FGPA is shown in algorithm 4.2.

### 4.3. Hybrid algorithm (HYBA)

As mentioned before, the EPIA cannot ensure to obtain a new vertex in each iteration nor a vertex that improves the leader's objective function. An intuitive idea is to use the FGPA to avoid this issue. The latter algorithm has the advantage of exploring solutions that does not belong to the inducible region, by doing this, unexplored vertices are obtained.

The main idea of the HYBA is to perform the steps of EPIA until stops. When this occurs, a subroutine that solves the MMP is performed to find a new vertex aiming to reach a different feasible solution. If the new exploration improves the leader's objective function,

---

**Algorithm 4.2** Follower's gap penalization algorithm

---

**Step 1.** Initialization:  $\varphi = \emptyset$ ,  $\rho = \infty$ ,  $M = \text{Max}\{1, D(q_i^B)\}$ ;

**Step 2.** Solve the modified master problem:  $\text{PMM}(\varphi, M) \rightarrow (\chi^k, \Upsilon^k)$ .  $\pi = F^*(\chi^k, \Upsilon^k)$ ;

**Step 3.** Find a vertex  $k$ :  $D(z^k) \rightarrow v_k$ .  $M = \text{Max}\{M, D(z^k)\}$ ;

**Step 4.** Evaluate leader's objective function.

- **If  $\rho > \pi$  then**  $\varphi = \varphi \cup \{v_k\}$ ,  $\rho = \pi$ . Return to **Step 2**;

- **If  $\rho \leq \pi$  then** go to **Step 5**.

**Step 5.** Obtain a feasible solution:  $\psi = D(z_k)$ .  $\text{RP}(\varphi, \psi, \hat{z}) \rightarrow (\chi^*, \Upsilon^*)$ .  $\pi = F(\chi^*, \Upsilon^*)$ ;

**Output:**  $\pi, (\chi^*, \Upsilon^*)$

---

315 then the FGPA continues until the stopping criterion is reached. The pseudocode of the  
316 proposed hybrid algorithm is shown in algorithm 4.3.

317

---

**Algorithm 4.3** Hybrid algorithm

---

**Step 1.** Initialization:  $\varphi = \emptyset$ ,  $\rho = \infty$ ;

**Step 2.** Find the initial vertices:  $D(\mathbf{0}) \rightarrow v_1$ .  $D(q_i^B) \rightarrow v_2$ .  $\varphi = \varphi \cup \{v_1, v_2\}$ ,  $M = \text{Max}\{D(\mathbf{0}), D(q_i^B)\}$ ;

**Step 3.** Solve the master problem:  $\text{MP}(\varphi) \rightarrow (\chi^k, \Upsilon^k)$ .  $\pi = F(\chi^k, \Upsilon^k)$ ;

**Step 4.** Find a vertex  $k$ :  $D(z^k) \rightarrow v_k$ .  $\varphi = \varphi \cup \{v_k\}$ ,  $M = \text{Max}\{M, D(z^k)\}$ ;

**Step 5.** Evaluate leader's objective function.

- **If  $\rho > \pi$  then** Go to **Step 6**;

- **If  $\rho \leq \pi$  then**

$\text{MMP}(\varphi, M) \rightarrow (\chi^{k^*}, \Upsilon^{k^*})$ .  $D(z^{k^*}) \rightarrow v_{k^*}$ .  $\varphi = \varphi \cup \{v_{k^*}\}$ ;

$\text{MP}(\varphi) \rightarrow (\chi^k, \Upsilon^k)$ .  $\pi = F(\chi^k, \Upsilon^k)$ .  $D(z^k) \rightarrow v_k$ ,  $\varphi = \varphi \cup \{v_k\}$ ,  $M = \text{Max}\{M, D(z^k)\}$ .

**Step 6.** Re-evaluating leader's objective function:

- **If  $\rho > \pi$  then**  $\rho = \pi$ . Return to **Step 3**;

- **If  $\rho \leq \pi$  then** Stops.

**Output:**  $\pi, (\chi^k, \Upsilon^k)$

---

## 318 5. Computational experimentation

319 In this section, the results obtained from an extensive computational experimentation of  
320 our solution methodologies are reported. To evaluate both single-level reformulations and



321 the performance of the proposed heuristic algorithms, a set of 360 instances was used. The  
322 first subset of 180 instances corresponds to a case based on a real situation occurred in the  
323 Mexican petrochemical industry between 1958-2014. The second subset of 180 instances was  
324 randomly generated to test the efficiency of our algorithms in more general data sets.

325  
326 All the computational experiments were conducted in a personal computer Dell Inspiron  
327 5558, with the following characteristics: a processor Intel(R) Core i3 with 2.10 GHz, 6.00  
328 GB of RAM under Windows 10 operative system. The mixed-integer linear reformulation  
329 based on the complementarity slackness condition (hereafter Ref.2), and the three proposed  
330 heuristic algorithms were implemented in C++ using Microsoft Visual Studio 2017. The  
331 optimizer used was CPLEX 12.7.1. On the other hand, the non-linear reformulation based  
332 on the strong duality condition (hereafter Ref.1) was implemented in AMPL using Baron  
333 18.5.8 as optimizer.

### 334 *5.1. Case study*

335 In Mexico, the government intervention in the petrochemical industry existed by many  
336 decades. The main characteristic was that the government monopolized the extraction of  
337 the main raw material for this industry, namely petroleum and some other derivatives called  
338 basic petrochemicals. Another characteristic of this situation was that state-owned firms  
339 competed against private firms in the market of final commodities (secondary petrochemi-  
340 cals).

341  
342 The institutional framework defined by law the petrochemical industry as the one that  
343 performs chemical or physical processes to product compounds from petroleum natural hy-  
344 drocarbons or from the products derived from refinement operations. Some of these products  
345 may serve as raw materials to the industry, and they were classified as part of the basic petro-  
346 chemicals. The remaining products were included into the secondary petrochemical category  
347 [19].

348  
349 Specifically, in a law from 1958 [20], it was established that only the government was  
350 allowed to exploit the hydrocarbons related with the oil industry, which was concerned with  
351 the production, warehousing, distribution, and sales of the petroleum derivatives that can  
352 be considered as raw materials (basic petrochemicals) for the industry. However, for the  
353 production of secondary petrochemicals products, state-owned and private firms were al-  
354 lowed to be involved in. Therefore, Petróleos Mexicanos (PEMEX) and its subsidiaries was  
355 in charge of this industry [21], and from 1992, also their decentralized departments [22].  
356 The classification of basic and secondary petrochemical products varies by law from year to  
357 year [23], [21], [19], [24], [25], [26], [27].

358  
359 To delimit the scope and size of our case study instances based on the above situation,  
360 we consider the number of economic units involved in the manufacturing of organic chemical  
361 basic products registered by the National Institute of Statistics and Geography in Mexico  
362 (INEGI by its acronym in Spanish) during the economical census conducted in 1999, 2004,

2009, and 2014. The biggest number of these units was registered in 2004, which was 159. This value is used as an upper bound of the private firms dedicated to secondary petrochemical industry. For delimiting the amount of commodities considered, we consider the number of commodities classified as basic and secondary petrochemical by the Mexican legislation in different years. This information is summarized in table 1.

Table 1: Classification of petrochemical commodities

Year of the classification	Basic petrochemicals	Secondary petrochemicals
1960	16	-
1986	34	26
1989	8	13
1992	20	35

Obtained by using data from [21], [24], [25], [26]

To complete the instances, we extracted data from the statistical report of the energy sector used by the Mexican Secretary of Energy. The data of 47 petrochemical commodities from 1980 to 2002 were obtained. The information consists of the demand, prices, and productive capacity. For each commodity  $i$ , the demand  $d_i$  was generated from a uniform distribution with the following parameters: the average of the demands in that period of time minus the standard deviation, and the average of the demands in that period of time plus the standard deviation. The price  $p_i$  and the government production capacity  $q_i^A$  were randomly generated between 1 and the maximum price or the maximum capacity for each commodity  $i$ , respectively.

The government costs  $c_i^G$  were obtained as the product of the price of each commodity  $i$  times a random number between 0.22 and 0.60. The latter range was defined based on the ratio between the production costs and the PEMEX (the Mexican state-owned firm) total income registered from the years 1988-2000, 2011-2013, and 2015-2016. The minimum profit that the state-owned firm must achieved was computed as the 30% of the total benefit that the government may obtain if all the petrochemicals market demand was satisfied by PEMEX.

The state-owned firm has a production capacity  $q_i^B$  regarding raw materials  $i$ , which was computed as the product of the maximum production coefficient of the private firms and the maximum capacity of the state-owned firm for producing commodities. Under this assumption, the private firm with best technology could match the production capacity of the state-owned firm for an specific commodity.

Production costs  $c_{ij}^E$  of private firm  $j$  are defined as the product of random coefficients in  $[0.784, 0.884]$  multiplied times the price of each commodity  $i$ . That range corresponds to

394 the minimum and maximum average values of the ration between the production costs and  
395 the income per year of five real private firms. These private firms operate in the Mexican  
396 petrochemical sector: Alpek, Kuo (UEN synthetic rubber), Kuo (UEN polystyrene), Mexi-  
397 chem, and Pochteca group.

398  
399 The production technical coefficients  $a_{ij}$  are randomly generated between  $[0.085, 2.111]$ ,  
400 which correspond to the range determined by the average plus/minus one standard deviation  
401 of all the petrochemical substances considered in the report presented by CEPAL [? ]. The  
402 production levels  $b_{ij}$  are generated from the interval  $[1, 95]$ . The lower bound is natural since  
403 these coefficients must be strictly positives, while the upper bound is equal to the estimated  
404 average of the production levels for each substance considered in the petrochemical facilities  
405 of PEMEX.

406  
407 Finally, the production capacity  $m_j$  of the private firms were randomly generated be-  
408 tween 4,665 and 20,825, which came from the average of the production capacity of five  
409 petrochemical complex of PEMEX.

410  
411 Based on the above factors and ranges, we generated 30 instances for each one of the  
412 following sizes:

- 413 •  $|I| = 10, |J| = 10$
- 414 •  $|I| = 25, |J| = 25$
- 415 •  $|I| = 25, |J| = 75$
- 416 •  $|I| = 50, |J| = 100$
- 417 •  $|I| = 75, |J| = 125$
- 418 •  $|I| = 150, |J| = 200$

419 As it is mentioned above, the instances generated to analyze the case study were taken  
420 as a basis for constructing another set of synthetic instances to test the performance of our  
421 algorithms on a different data set. This new set of instances has the same sizes but were  
422 randomly generated using arbitrary ranges for each parameter. Furthermore, the generation  
423 scheme guarantees feasibility of the problem, that is, ensure that the state-owned firm has  
424 capacity to produce all the demand.

## 425 5.2. Numerical results

426 Our computational experiment consists on solving both subsets of instances described  
427 above, that is, the case-study and the synthetic instances. All of them were solved by the  
428 two reformulations presented in section 3 and by the three heuristic algorithms proposed  
429 in section 4. The leader's objective function value and the required time were registered.  
430 Among the two exact methods, only Ref.2 was able to optimally solve all the tested instances.

431 Hence, these values were used to compute the optimality gap obtained by the heuristic al-  
 432 gorithms. A maximum CPU time of 1000 seconds was set to the solver for solving Ref.1,  
 433 while Ref.2 and the heuristic algorithms we did not fix a time limit since in all cases the  
 434 required time was rather small.

435  
 436 We observe from our results, that Ref.1 was not able to solve all the instances within the  
 437 time limit. We also observe that the case-study instances are harder to be solved than the  
 438 synthetic ones. For example, for the larger instance sizes ( $150 \times 200$ ), Ref.1 did not even find  
 439 feasible solutions for the problem. This may be due to the fact that synthetic instances were  
 440 well-balanced so that, in most cases, shortages and surplus are both zero. Summarizing, the  
 441 number of instances solved to optimality by each reformulation is shown in table 2.

Table 2: Number of instances solved to optimality

Instances	Size	Ref.1	Ref.2
Case-study	$10 \times 10$	15	30
	$25 \times 25$	0	30
	$25 \times 75$	0	30
	$50 \times 100$	0	30
	$75 \times 125$	0	30
	$150 \times 200$	0	30
Synthetic	$10 \times 10$	30	30
	$25 \times 25$	29	30
	$25 \times 50$	30	30
	$50 \times 100$	29	30
	$75 \times 125$	28	30
	$150 \times 200$	13	30

442 The following tables 3 and 4 report all the results of our experiment. The results are or-  
 443 ganized in five column blocks corresponding to the five solution methods that are compared.  
 444 In each block, we include two columns with the objective function values (F) and the CPU  
 445 time (t). In these tables, averages of the registered values for each size of the case-study and  
 446 synthetic instances are shown. It is worth mentioning that when Ref.1 was not able to solve  
 447 the instance, that instance was omitted for the computation of the corresponding average  
 448 value.

Table 3: Summarized results obtained from the case-study instances

Size	Ref.1		Ref.2		EPIA		FGPA		HYBA	
	$\bar{F}$	$\bar{t}$ (s)	$\bar{F}$	$\bar{t}$ (s)	$\bar{F}$	$\bar{t}$ (s)	$\bar{F}$	$\bar{t}$ (s)	$\bar{F}$	$\bar{t}$ (s)
10 × 10	0.561	255.56	0.506	0.18	0.774	0.23	0.825	0.16	0.524	0.41
25 × 25	14.225	500.26	1.000	1.86	1.570	0.30	1.171	0.35	1.020	0.71
25 × 75	12.435	500.46	0.527	4.36	1.025	0.48	0.565	0.39	0.552	1.02
50 × 100	211.696	558.80	1.051	53.15	2.041	2.39	1.132	1.14	1.091	5.18
75 × 125	431.138	501.24	1.546	268.89	2.339	8.89	1.811	3.68	1.621	20.69
150 × 200	-	-	3.279	2754.68	4.129	61.82	5.037	12.06	3.443	112.25

Table 4: Summarized results obtained from the synthetic instances

Size	Ref.1		Ref.2		EPIA		FGPA		HYBA	
	$\bar{F}$	$\bar{t}$ (s)	$\bar{F}$	$\bar{t}$ (s)	$\bar{F}$	$\bar{t}$ (s)	$\bar{F}$	$\bar{t}$ (s)	$\bar{F}$	$\bar{t}$ (s)
10 × 10	0	0.26	0	0.14	0.007	0.08	0.006	0.10	0	0.09
25 × 25	0.012	29.01	0	1.43	0	0.11	0.021	0.14	0	0.11
25 × 75	0	10.94	0	3.65	0	0.17	0.019	0.20	0	0.17
50 × 100	0.057	73.60	0	35.52	0	0.29	0.170	0.31	0	0.28
75 × 125	0	118.76	0	90.48	0	0.60	0.086	0.76	0	0.57
150 × 200	1.657	414.55	0	193.48	0	1.78	0.443	1.74	0	1.57

449 We observe from tables 3 and 4 that Ref.1 (recall that Ref.1 is a continuous non-linear  
450 global optimization problem, while Ref.2 is a MILP) gets values, on average, closed to those  
451 obtained by Ref.2 for the synthetic instances. However, for the case-study instances the  
452 situation is different and the average solution values and times move away from the optimal  
453 averages as the size of the instance increases. Note that for most of the synthetic instances,  
454 the optimal value is zero, i.e., the supply is perfectly balanced with the demand. Hence,  
455 neither shortages nor surplus exist due to the generation mechanism of these instances.

456  
457 Regarding the heuristic algorithms, it can be observed from table 4 that EPIA and  
458 HYBA report values, on average very close, to the optimal ones for all the sizes of the  
459 synthetic instances. However, for the case-study instances, HYBA reports the best results  
460 among all the three heuristic algorithms (see table 3). Moreover, the performance of EPIA  
461 and FGPA significantly depends on the size of the instance. In terms of the average re-  
462 quired time, FGPA requires less time for the case-study instances. This finding may be  
463 derived from the fact that the algorithm explores points not necessarily in the inducible  
464 region, which may help to approach the optimal solution in a faster manner. On the con-  
465 trary, HYBA is the algorithm that requires the more time, but this may be obvious since it  
466 performs all the steps of EPIA and at each iteration it may solve upon four extra mathemat-  
467 ical programs. Nevertheless, the registered times are acceptable for a problem of this nature.

468

To compare the efficiency and the quality of the solutions obtained by the heuristic algorithms, the optimality gap is computed using the values obtained by Ref.2. Given that there are many optimal values equal to zero, the optimality gap (GAP) measures the relative deviation from the optimal value, and it is computed by the following formula:

$$\text{GAP} = \left| \frac{(\text{Optimal value}) - (\text{Obtained value})}{(\text{Obtained value})} \right| \times 100\% \quad (105)$$

469 Also, an analogous formula is used to compute the relative savings in time (%*t*). That is,  
 470 the reduction in computational time when solving the problem using the heuristics rather  
 471 than Ref.2. A negative value indicates that the heuristic consumes more time than the exact  
 472 reformulation. The formula used is presented next:

473

$$\%t = \frac{(\text{Required time by Ref.2}) - (\text{Required time by a heuristic})}{(\text{Required time by a heuristic})} \times 100\% \quad (106)$$

474 The GAP and %*t* average values obtained for both type of instances are shown in tables  
 475 5 and 6, respectively.

Table 5: Evaluating the effectiveness of the heuristic algorithms in the case-study instances

Size	EPIA		FGPA		HYBA	
	GAP	% <i>t</i>	GAP	% <i>t</i>	GAP	% <i>t</i>
10 × 10	42.27%	12%	29.60%	15%	7.68%	-48%
25 × 25	40.15%	746%	11.70%	534%	1.67%	190%
25 × 75	53.78%	1156%	19.67%	1101%	7.04%	445%
50 × 100	54.29%	3418%	10.72%	5418%	6.02%	1176%
75 × 125	33.20%	5269%	12.26%	10696%	5.64%	1499%
150 × 200	19.29%	9210%	17.20%	30206%	4.92%	3324%

Table 6: Evaluating the effectiveness of the heuristic algorithms in the synthetic instances

Size	EPIA		FGPA		HYBA	
	GAP	% <i>t</i>	GAP	% <i>t</i>	GAP	% <i>t</i>
10 × 10	6.67%	79%	10.00%	42%	0%	69%
25 × 25	0%	1332%	46.67%	1038%	0%	1283%
25 × 75	0%	2189%	50.00%	1895%	0%	2192%
50 × 100	0%	13714%	90.00%	12176%	0%	14466%
75 × 125	0%	16773%	90.00%	14272%	0%	18485%
150 × 200	0%	11561%	100.00%	11377%	0%	12781%

476 HYBA is the algorithm that shows the best quality in the feasible solutions obtained.  
477 For the case study instances, the average optimality gap was lower than 8% for all the sizes  
478 (see table 5); while for the synthetic instances, the average optimality gap was zero (see ta-  
479 ble 6). FGPA presents a very good average gap in comparison with EPIA in the case-study  
480 instances, but this behavior is opposite in the synthetic instances. Concerning the savings  
481 in the required time, the three heuristic algorithms showed significant savings, which is im-  
482 proved as the size of the instance increases. FGPA is the one that evidence more savings for  
483 the case-study instances. It is convenient to mention that for the synthetic instances of size  
484  $10 \times 10$ , HYBA showed a negative saving, this implies that Ref.2 was faster to solve these  
485 instances. But, this is expected since HYBA solves at least one linear model and one MILP  
486 model, and Ref.2 only solves one MILP model. In spite of that, the advantage of HYBA  
487 over the reformulation is evident for medium and large-size instances.

488  
489 To support the latter findings, the number of optimal solutions obtained by each heuris-  
490 tic algorithm are displayed in table 7. It can be seen from that table that HYBA is the  
491 algorithm that was able to obtain more optimal solutions. Also, it an be observed that  
492 FGPA obtained more optimal solutions for the case-study instances than EPIA, but for the  
493 synthetic instances the behavior was the opposite. These results confirm the suitability for  
494 combining both heuristic algorithms to create HYBA.

Table 7: Number of optimal solutions obtained for each heuristic algorithm

Instance	Size	EPIA	FGPA	HYBA
Case-study	$10 \times 10$	7	11	20
	$25 \times 25$	0	12	19
	$25 \times 75$	5	13	19
	$50 \times 100$	1	13	15
	$75 \times 125$	0	4	9
	$150 \times 200$	0	0	1
Synthetic	$10 \times 10$	28	27	30
	$25 \times 25$	30	16	30
	$25 \times 75$	30	15	30
	$50 \times 100$	30	3	30
	$75 \times 125$	30	3	30
	$150 \times 200$	30	0	30

### 495 5.3. Using heuristic solutions to obtain the optimal

496 The good performance of the heuristic algorithms lead us to analyze the idea of use the  
497 near-optimal obtained solutions as an input to Ref.2 seeking to enhance their process. By  
498 doing this, we expect to significantly reduce the computational time required to optimally  
499 solve an instance. The results obtained from this experimentation are shown in table 8.  
500 The columns with label “Ref.2”, “HYBA”, and “Ref.2 w/initial sol.” display the average

501 required time for each solution scheme. The final column “Total” sums the time required  
502 for the hybrid algorithm and Ref.2 using an initial good solution. As it was expected, the  
503 advantages of using this scheme are more notorious for medium and large-size instances. The  
504 best example of this finding is given by the instances of size  $150 \times 200$ , in which the time  
505 was reduced from 2754.68s to 586.53s, and from 193.48s to only 7.28s, for the case-study  
506 and synthetic instances, respectively.

Table 8: Average required time when Ref.2 uses an initial heuristic solution

Instance	Size	Ref.2	HYBA	Ref.2 w/initial sol.	Total
Case study	$10 \times 10$	0.18	0.33	0.07	0.40
	$25 \times 25$	1.86	0.82	0.39	1.21
	$25 \times 75$	4.36	1.54	0.80	2.34
	$50 \times 100$	53.15	6.07	4.24	10.31
	$75 \times 125$	268.89	20.24	19.73	39.97
	$150 \times 200$	2754.68	196.38	390.15	586.53
Random	$10 \times 10$	0.14	0.09	0.04	0.13
	$25 \times 25$	1.43	0.12	0.10	0.22
	$25 \times 75$	3.65	0.19	0.49	0.68
	$50 \times 100$	35.52	0.33	1.15	1.48
	$75 \times 125$	90.48	0.46	2.22	2.68
	$150 \times 200$	193.48	1.95	7.28	9.23

## 507 6. Conclusions and research directions

508 This paper studies a market regulation situation, in which the government controls the  
509 distribution of raw material and competes against private firms in the production of final  
510 commodities. In this market, the government has a privileged position since it determines its  
511 production and the amount of raw material offered to private firms and its goal is to balance  
512 the market. Although this type of market situation is common in some economies, to the  
513 best of our knowledge and after an intensive literature revision, we did not find a similar  
514 problem in the literature. The situation that motivated our paper comes from the actual  
515 situation in Mexican petrochemical industry with PEMEX from 1958 till 2014. However,  
516 this case is not exclusive from the petrochemical industry. Also in the agricultural industry,  
517 the Mexican government monopolized the production, imports, and distribution of fertil-  
518 izers from 1970 to 1986 through the state-owned company named FERTIMEX [28]. The  
519 government also regulated the water supply for agricultural consumption by fixing prices,  
520 defining production goals, and restricting the land transactions [29]. In addition, this type  
521 of economic regulation has been applied in other countries. For example, in [3], it is stated  
522 that there is a tendency to use government instruments in developing countries, mainly  
523 in raw material industries. An indicator of the latter can be seen in the large amount of  
524 oil expropriations in the 70’s (see [4], [5], [6]). Also, nationalizations occurred in Bolivia,  
525 Ecuador, Venezuela, and Russia in 2006 (see [6]).

526



527 Motivated by the lack of models to analyze these type of markets, this paper develops  
528 a bilevel programming approach that fits naturally to the problem. To solve the bilevel  
529 problem herein proposed, two single-level reformulations are presented. The first reformula-  
530 tion uses the lower level necessary and sufficient optimality conditions to ensure the global  
531 optimum of the bilevel problem. As a result of this reformulation, a continuous non-linear  
532 problem is obtained. The second reformulation substitutes the above-mentioned conditions  
533 by the complementarity slackness conditions generating a MILP. These reformulations are  
534 able to solve optimally limited size instances. Therefore, three heuristic algorithms are pro-  
535 posed to find good quality feasible solutions in a reasonable computational time. The first  
536 two algorithms (EPIA and FGPA) are tailored for this problem, and exploits the fact that  
537 the polyhedron associated to the dual problem of the lower level remains the same for any  
538 leader’s decision. The third algorithm (HYBA) is a combination of the previous ones.

539  
540 Extensive computational experimentation is performed to test the exact and heuristic  
541 approaches proposed in this paper. Two sets of instances are used: case-study instances,  
542 that arise from the Mexican petrochemical industry, and synthetic instances adapted from  
543 the previous ones. As a result from the computational experimentation, it can be stated that  
544 the reformulation that applies the complementarity slackness constraints is more efficient for  
545 the tested instances. On the other hand, the performance of EPIA and FGPA strongly de-  
546 pends on the characteristics of the instance. In addition, the hybridization of both heuristic  
547 algorithms (HYBA) yields to the best results in terms of effectiveness and solutions quality.  
548 Moreover, HYBA found the optimal solution for all the synthetic instances, and obtained  
549 optimality gaps less than 8% in the case-study ones. It is worth to mention that HYBA  
550 found the optimal value in at least 50% for the first four types of instance sizes.

551  
552 Finally, it can be concluded from our computational experiment, that the most efficient  
553 manner to find optimal solutions for the bilevel problem is by obtaining a good quality  
554 feasible solution through HYBA and then using it as an initial solution to solve Ref.2. By  
555 doing this, a significant computational time reduction is achieved.

556  
557 Note that this bilevel problem may have multiple optimal solutions, each of them offers  
558 different alternatives to the upper level decision-maker (government) for producing and  
559 offering raw material to the lower level decision-makers (private firms). Therefore, this  
560 characteristic is economically worthy to be explored. Another possible research direction  
561 is to neglect, in the model, the assumption there exists an organism that manages the  
562 cooperation among private firms. Hence, a new element will emerge in the model since  
563 competition among private firms will also exist. As a result of this, a bilevel problem with  
564 one leader and multiple followers will appear. Then, in a natural way, a Stackelberg game  
565 in which the follower’s problem consists of a generalized Nash game fits to analyze this  
566 situation.

567 **Acknowledgements**

568 José-Fernando Camacho-Vallejo has been partially supported by the program of profes-  
569 sional development of professors with the grant PRODEP/511-6/17/7425 for research stays  
570 during his sabbatical year. Justo Puerto has been partially supported by Spanish/FEDER  
571 grants number MTM2016-74983-C02-01 (MINECO). Martine Labbé has been partially sup-  
572 ported by the Fonds de la Recherche Scientifique - FNRS under Grant(s) no PDR T0098.18.

573 **References**

- 574 [1] E. G. Wiens, Government firm regulation of a vertically integrated industry, Carleton Economic Paper  
575 (1978) 78–109.
- 576 [2] J. E. Stiglitz, Regulation and failure, in: D. Moss, J. Cisternino (Eds.), *New Perspectives on Regulation*,  
577 Tobin Project, 2009.
- 578 [3] R. Vernon, Uncertainty in the resource industries: the special role of state-owned enterprises, in: D. W.  
579 Pearce, H. Siebert (Eds.), *Risk and the Political Economy of Resource Development*, Springer, 1984,  
580 pp. 207–223.
- 581 [4] S. J. Kobrin, The nationalisation of oil production, 1918-80, in: D. W. Pearce, H. Siebert (Eds.), *Risk*  
582 *and the Political Economy of Resource Development*, Springer, 1984, pp. 137–164.
- 583 [5] M. S. Minor, The demise of expropriation as an instrument of ldc policy 1980-1992, *Journal of Inter-*  
584 *national Business Studies* 25 (1) (1994) 177–188.
- 585 [6] S. Guriev, A. Kolotilin, K. Sonin, Determinants of nationalization in the oil sector: A theory and  
586 evidence from panel data, *The Journal of Law, Economics, and Organization* 27 (2) (2011) 301–323.
- 587 [7] W. C. Merrill, N. Schneider, Government firms in oligopoly industries: A short-run analysis, *The*  
588 *Quarterly Journal of Economics* 80 (3) (1966) 400–412.
- 589 [8] B. Blackwell, The state and the macro-industrial economy: Lessons from the aerospace sector, *Omega*  
590 14 (6) (1986) 421–442.
- 591 [9] R. G. Harris, E. G. Wiens, Government enterprise: An instrument for the internal regulation of industry,  
592 *The Canadian Journal of Economics* 13 (1) (1980) 125–132.
- 593 [10] M. R. Sertel, Regulation by participation, *Journal of Economics* 48 (2) (1988) 111–134.
- 594 [11] H. Cremer, M. Marchand, J.-F. Thisse, The public firm as an instrument for regulating an oligopolistic  
595 market, *Oxford Economic Papers* 41 (2) (1989) 283–301.
- 596 [12] G. De Fraja, F. Delbono, Game theoretic models of mixed oligopoly, *Journal of Economic Surveys* 4 (1)  
597 (1990) 1–17.
- 598 [13] G. De Fraja, F. Delbono, Mixed oligopoly: An overview, Tech. rep., Quaderni-Working Paper DSE  
599 (1988).
- 600 [14] L. Netti, Mixed oligopoly with homogeneous goods, *Annals of Public and Cooperative Economics* 64 (3)  
601 (1993) 367–393.
- 602 [15] E. Z. Demirci, N. K. Erkip, Designing an intervention strategy for public-interest goods: The california  
603 electric vehicle market case, *Omega* 69 (2017) 53–69.
- 604 [16] G. Whittaker, R. Färe, S. Grosskopf, B. Barnhart, M. Bostian, G. Mueller-Warrant, S. Griffith, Spatial  
605 targeting of agri-environmental policy using bilevel evolutionary optimization, *Omega* 66 (2017) 15–27.
- 606 [17] S. Dempe, *Bilevel programming: A survey*, Dekan der Fak. für Mathematik und Informatik, 2003.
- 607 [18] J. F. Bard, *Practical Bilevel Optimization, Nonconvex Optimization and Its Applications*, Springer US,  
608 1998.
- 609 [19] Reglamento de la ley reglamentaria del artículo 27 constitucional en el ramo del petróleo, en materia  
610 de petroquímica, Diario Oficial, México, 9 de febrero de 1971.
- 611 [20] Ley reglamentaria del artículo 27 constitucional en el ramo del petróleo, Diario Oficial, México, 29 de  
612 noviembre de 1958.

- 613 [21] Acuerdo que dispone que corresponde a petróleos mexicanos u organizaciones o empresas subsidiarias  
614 o asociadas a la misma institución, la elaboración de diversos productos de petroquímica básica, Diario  
615 Oficial, México, 9 de abril de 1960.
- 616 [22] Ley orgánica de petróleos mexicanos y organismos subsidiarios, Diario Oficial de la Federación, México,  
617 16 de julio de 1992.
- 618 [23] Reglamento de la ley reglamentaria del artículo 27 constitucional en el ramo del petróleo, Diario Oficial,  
619 México, 25 de agosto de 1959.
- 620 [24] Resolución que clasifica los productos petroquímicos que se indican, dentro de la petroquímica básica  
621 o secundaria, Diario Oficial de la Federación, México, 13 de octubre de 1986.
- 622 [25] Resolución que clasifica los productos petroquímicos que se indican, dentro de la petroquímica básica  
623 o secundaria, Diario Oficial de la Federación, México, 15 de agosto de 1989.
- 624 [26] Resolución que clasifica los productos petroquímicos que se indican, dentro de la petroquímica básica  
625 o secundaria, Diario Oficial de la Federación, México, 17 de agosto de 1992.
- 626 [27] Decreto por el que se reforma la ley reglamentaria del artículo 27 constitucional en el ramo del petróleo,  
627 Diario Oficial de la Federación, México, 13 de noviembre de 1996.
- 628 [28] A. De Janvry, I. F. for Agricultural Development, I.-A. I. for Cooperation on Agriculture, Reformas  
629 del sector agrícola y el campesinado en México, Estrategias para mitigar la pobreza rural en América  
630 Latina y El Caribe, Fondo Internacional de Desarrollo Agrícola, 1995.
- 631 [29] S. Johnson, La transferencia del manejo de la irrigación en México: una estrategia para lograr la  
632 sostenibilidad de los distritos de riego., IWMI Research Report, Instituto Internacional del Manejo de  
633 la Irrigación, 1997.