Impact of Government Policies
on Sustainable Petroleum Supply Chain (SPSC):
A Case Study – Part I (Models)

Davoud Ghahremanlou*, Wieslaw Kubiak*

Abstract. Environmental concerns and energy security have led governments to establish legislation to convert a Conventional Petroleum Supply Chain (CPSC) to a Sustainable Petroleum Supply Chain (SPSC). The United States (US), one of the biggest oil consumers in the world, has created regulations to manage ethanol production and consumption over the last half century. Although these regulations have created new opportunities, they have also added new burdens to the obligated parties. It is thus key for the government, the obligated parties, and related businesses to study the impact of the policies on the SPSC. We develop a two-stage stochastic programming model, General Model (GM), which incorporates Renewable Fuel Standard 2 (RFS2), Tax Credits, Tariffs, and Blend Wall (BW) to study the policy impact on the SPSC using cellulosic ethanol. The model, as with any other general model available in the literature, makes it highly impractical to study the policy impact due to the model's computational complexity. We use the GM to derive a Lean Model (LM) to study the impact by running computational experiments more efficiently and consequently by arriving much faster at robust managerial insights. We present a case study of the policy impact on the SPSC in the State of Nebraska using the LM in the accompanying part II (Ghahremanlou and Kubiak, 2020).

1. INTRODUCTION

1.1. Context and Motivation

Crude oil is the main global source of vehicle transportation fuel (Independent Statistics & Analysis, 2015; Kessel, 2000). Global warming, the uneven distribution of worldwide crude oil reservoirs, and political instability of the countries owning almost half of the

---

* Faculty of Business Administration, Memorial University of Newfoundland, St. John's, NL, Canada

known reservoirs, have all compelled many countries to move towards local renewable energy sources (Sahebi *et al.*, 2014; Agarwal, 2007; Yan, 2012). Ethanol produced from biological materials is considered a replacement for gasoline (El-Naggar *et al.*, 2014; Humbird *et al.*, 2011; Baeyens *et al.*, 2015). However, due to market restrictions and infrastructure compatibility, ethanol is currently used mainly as an additive to gasoline in most countries (Agarwal, 2007; Yacobucci, 2010; HIS Markit, 2019). The Ethanol Supply Chain (ESC) is often merged with CPSC in its downstream, where ethanol is blended with gasoline.

The US was the biggest oil producer and consumer in 2016, and gasoline made up to 60% of total transportation fuel demand in the US. Environmental concerns and energy security led the US government to establish policies to stimulate ethanol production and consumption as an additive to gasoline. The Volumetric Ethanol Excise Tax Credit (VEETC) was created by the America Job Act in 2004, and its amount was reduced in 2009. Based on this tax credit, blenders received 0.45 $\frac{\$}{\text{gal}}$ of ethanol blended with gasoline; since imported ethanol was eligible for this credit it was subject to a 0.54 $\frac{\$}{\text{gal}}$ tariff (McPhail *et al.*, 2011). These rules expired in 2011.

The Energy Independence and Security Act (EISA) was established in 2007. It determines the Renewable Fuel Standard 2 (RFS2). The RFS2 requires gasoline refiners and gasoline importers in the US, called *obligated* parties (Legal Information Institute, 2010; McPhail *et al.*, 2011), to blend at least a minimum amount of renewable fuels, referred to as Renewable Volume Obligations (RVOs), or *mandate*, with their gasoline annually (Duffield *et al.*, 2008; Thompson *et al.*, 2009). According to RFS2, the biofuels are categorized based on their feedstock types and lifecycle Greenhouse Gas (GHG) emissions reduction (Thompson *et al.*, 2009). Ethanol can be produced from different types of feedstocks such as corn, sugar (called first generation) with a 20% reduction in GHG emission; corn stover, straw (called second generation or cellulosic) with a 60% GHG emission reduction; and algae (called third generation) with a 50% GHG emission reduction (Baeyens *et al.*, 2015). Based on the nested structure of RFS2, the fuels with higher emissions reduction can be used to meet the mandate for lower reduction categories (Environmental Protection Agency, 2017). The first generation biofuel production is limited in order to maintain the food security (Sharma *et al.*, 2013), and the third generation is still under research and development, and not yet commercialized (Baeyens *et al.*, 2015). Hence, the second generation has received considerable attention from different entities like governments and investors (Gupta, Verma, 2015). According to RFS2, the mandate to blend cellulosic ethanol began in 2016 which means the obligated parties *must* comply with it. Therefore, there clearly has been a need for local cellulosic ethanol production or cellulosic ethanol import. Also, since ethanol gets easily contaminated by water, investment in the infrastructure is required for its storage and blending with gasoline.

Apart from the aforementioned incentives and obligations for blending more ethanol with gasoline, there is another control factor limiting the amount of ethanol blended. According to the US Clean Air Act 1963 (CAA), all gasoline engine vehicles are permitted to use up to 10% ethanol blended with gasoline (the E10 blend); however, Flex-Fuel Vehicles (FFVs) are allowed to use up to 85% blend (E85). The maximum amount of ethanol (e.g., 10%) which can be blended with each gallon of gasoline to

be used in all gasoline engine vehicles is called Blend Wall (BW) (Renewable Fuels Association, 2015). Other intermediate blends, e.g., 15% blend (E15), can be produced by waiving the CAA under certain conditions. For instance, consumption of the E15 fuel for vehicles model year 2001 and later was allowed by the US Environmental Protection Agency in 2011. Nevertheless, almost all fuel distributed in the US is E10, though ethanol producers are interested in increasing the BW to 15% (Yacobucci, 2010).

Philadelphia Energy Solutions, the largest U.S. East Coast oil refinery, went bankrupt in 2018. The bankruptcy resulted in job loses, and calls for the amendment of some laws (Renshaw, 2018; Willette, 2018). The company blamed the RFS2 for the bankruptcy, see DiNapoli and Renshaw (2018), Simeone (2018), and Stein (2018). This provides further motivation for us to address the following questions with clear applications for production systems:

- What has been the impact of the following government policies: Renewable Fuel Standard 2 (RFS2), Tax Credit for Local ethanol blended with gasoline (TCL), Tax Credit for Imported ethanol blended with gasoline (TCI), Tariff for Local ethanol blended with gasoline (TL), Tariff for Imported ethanol blended with gasoline (TI), and Blend Wall (BW), on the SPSCs from the economic, social, and environmental points of view?
- How to determine most robust, i.e. resilient to policy change, location and production capacities for bio-refineries and blending sites in the SPSCs?

Addressing these questions will (1) shed light on how the US government may change the policies to create the SPSCs that are expected to achieve highest profit, most positive social impact, and most environmentally friendly fuel, with a minimum expenditure from its budget; (2) provide managerial insights to the SPSC investors on how to mitigate the bankruptcy chance due to policy change by creating the most robust SPSCs. To our knowledge these questions have not been addressed directly in the existing literature. However, there have been substantial literature published already on various aspects of the SPSC. We review this literature in the following subsection, and we then summarize the contributions of this paper, finally we outline the paper's content.

1.2. Literature Review and Gaps

The government policies were studied from the economic perspective, not the supply chain perspective as in this paper, by Whistance *et al.* (2016), Qiu *et al.* (2014), Aguilar *et al.* (2015), Thompson *et al.* (2009), and Babcock (2012). Whistance *et al.* (2016) study the impact of RIN price information on petroleum, biofuel, and agricultural commodity markets. Qiu *et al.* (2014) recommend directing the government policies towards increasing the demand for E85. Aguilar *et al.* (2015) argue that the majority of Americans are willing to purchase fuel with a higher amount of ethanol blend, e.g., E85. Thompson *et al.* (2009) employ a demand and supply curve to show under which conditions the RFS2 mandate is binding. The study done by Babcock (2012) shows that increasing RFS2, TCL, and TI increases the US corn price. Although all these

studies are policy related, none of them addresses the impact of government policies on the SPSC, and none uses optimization methods in contrast to this paper.

The CPSC and ESC supply chains have been the subjects of a number of recent reviews. In particular, Sahebi *et al.* (2014) and Lima *et al.* (2016) review research on the CPSCs. Both reviews emphasize the study of real-life CPSCs, new incentive schemes, and development of efficient algorithms for solving real-life CPSC optimization problems as the most promising research avenues to pursue. Mafakheri and Nasiri (2014) and Ghaderi *et al.* (2016) review literature on Biofuel Supply Chains (BSCs), which include the ESCs, they underline the importance of incorporating government policies in the models of the BSCs. Ghaderi *et al.* (2016) argue that conducting case studies of BSCs in regions with different climates, economic and political situations is a new research direction. Ba *et al.* (2016) review biomass supply chains which are parts of BSCs. Their findings show the need for optimization and efficient algorithms for large biomass supply chains which clearly apply to the SPSCs. Chukwuma (2019) advocates the integration of GIS data in mathematical programming models. This research is aimed at filling in the gaps identified by these publications by addressing the issue of developing an efficient optimization algorithm and a real-life case study presented in the accompanying part II (Ghahremanlou and Kubiak 2020) for a US state, which has not been studied from the SPSC point of view, using available real-life data, and GIS data in particular.

In order to convert a CPSC to an SPSC gasoline can be blended either with ethanol or with drop-in biofuel (i.e. biofuel compatible with the existing infrastructure) (Yue *et al.* 2014). Tong, You, and Rong (2014), Tong *et al.* (2013), and Tong, Gleeson, Rong, and You (2014) study design and operation of the SPSC with drop-in biofuel; Najmi *et al.* (2016) focus on the equilibrium models for the SPSC with drop-in biofuel. However, currently the drop-in biofuel is not being used in the US, and it is still often referred to as an energy for future (BETO 2013). Therefore, in this paper we focus on the SPSC which is created by merging an ESC with a CPSC. Andersen *et al.* (2013) and Kazemzadeh and Hu (2015) study such SPSC. Andersen *et al.* (2013) propose a strategic model to investigate the regions of the US which require investment to implement the SPSC. They also propose a detailed model to study distribution of fuel within a state. Their models focus on cost minimization, do not include uncertainty which is present in the ESCs (Awudu, Zhang, 2012; Meyer, 2007; Yue *et al.*, 2014), and do not consider ethanol imports. Kazemzadeh and Hu (2015) incorporate RFS2 and TCL in their stochastic programming model but this does not consider ethanol imported into the US. They run their computational experiments for 9 instances, however, they do not report whether they were able to find optimal solutions for those instances in reasonable time. Their research again underlines the need for new models that not only provide solutions for a large number of instances in reasonable time but also provide key insights into the policy through those solutions. This paper aims to fill the gaps by including uncertain factors, imported ethanol from other states and abroad, and a broader range of policies in the proposed models.

The SPSC needs to establish bio-refineries in some locations to process a limited supply of corn stover. The ethanol produced in those bio-refineries may be exported or shipped out to blending sites which also need to be established in some locations

for blending ethanol with gasoline and for storage. Finally, the fuel (blend) is shipped from blending sites to distribution centers to meet the demand for fuel. From that perspective, the creation of the SPSC requires solving a Multi-echelon Location--Allocation (LA) problem (Cooper, 1963; Wang, Lee, 2015; Shankar, *et al.* 2013). The LA problem is also key to applications in many other operation research (OR) areas, e.g., healthcare (Mestre *et al.*, 2015), and energy (Chukwuma, 2019; Chen, Fan, 2012; Gebreslassie *et al.*, 2012; Liu *et al.*, 2010; Serrano-Hernandez *et al.*, 2017). Azarmand and Neishabouri (2009) provide a classification of the LA problems.

The LA problem, even deterministic and single-echelon, is NP-hard. Due to the computational complexity of the LA problem the instances with large numbers of potential locations cannot be solved to optimality in reasonable time by standard solvers like Gurobi. Moreover, though some customized algorithms like a branch and bound algorithm of Kuenne and Soland (1972) have been proposed in the literature it remains to be seen whether they can even compete with off-the-shelf optimization solvers, e.g., Gurobi. Therefore, the optimal solutions for the instances with large numbers of potential locations are out of reach in practice, and thus various heuristics have been proposed in the literature (Murray, Church, 1996; Bischoff, Dächert 2009). The computational complexity grows further if the uncertainty is introduced in LA problems which is the case for this paper. To provide the reader with a perspective it is worth pointing out that Chen and Fan (2012) employ the Progressive Hedging (PH) algorithm for solving a stochastic programming model with 8 scenarios, but they only reach a solution within 0.131% from the optimum after 2 hours, though without proving that the solution is optimal. Clearly showing that the solution found may be relatively close to the optimum (more precisely to either a lower or an upper bound obtained by relaxations) in reasonable time does not mean that the optimum itself can also be found in reasonable time since the proof of optimality is typically much more time consuming due to the problem NP-hardness. Computational complexity poses a formidable barrier in policy impact analysis based on optimization since the analysis requires a large number of instances to be solved to optimality. This paper proposes a lean model to overcome this barrier.

1.3. Paper Contributions and Outline

We identified a number of gaps in the existing literature in the previous subsection. Those gaps will be filled in by our contributions that we briefly describe in this subsection leaving their details for the remaining sections and to the accompanying part II (Ghahremanlou and Kubiak 2020).

This paper studies the impact of US government policies concerning cellulosic ethanol (RFS2, TCL, TCI, TL, TI, and BW) on the SPSC. We call a six-tuple (RFS2, TCL, TCI, TL, TI, BW) of values for each RFS2, TCL, TCI, TL, TI, and BW a policy combination or just a policy. This requires *multiple* instances, thousands in this study, obtained by changing the values of mandate (RFS2), tax credits (TCL and TCI), tariffs (TL and TI), and blend wall (BW), to be solved to optimality efficiently in computational experiments. That task is impossible at the moment for the General Model (GM) based on two-stage stochastic programming presented in this paper, and

also for any other similar model presented in the literature. This becomes clear from the literature review showing that solving multi-echelon location-allocation stochastic programs to optimality is practically beyond reach even for a *single* real-life instance with close to a hundred potential locations. Therefore, in this paper we propose a Lean Model (LM) based on two-stage stochastic programming to study the impact. The LM proposes a macro level view on the flows of corn stover, ethanol and fuel which significantly reduces time required by computational experiments. We prove key relationships between optimal solutions to the GM and the LM, which help in making more robust decisions.

The rest of the paper is organized as follows: Section 2 informally describes the problem; Section 3 gives mathematical programming formulation of two models of the problem. Section 3.1 formulates the General Model (GM) as a two-stage stochastic programming model. The notations, including variables and parameters, for the formulation are defined in Appendix 5.1. Section 3.2 gives the formulation of a Lean Model (LM) which aggregates the flow variables of the GM. The aggregated variables are defined in Appendix 5.2.

2. PROBLEM STATEMENT

The problem consists in establishing an SPSC in a state of the US to meet the demand for fuel (ethanol-gasoline blend) in that state in such a way that the annual expected profit is maximized. The investment in the SPSC has been accelerated by the market created by the RFS2 mandate. Here we assume the investors take the lead in creating what is required to convert a CPSC to the SPSC according to the legislations. They also manage the SPSC. The design and operation of the SPSC are subject to various regulations: RFS2, TCL, TCI, TI, TL, and BW. The impact of these regulations on the SPSC is the main focus of this paper.

The SPSC includes harvesting sites, bio-refineries, blending sites, ethanol exporters, ethanol importers from other states and abroad, refineries, gasoline importers, and distribution centers (Fig. 1). Each county of the state has its own harvesting site and distribution center both located in the center of the county. The harvesting sites and distribution centers have their own amounts of feedstock and fuel demand respectively. Furthermore, the center of each county is a potential location for bio-refineries and blending sites. These are established by a US government loan which will be repaid during t years with an interest rate ϕ. Therefore, the problem resembles a project management type of problem which requires network design within a limited budget, where initially, the facilities locations (bio-refineries, and blending sites) are decided, and then the flows (of feedstock, ethanol, and fuel) are determined. The bio-refineries and blending sites have the same technology but different capacities and accordingly different costs to establish.

The feedstock is purchased from farmers. To keep the land fertile, only a specific portion of feedstock available can be considered for shipping to the bio-refineries; out of this amount a portion is lost due to baling and loading. The transportation network for feedstock depends on the location of bio-refineries. Bio-refineries convert a specific portion of the feedstock to ethanol which can then be shipped to blending sites and /

or sold to ethanol exporters. The transportation network for ethanol depends on the location of bio-refineries and blending sites. Blending sites receive ethanol from bio-refineries, other states, and abroad, and gasoline from refineries, and other countries, and blend the two according to the BW. The imported ethanol (from other states, and abroad) and gasoline (from refineries, and other countries) are purchased to be delivered to blending sites locations. The fuel is shipped to distribution centers to be sold to the customers. The transportation network for the fuel depends of the location of blending sites. The transportation of materials (feedstock, ethanol, and fuel) includes distance-fixed cost and distance-variable cost, and it is done by truck. We incorporate all regulations: RFS2, TCL, TCI, TI, TL, and BW in the model.

The uncertain factors in the model are: feedstock availability, feedstock price, variable transportation cost, ethanol import prices, fuel price, gasoline price, ethanol exporting price, fuel demand, number of jobs created due to different activities (construction of bio-refineries and blending sites, feedstock to ethanol conversion and ethanol-gasoline blending operations, and transportation of feedstock, ethanol and fuel).

We consider two objectives. The main objective is maximization of the annual expected profit, and the secondary is maximization of the expected number of jobs created in the state within the project lifetime of Q years. We assume the jobs are created only for construction of bio-refineries and blending sites, their operations (feedstock to ethanol conversion, and ethanol-gasoline blending), and transportation (feedstock, ethanol, and fuel).
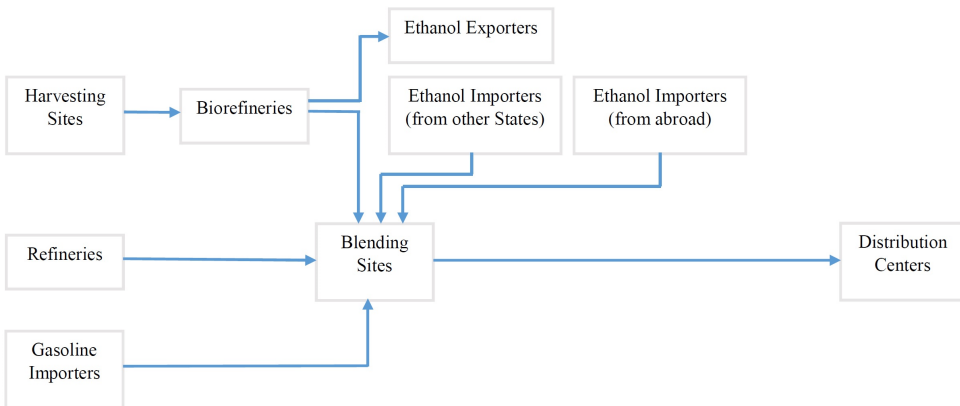

Fig. 1. *Sustainable Petroleum Supply Chain Network*

3. FORMULATION OF MODELS

In this section two types of mathematical programming models are explained. The first one is the GM which includes the details. The second one is the LM which is a conceptual model based on the GM.

3.1. General Model (GM)

We develop a two-stage stochastic programming model for the problem in this section. At the first stage the decisions regarding the locations and capacities of bio-refineries and blending sites are made before the realization of uncertain factors. At the second stage all the uncertain factors are realized and then the flow decisions are made. Therefore, the flow decisions are optimal.

The design constrains are formulated in Subsection 3.1.1, the flow constraints are given in Subsection 3.1.2, finally the objective functions are formulated in the Subsections 3.1.3–3.1.5. To streamline the presentation we leave quite heavy notations required by the variables and the parameters of the model to Appendix 5.1. The model needs to handle three different types of facilities: bio-refineries, blending sites, and distribution centers; and the flows of three different products: corn stover, ethanol, and fuel. To facilitate the presentation of the model we adopt the convention represented by the following upstream-downstream path (harvesting site j)$\rightarrow$ (bio-refinery i)$\rightarrow$ (blending site j)$\rightarrow$ (distribution center i). That is, corn stover flows from j to i, thus f_{jis} in scenario s, ethanol flows from i to j, thus e_{ijs} in scenario s, and fuel flows from j to i, thus x_{jis} in scenario s. Consequently, for instance, we use the notation d_{ji} for the distance between the harvesting side in county j and bio-refinery in county i, while d_{ij} for the distance between bio-refinery in county i and blending site in county j.

3.1.1. Design Constraints

The design constraints are related to the locations and capacities of bio-refineries and blending sites. The Constraint (1) guarantees that the total investment in the construction of bio-refineries and blending sites in the state does not exceed B, the available budget. The Constraints (2) and (3) guarantee that at most one bio-refinery and at most one blending site, respectively, is established in each county of the state.

$$\sum_m C_m \cdot \sum_i r_{mi} + \sum_n W_n \cdot \sum_j b_{nj} \leqslant B \tag{1}$$

$$\sum_m r_{mi} \leqslant 1, \qquad \forall i \in N \tag{2}$$

$$\sum_n b_{nj} \leqslant 1, \qquad \forall j \in N. \tag{3}$$

3.1.2. Flow Constraints

Suppose the production of ethanol takes place in the state. This will generate three types of flows between the counties of the state: the flow of *feedstock*, the flow of *ethanol*, and the flow of *fuel*. In this section, constraints about the flows are introduced and discussed.

The out-flow of feedstock from each county j must not exceed the amount of feedstock available for shipping from that county. This amount depends on the

sustainability factor (F), and the feedstock loss factor (L). These two factors are considered to be the same for all counties, as all the counties are located in one state, and the collection, baling, and loading method is the same. Therefore, the left-hand side of the Constraint (4) shows total feedstock available for shipping from county j in scenario s, and the right-hand side shows the out-flow of feedstock from j in scenario s.

$$(1 - L) \cdot [(1 - F) \cdot A_{js}] \geqslant \sum_i f_{jis}, \qquad \forall j \in N, \forall s \in S \qquad (4)$$

The in-flow of feedstock to county i must not exceed the capacity of bio-refinery in county i in scenario s which is guaranteed by Constraint (5). In particular this constraint guarantees that the feedstock does not flow from any other county j to i without bio-refinery in any scenario s.

$$\sum_j f_{jis} \leqslant \sum_m U_m \cdot r_{mi}, \qquad \forall i \in N, \forall s \in S \qquad (5)$$

The percentage V of all feedstock available to the bio-refinery located in county i is converted to ethanol, the left-hand side of Constraint (6). This amount of ethanol either flows from i to the counties of the state (possibly including i) with blending sites or it is sold to the exporters (o_{is}), by county i in scenario s, the right-hand side. Observe that this constraint along with Constraint (5) guarantees that ethanol flow out of a county without a bio-refinery is forbidden.

$$V \cdot \sum_j f_{jis} = \sum_j e_{ijs} + o_{is}, \qquad \forall i \in N, \forall s \in S \qquad (6)$$

The in-flow of ethanol to county j must not exceed the capacity of the blending site established in county j in scenario s which is guaranteed by Constraint (7). In particular this constraint guarantees no ethanol, either from the bio-refinery located in county i or purchased from other states or abroad, flows to j without a blending site in any scenario s.

$$\left[\sum_i e_{ijs} + h_{js} + k_{js} + g_{js} \right] \leqslant \sum_n H_n \cdot b_{nj}, \qquad \forall j \in N, \forall s \in S \qquad (7)$$

The total amount of ethanol that flows into a blending site in county j must not exceed the fraction α, the BW, of the total in-flow, ethanol and gasoline, into the blending site. This is guaranteed by Constraint (8).

$$\left[\sum_i e_{ijs} + h_{js} + k_{js} \right] \leqslant \alpha \cdot \left[\sum_i e_{ijs} + h_{js} + k_{js} + g_{js} \right], \qquad \forall j \in N, \forall s \in S \quad (8)$$

The following Constraint (9) guarantees that the total amount of ethanol purchased annually from other states ($\sum_j h_{js}$) will not exceed their total annual ethanol production capacity (E) in any scenario s.

$$\sum_j h_{js} \leqslant E, \qquad \forall s \in S \tag{9}$$

Finally consider the fuel flows. The left-hand side of the Constraint (10) equals the total amount of fuel blended by the blending site located in county j in scenario s. The right-hand side of the constraint equals the total out-flow of fuel from j to the distribution centers of counties (including j) in scenario s.

$$\left[\sum_i e_{ijs} + h_{js} + k_{js} + g_{js} \right] = \sum_i x_{jis}, \qquad \forall j \in N, \forall s \in S \tag{10}$$

The in-flow of fuel to the distribution center in county i must meet demand for fuel in i in scenario s. This is guaranteed by Constraint (11).

$$\sum_j x_{jis} = D_{is}, \qquad \forall i \in N, \forall s \in S \tag{11}$$

It is worth observing that the constraints of the model, in particular the Constraints (8), (10), and (11), do not guarantee that a feasible solution requires any positive amount of ethanol to be produced. Thus some feasible solutions may not require any production of ethanol. It would however be the mandate's task to impose the penalty on the obligated parties in order to provide the incentive for investors to produce ethanol and to establish bio-refineries and blending sites in the state. Therefore the mandate would make the solutions which do not require ethanol production to be unlikely candidates for optimal solutions. The mandate is discussed next.

3.1.3. The Mandate

The mandate is calculated as a fraction of total gasoline consumption ($\sum_j g_{js}$) in the state. The fraction is determined by the current renewable standards $\overline{R}$ and $\overline{\overline{R}}$ for the first and the second generation ethanol respectively. The mandate has been waived or changed by the government due to hitting the BW or to immaturity of the conversion technologies by adjusting the standards $\overline{R}$ and $\overline{\overline{R}}$. Since the conversion technology for the first generation ethanol has been completely developed and matured, which is not the case for the second generation we use coefficient β for $\overline{\overline{R}}$. Due to the nested structure of the RFS2 regulations the total coefficient $\overline{R} + \beta \cdot \overline{\overline{R}}$ may apply to the second generation ethanol only. Thus the mandate is defined as follows

$$M_s := \left[\overline{R} + \beta \cdot \overline{\overline{R}} \right] \cdot \sum_j g_{js}, \qquad \forall s \in S \tag{12}$$

The mandate is met by having enough Renewable Identification Numbers (RINs). One gallon ethanol is counted as one RIN. The RINs are detached when ethanol is

blended with gasoline. The number of detached RINs compared to the mandate will have one of these three outcomes: (1) the number of RINs equals the mandate; (2) the number of RINs is less than the mandate and the deficiency must be purchased from other obliged parties; (3) the number of RINs is greater than the mandate and the surplus is sold to other obliged parties. We define a variable RIN_s as the deviation of the amount of ethanol blended with gasoline from the mandate.

$$RIN_s := \left[\sum_i \sum_j e_{ijs} + \sum_j h_{js} + \sum_j k_{js} \right] - M_s, \qquad \forall s \in S \qquad (13)$$

This variable is part of the objective function we define in the next section, its contribution to the value of the objective could be zero, negative or positive depending of the Outcome (1), (2) or (3) respectively.

3.1.4. Expected Profit Maximization Objective Function

The primary objective function maximizes annual expected profit. It includes expected revenues and expected costs. The expected revenues are as follows:

– The revenue from selling extra RINs (if $RIN_s > 0$), or the cost of purchasing the deficiency (if $RIN_s < 0$)

$$R^R = P^R \cdot \sum_s RIN_s \cdot \omega_s \qquad (14)$$

where P^R is the RIN price, RIN_s is the number of RINs in scenario s defined in (13), and ω_s as the probability of scenario s. The mandate's task is to impose the penalty reflected in (12), (13), and (14) on the obligated parties in order to provide the incentive for investors to produce ethanol and to establish bio-refineries and blending sites in the state.

– The revenue from selling fuel

$$R^S = \sum_s P_s \cdot \omega_s \cdot \sum_i D_{is} \qquad (15)$$

where P_s is the fuel market price in scenario s. Observe that the revenue does not depend on any variable of the model, however it does depend on fuel demand in the state.

– The revenue generated by TCL (R^{TL}) includes two parts: one represents the ethanol produced in the state ($\sum_i \sum_j e_{ijs}$), while the other one is the ethanol imported from other states ($\sum_j h_{js}$) in scenario s. The Tax Credit per gallon of the US ethanol equals T, and η is a coefficient to take care of the government decisions to change the TCL. When $\eta < 0$, R^{TL} becomes the TL.

$$R^{TL} = \eta \cdot T \cdot \left[\sum_s \omega_s \cdot \sum_i \sum_j e_{ijs} + \sum_s \omega_s \cdot \sum_j h_{js} \right] \qquad (16)$$

– The revenue generated by TCI (R^{TC}) depends on the amount of ethanol imported from other countries ($\sum_j k_{js}$) in scenario s. The government may decide to change the Tax Credit per gallon of foreign ethanol $\overline{T}$, and thus we consider coefficient θ to take care of this change. When $\theta < 0$, R^{TC} becomes the TI.

$$R^{TC} = \theta \cdot \overline{T} \cdot \sum_s \omega_s \cdot \sum_j k_{js} \tag{17}$$

– The revenue generated by selling the ethanol produced by bio-refineries to the exporters is R^{EE}. The amount of the ethanol sold equals $\sum_i o_{is}$, and the selling price P_s^E per gallon in scenario s.

$$R^{EE} = \sum_s P_s^E \cdot \omega_s \cdot \sum_i o_{is} \tag{18}$$

The expected costs are as follows:

– This annual loan payment with an interest rate ϕ will be continued for t years.

$$C^A = \left[\frac{\phi \cdot (1+\phi)^t}{(1+\phi)^t - 1} \right] \cdot \left[\sum_m C_m \cdot \sum_i r_{mi} + \sum_n W_n \cdot \sum_j b_{nj} \right] \tag{19}$$

– The cost of purchasing feedstock

$$C^{FP} = \sum_s P_s^F \cdot \omega_s \cdot \sum_j \sum_i f_{jis} \tag{20}$$

where $\sum_j \sum_i f_{jis}$ – the total amount of feedstock shipped from harvesting sites to bio-refineries// P_s^F – price per ton (MT).

– The operating costs

$$C^O = C^{FE} \cdot \sum_s \omega_s \cdot \sum_i \left(o_{is} + \sum_j e_{ijs} \right) + C^B \cdot \sum_s \omega_s \cdot \sum_i D_{is} \tag{21}$$

include the costs of conversion of feedstock into ethanol at bio-refineries, and costs of blending ethanol and gasoline at blending sites. The former depends on the amount of ethanol produced in the state ($\sum_i (\sum_j e_{ijs} + o_{is})$), the latter on the fuel demand ($\sum_i D_{is}$) in the state in scenario s.

– The transportation cost of feedstock C^{TF}, ethanol C^{TE}, and fuel C^{TEG}

$$C^{TF} = C^{FTF} \cdot \sum_s \omega_s \cdot \sum_j \sum_i f_{jis} + \tau \cdot \sum_s C_s^{VTF} \cdot \omega_s \cdot \sum_j \sum_i f_{jis} \cdot d_{ji} \tag{22}$$

$$C^{TE} = C^{FTE} \cdot \sum_s \omega_s \cdot \sum_j \sum_i e_{ijs} + \tau \cdot \sum_s C_s^{VTE} \cdot \omega_s \cdot \sum_j \sum_i e_{ijs} \cdot d_{ij} \tag{23}$$

$$C^{TEG} = C^{FTEG} \cdot \sum_s \omega_s \cdot \sum_j \sum_i x_{jis} + \tau \cdot \sum_s C_s^{VTEG} \cdot \omega_s \cdot \sum_j \sum_i x_{jis} \cdot d_{ji} \quad (24)$$

are calculated using fixed unit costs C^{FTF}, C^{FTE}, and C^{FTEG} respectively, and variable unit cost C_s^{VTF}, C_s^{VTE}, and C_s^{VTEG} respectively. To better approximate distances d_{ji} between the counties of the state the tortuosity factor (τ) is included in the calculations.

– The cost of importing ethanol

$$C^I = \sum_s P_s^{EI} \cdot \omega_s \cdot \sum_j h_{js} + \sum_s P_s^{EE} \cdot \omega_s \cdot \sum_j k_{js} \quad (25)$$

purchasing $\sum_j h_{js}$ of ethanol from other states with unit cost of P_s^{EI}, and importing $\sum_j k_{js}$ of ethanol from other countries with unit cost of P_s^{EE} in scenario s.

– The cost of purchasing petroleum gasoline to blend with ethanol

$$C^G = \sum_s P_s^G \cdot \omega_s \cdot \sum_j g_{js} \quad (26)$$

at unit price P_s^G in scenario s.

Therefore, the primary objective function is as follows:

$$G_1 = (R^R + R^S + R^{TL} + R^{TC} + R^{EE}) - (C^A + C^{FP} + C^O + C^{TF} + C^{TE} + C^{TEG} + C^I + C^G) \quad (27)$$

3.1.5. Expected Job Created Maximization Objective Function

The secondary objective is the maximization of the expected number of jobs created in the state during the Q years lifetime of the project.

– The number of jobs created for the construction (J^C) are calculated based on the amount of the investment in the construction of bio-refineries and blending sites.

$$J^C = \sum_s J_s^{Co} \cdot \omega_s \cdot \left[\sum_m C_m \cdot \sum_i r_{mi} + \sum_n W_n \cdot \sum_j b_{nj} \right] \quad (28)$$

– The number of jobs created by the transportation of feedstock

$$J^{TF} = Q \cdot \tau \cdot \left[\sum_s \omega_s \cdot J_s \cdot \sum_j \sum_i f_{jis} \cdot d_{ji} \right] \quad (29)$$

– The number of jobs created by the transportation of ethanol

$$J^{TE} = Q \cdot \tau \cdot \left[\sum_s \omega_s \cdot J_s^{TE} \cdot \sum_i \sum_j e_{ijs} \cdot d_{ij} \right] \tag{30}$$

– The number of jobs created by the transportation of fuel

$$J^{TEG} = Q \cdot \tau \cdot \left[\sum_s \omega_s \cdot J_s^{TEG} \cdot \sum_j \sum_i x_{jis} \cdot d_{ji} \right] \tag{31}$$

– The number of jobs created by the ethanol production and blending in the state

$$J^O = Q \cdot \left[\sum_s J_s^{FE} \cdot \omega_s \cdot \sum_m \sum_i C_m \cdot r_{mi} + \sum_s J_s^B \cdot \omega_s \cdot \sum_n \sum_j W_n \cdot b_{nj} \right] \tag{32}$$

The secondary objective is as follows

$$G_2 = J^C + J^{TF} + J^{TE} + J^{TEG} + J^O. \tag{33}$$

Observe from Equation (33) that the secondary objective does not depend on the RFS2 mandate, the Blend Wall, BW, the Tax Credit for the US produced ethanol, TCL, the Tax Credit for the foreign produced ethanol, TCI, the Tariff for the US produced ethanol, TL, or the Tariff for the foreign produced ethanol, TI. Thus the changes in the government policies affecting these do not affect the maximization of the secondary objective. Also, the objective is in conflict with the primary objective since the increase in the flows weighted by the distances, which may be a result of changes in bio-refineries or blending site locations, reduces the the expected profit while at the same time it increases expected number of jobs created.

3.2. Lean Model (LM)

The GM belongs to the class of NP-hard problems and thus its optimization is very unlikely to be done efficiently. Even the design problem itself, i.e. the decision where to locate bio-refineries and blending sites in order to minimize the transportation costs of feedstock, ethanol, and fuel is NP-hard and thus difficult to solve efficiently. Therefore, it is unlikely that a single instance of the problem could be solved efficiently, even more so when multiple instances are required to be solved to show the impact of policy change by changing the values of α, β, η and θ. This motivates us to come up with a model that captures the main features of the problem, and thus makes it relevant for strategic decision making and policy analysis, though it does so at a cost of hiding less relevant details for these purposes. We propose a LM in this section that does

just that by aggregating variables over counties of the state thus hiding particular flows between them. The aggregated variables are defined in Appendix 5.2. The flows may be irrelevant at the project management level, although the total flow obtained by the aggregation is and will be part of the LM. One could argue that replacing the GM by the LM leads to the loss of precision in determining the value of the optimal solution. However, we need to keep in mind that the parameters of the models are often estimates, see for instance the discussion of the corn stover price and conversion rate parameters in Humbird *et al.* (2011), or consider the fact the cellulosic ethanol production is still in the process of commercialization which explains the lack of data pertaining to its performance. The British economist John Maynard Keynes once remarked: "it is much better to be roughly right than precisely wrong" (Dios Ortúzar, Willumsen, 2011). . Therefore, a general rule acceptable in model building is the fewer parameters the better. The LM is less parameter hungry than the GM since it does not require the unit transportation costs C^{TF}, C^{TE}, and C^{TEG}, and unit job rates J^{TF}, J^{TE}, and J^{TEG}. The LMs take much less time to solve by standard universal solvers like Gurobi.

3.2.1. Constraints

To write the LM constraints, we employ the aggregated variables, see Appendix 5.2 for definition, and closely mirror the constraints of the GM.

The number of bio-refineries with capacity level m, denoted by r_m, and the number of blending sites with capacity level n, denoted by b_n, to set up must not exceed the budget B. This is guaranteed by the following constraint

$$\sum_m C_m \cdot r_m + \sum_n W_n \cdot b_n \leqslant B \tag{34}$$

Constraints (35) and (36) guarantee that the number of bio-refineries and the number of blending sites do not exceed the number of counties, $|N|$, respectively.

$$\sum_m r_m \leqslant |N| \tag{35}$$

$$\sum_n b_n \leqslant |N| \tag{36}$$

The total shipments of corn stover to bio-refineries, denoted by f_s, in scenario s must not exceed the limit set by the supply of the corn stover in the state in s after factoring in the corn stover loss, L, and the sustainability, F, factors. This is guaranteed by the following constraint

$$f_s \leqslant A_s \cdot (1 - L) \cdot (1 - F), \qquad \forall s \in S \tag{37}$$

and they must respect the limit imposed by the total bio-refineries capacity

$$f_s \leqslant \sum_m U_m \cdot r_m, \qquad \forall s \in S \tag{38}$$

The total production of ethanol by bio-refineries, $V \cdot f_s$, in scenario s is either used in the state, e_s, or exported, o_s, which is guaranteed by the following constraint

$$V \cdot f_s = e_s + o_s, \qquad \forall s \in S \tag{39}$$

The ethanol available in the state in scenario s which is made up of the ethanol purchased from other states, h_s, or abroad, k_s, or produced internally in the state, e_s, must not exceed the fraction α, of the total demand for fuel in the state in s. This is guaranteed by

$$e_s + h_s + k_s \leqslant \alpha \cdot D_s, \qquad \forall s \in S \tag{40}$$

The amount of ethanol purchased from other states must not exceed the limit E

$$h_s \leqslant E, \qquad \forall s \in S \tag{41}$$

and the state demand must not exceed the blending capacity of the state which is guaranteed by

$$D_s \leqslant \sum_n H_n \cdot b_n, \qquad \forall s \in S \tag{42}$$

The next three constraints recognize that the shipments of corn stover can be split between the shipments between the counties, f_s^E, and internal within the counties f_s^I, Constraint (43), the latter occurs only in those counties with bio-refineries

$$f_s = f_s^E + f_s^I, \qquad \forall s \in S \tag{43}$$

the shipments of ethanol can be split between the shipments between the counties, e_s^E, and internal within the counties e_s^I, (44), the latter occurs only in those counties with both bio-refineries and blending sites

$$e_s = e_s^E + e_s^I, \qquad \forall s \in S \tag{44}$$

the shipments of fuel can be split between the shipments between the counties, x_s^E, and internal within the counties x_s^I, (45), the latter occurs only in those counties with blending sites

$$D_s = x_s^E + x_s^I, \qquad \forall s \in S \tag{45}$$

The new variables introduced in Constraints (43–45) are required to better approximate the solution of the GM by the solution to the LM which can be solved more efficiently than the GM. Finally, we add constraints that upper bound the internal shipments of feedstock, ethanol, and fuel so that a disaggregation with the same flows, internal in particular, would be possible. Section 3.3 gives more details on this.

Define $B_{is}^m = \min\{A_{is} \cdot (1 - F) \cdot (1 - L), U_m\}$ for $m = 1, 2, 3$, $i = 1, ..., |N|$, and $s \in S$. We add the following constraints

$$\sum_i r_{mi} = r_m, \qquad \forall m = 1, 2, 3 \tag{46}$$

$$\sum_m r_{mi} \leqslant 1, \qquad \forall i = 1, ..., |N| \tag{47}$$

$$f_s^I \leqslant \sum_i \sum_m B_{is}^m \cdot r_{mi}, \qquad \forall s \in S \tag{48}$$

Define $C_{js}^n = \min\{D_{js}, H_n\}$ for $n = 1, ..., 6$, $j = 1, ..., |N|$, and $s \in S$. We add the following constraints

$$\sum_j b_{nj} = b_n, \qquad \forall n = 1, ..., 6 \tag{49}$$

$$\sum_n b_{nj} \leqslant 1, \qquad \forall j = 1, ..., |N| \tag{50}$$

$$x_s^I \leqslant \sum_j \sum_n C_{js}^n \cdot b_{nj}, \qquad \forall s \in S \tag{51}$$

It is worth observing that constraints (48) and (51) place upper bounds on the internal flows of corn stover and blend respectively. These bounds may not allow us to take the advantage of the economy of scale implied by the strict concavity of bio-refineries and blending sites cost functions in general, cost-level$_k = k^{0.6} \cdot$ base cost for $k = 1, 2, 3$ (Ghahremanlou, Kubiak, 2020). The following example explains why this may happen for corn stover. Assume $U_1 = 100$, $U_2 = 200$, and the supply of corn stover, $A_i \cdot (1 - F) \cdot (1 - L)$, from the top corn stover supply counties are 80, 70, and 50. Now consider the following two solutions:

- $r_1 = 2, r_2 = 0, r_3 = 0$. By Constraint (46)

$$\sum_i r_{1i} = 2, \sum_i r_{2i} = 0, \sum_i r_{3i} = 0, \tag{52}$$

and by Constraint (47)

$$r_{1i} + r_{2i} + r_{3i} \leqslant 1, \forall i \Rightarrow r_{1j} \leqslant 1, \forall i. \tag{53}$$

Thus by (48), and $B_i^1 = \min\{A_i \cdot (1 - F) \cdot (1 - L), U_1 = 100\}$, we have $f_s^I \leqslant \sum_j \sum_m B_{is}^m \cdot r_{mi} = 150$, and clearly the internal flow of $f_s^I = 150$ is achievable by locating the two bio-refineries of capacity $U_1 = 100$ each in the two counties with the highest corn stover supplies 80 and 70 respectively.

- $r_1 = 0, r_2 = 1, r_3 = 0$. By Constraint (46)

$$\sum_i r_{1i} = 0, \sum_i r_{2i} = 1, \sum_i r_{3i} = 0, \tag{54}$$

and by Constraint (47)

$$r_{1i} + r_{2i} + r_{3i} \leqslant 1, \forall j \Rightarrow r_{2i} \leqslant 1, \forall i. \tag{55}$$

Thus by Constraint (48), and $B_i^2 = \min\{A_i \cdot (1 - F) \cdot (1 - L), U_2 = 200\}$, we have $f_s^I \leqslant \sum_i \sum_m B_{is}^m \cdot r_{mi} = 80$, and clearly the internal flow of $f_s^I = 80$ is achievable by locating a single bio-refinery of capacity $U_2 = 200$ in the county with the highest corn stover supply 80.

The former solution increases the cost of establishing bi-refineries by ignoring the economy of scale. However, the two bio-refineries established in two different counties permit higher internal flows. The latter solution, on the other hand, takes advantage of the economy of scale to reduce the cost of establishing bio-refineries, however, it reduces the internal flow since such flow is now limited to a single county where the bio-refinery is located. Therefore the reduction in transportation costs due to higher internal flows may outweigh the increase in the costs of establishing bio-refineries, thus the former solution may result in higher expected profit than the latter.

Define $E_{mn} = \min\{V \cdot U_m, \alpha \cdot H_n\}$ for $m = 1, 2, 3$, and $n = 1, ..., 6$. The following constraints limit the internal flow of ethanol

$$e_s^I \leqslant \sum_m \sum_n \sum_j E_{mn} \cdot P_{mn}^j, \qquad \forall s \in S \tag{56}$$

$$\sum_n P_{mn}^j \leqslant r_{mj}, \qquad \forall j = 1, ..., |N|, \forall m = 1, 2, 3 \tag{57}$$

and

$$\sum_m P_{mn}^j \leqslant b_{nj}, \qquad \forall j = 1, ..., |N|, \forall n = 1, ..., 6 \tag{58}$$

Observe that the two Constraints (57) and (58) imply that for $P_{mn}^j = 1$ it is necessary, but not sufficient, that both a bio-refinery of size U_m and blending site of size H_n are established in j. However, in optimality, when $r_{mj} = 1$ and $b_{nj} = 1$ then $P_{mn}^j = 1$, since e_s^I will be maximized and therefore P_{mn}^j has to reach its cap.

Finally, we define

$$M_s := \left[\overline{R} + \beta \cdot \overline{\overline{R}}\right] \cdot \left[D_s - (e_s + h_s + k_s)\right], \qquad \forall s \in S \tag{59}$$

$$RIN_s := [e_s + h_s + k_s] - M_s, \qquad \forall s \in S \tag{60}$$

Observe that only one Constraint (40) includes α, a policy parameter out of α, β, η or θ. This constraint represents the impact of BW, α, on the space of feasible solutions. A change in α may cause change in at least one of the variables e_s, h_s and k_s. Furthermore, e_s is tied with the design variable r_m and b_n. Therefore, a change in the BW may impact the long term strategic design decisions in the SPSC.

3.2.2. LM Objective Functions

The following revenues, costs, and number of jobs components of the LM objective functions, defined in Equations (61)–(78), exactly mirror those of the GM objective functions, defined in (14-26) and (28-32). The former are essentially obtained from the latter by replacing the variables of the latter by their aggregations defined in Appendix 5.2.

$$R^R = P^R \cdot \sum_s RIN_s \cdot \omega_s \tag{61}$$

$$R^S = \sum_s P_s \cdot \omega_s \cdot D_s \tag{62}$$

$$R^{TL} = \eta \cdot T \cdot \left[\sum_s \omega_s \cdot e_s + \sum_s \omega_s \cdot h_s \right] \tag{63}$$

$$R^{TC} = \theta \cdot \overline{T} \cdot \sum_s \omega_s \cdot k_s \tag{64}$$

$$R^{EE} = \sum_s P_s^E \cdot \omega_s \cdot o_s \tag{65}$$

$$C^A = \left[\frac{\phi \cdot (1+\phi)^t}{(1+\phi)^t - 1} \right] \cdot \left[\sum_m C_m \cdot r_m + \sum_n W_n \cdot b_n \right] \tag{66}$$

$$C^{FP} = \sum_s P_s^F \cdot \omega_s \cdot f_s \tag{67}$$

$$C^O = C^{FE} \cdot \sum_s \omega_s \cdot (e_s + o_s) + C^B \cdot \sum_s \omega_s \cdot D_s \tag{68}$$

$$C^{TF} = C^{FTF} \cdot \sum_s \omega_s \cdot f_s + \tau \cdot \overline{d} \cdot \sum_s C_s^{VTF} \cdot \omega_s \cdot f_s^E \tag{69}$$

$$C^{TE} = C^{FTE} \cdot \sum_s \omega_s \cdot e_s + \tau \cdot \overline{d} \cdot \sum_s C_s^{VTE} \cdot \omega_s \cdot e_s^E \tag{70}$$

$$C^{TEG} = C^{FTEG} \cdot \sum_s \omega_s \cdot D_s + \tau \cdot \overline{d} \cdot \sum_s C_s^{VTEG} \cdot \omega_s \cdot x_s^E \tag{71}$$

$$C^I = \sum_s P_s^{EI} \cdot \omega_s \cdot h_s + \sum_s P_s^{EE} \cdot \omega_s \cdot k_s \tag{72}$$

$$C^G = \sum_s P_s^G \cdot \omega_s \cdot [D_s - (e_s + h_s + k_s)] \tag{73}$$

$$J^C = \sum_s J_s^{Co} \cdot \omega_s \cdot \left[\sum_m C_m \cdot r_m + \sum_n W_n \cdot b_n \right] \tag{74}$$

$$J^{TF} = Q \cdot \tau \cdot \overline{d} \cdot \left[\sum_s J_s \cdot \omega_s \cdot f_s^E \right] \tag{75}$$

$$J^{TE} = Q \cdot \tau \cdot \overline{d} \cdot \left[\sum_s J_s^{TE} \cdot \omega_s \cdot e_s^E \right] \tag{76}$$

$$J^{TEG} = Q \cdot \tau \cdot \overline{d} \cdot \left[\sum_s J_s^{TEG} \cdot \omega_s \cdot x_s^E \right] \tag{77}$$

$$J^O = Q \cdot \sum_s J_s^{FE} \cdot \omega_s \cdot \left[\sum_m C_m \cdot r_m \right] + Q \cdot \sum_s J_s^B \cdot \omega_s \cdot \left[\sum_n W_n \cdot b_n \right] \tag{78}$$

where the distance approximation $\bar{d}$ equals either $\delta = \min_{i \neq j} d_{ij} > 0$ or $\Delta = \max_{i \neq j} d_{ij} > 0$ will be used in the next section. The Constraints (46)–(58) ensure that only external flows of corn stover, f_s^E, ethanol, e_s^E, and fuel, x_s^E, incur positive transportation costs, whereas the internal flows f_s^I, e_s^I, and x_s^I of corn stover, ethanol, and fuel respectively incur no such costs.

Therefore, the expected profit objective function and expected jobs created objective function are as follows:

$$
\begin{aligned}
L_1 = &(R^R + R^S + R^{TL} + R^{TC} + R^{EE}) - \\
&(C^A + C^{FP} + C^O + C^{TF} + C^{TE} + C^{TEG} + C^I + C^G)
\end{aligned}
\tag{79}
$$

$$
L_2 = J^C + J^{TF} + J^{TE} + J^{TEG} + J^O
\tag{80}
$$

respectively.

Observe that by Equation (80) the objective L_2 does not depend on the RFS2 mandate, the Blend Wall, BW, the Tax Credit for the US produced ethanol, TCL, the Tax Credit for the foreign produced ethanol, TCI, the Tariff for the US produced ethanol, TL, or the Tariff for the foreign produced ethanol, TI. Thus the changes in the government policies affecting these do not affect the maximization of L_2. Also, the L_2 is in conflict with the L_1 since the increase in the flows weighted by the distances, which may be a result of changes in bio-refineries or blending site locations, reduces L_1, the expected profit, while at the same time it increases L_2, the expected number of jobs created.

3.3. Relationship Between GM and LM

We have approximated the distance from one county to other county, d_{ij}, with $\bar{d}$ to make the GM completed converted to the aggregated model which is independent to counties, since without that C^{TF}, C^{TE}, and C^{TEG} in G_1, and J^{TF}, J^{TE}, and J^{TEG} in G_2 of the GM prevent full aggregation. Furthermore, to have a better bound we have used f_s^E, e_s^E, and x_s^E in these equations. To find the best value for $\bar{d}$, we used the following relations which exist between the GM and the LM in each scenario $s \in S$. Their proofs are given in Appendix 5.3:

Observation 1 *(Aggregation).* *Each feasible solution* $Y = (r_{mi}, b_{nj}, f_{jis}, e_{ijs}, o_{is}, h_{js}, k_{js}, g_{js}, x_{jis})$ *for the GM, can be converted into a feasible solution* $X = (r_m, b_n, f_s, f_s^I, f_s^E, e_s, e_s^I, e_s^E, o_s, h_s, k_s, g_s, x_s, x_s^I, x_s^E)$ *for the LM using the equations in Appendix 5.2.*

Observation 2 *(Disaggregation).* *Each optimal solution* $X = (r_m, b_n, f_s, f_s^I, f_s^E, e_s, e_s^I, e_s^E, o_s, h_s, k_s, x_s, x_s^I, x_s^E)$ *for the LM, can be converted into a feasible solution* $Y = (r_{mi}, b_{nj}, f_{jis}, e_{ijs}, o_{is}, h_{js}, k_{js}, g_{js}, x_{jis})$ *for the GM. The conversion is not unique.*

Observation 3 *For any α, β, η and θ. Let X_{min} and X_{max} be optimal solutions to the LM with $\bar{d} = \delta$ and $\bar{d} = \Delta$ respectively, then $L_1(X_{\min}) \geqslant G_1(Y) \geqslant L_1(X_{\max})$ for an optimal solution Y to G_1 of the GM.*

Observation 4 *For any α, β, η and θ. Let Z_{min} and Z_{max} be optimal solutions to the LM with $\bar{d} = \delta$ and $\bar{d} = \Delta$ respectively, then $L_2(Z_{\min}) \leqslant G_2(V) \leqslant L_2(Z_{\max})$ for an optimal solution V to G_2 of the GM.*

Finally, let us define L_1^{δ} and L_1^{Δ} to be the LM model with the objective L_1 where $\bar{d}$ is set to δ and Δ respectively. Similarly, let us define L_2^{δ} and L_2^{Δ} to be the LM model with the objective L_2 where $\bar{d}$ is set to δ and Δ respectively. We have the following observation.

Observation 5 *For any α, β, η and θ. Let Y_{min} and Y_{max} be optimal solutions to the LM with the objective $L_1^{\Delta} + L_2^{\delta}$ and $L_1^{\delta} + L_2^{\Delta}$ respectively, then $L_1^{\Delta}(Y_{\min}) + L_2^{\delta}(Y_{\min}) \leqslant G_1(Y) \leqslant L_1^{\delta}(Y_{\max}) + L_2^{\Delta}(Y_{\max})$ for an optimal solution Y to $G_1 + G_2$ of the GM.*

4. CONCLUSIONS AND FURTHER RESEARCH

We studied the impact of Renewable Fuel Standard 2 (RFS2), Tax Credit (TCL and TCI), Tariff (TL and TI), and the Blend Wall (BW) on the SPSC. We proposed the General Model (GM) for the creation of the SPSC, which falls in the category of Multi-echelon Location-Allocation (LA) problems with uncertainty. The LA problem, even deterministic and a single-echelon, is NP-hard in the strong sense, thus computationally intractable. Therefore, the GM along with all other models of general nature proposed in the literature is NP-hard in the strong sense. Hence it may be very time consuming to find an optimum for the GM for a *single* problem instance. This computational complexity makes those general models impractical as models to study policy impacts where *thousands* of instances need to be solved to optimality in computational experiments in reasonable time. Thus, we proposed a Lean Model (LM) to study the impact. The leanness comes at the cost of loosing some details about flows in the SPSC which, however, may not be that important at the stage when the SPSC is created. The concept of aggregation behind the LM as well as the model itself stand on their own and seem worthy of further research in the context of other optimization problems where general models are too time consuming to solve to optimality. We did a case study for the State of Nebraska, one of the main corn stover producers in the US, using the LM, please see the accompanying part II (Ghahremanlou, Kubiak, 2020), to provide insights for decision makers and investors who are willing to invest in the SPSC in order to make profit, to fulfill the US government regulations, and at the same time, meet the demand for fuel in the State. These insights help in arriving at robust decisions.

We would like to emphasize the need for optimization algorithms for the multi-echelon location-allocation problems, both deterministic and stochastic, capable of competing with the standard solvers like Gurobi at solving real-life instances with close to a hundred potential locations in reasonable time. The challenge has not yet

been met which creates ample opportunities for research that could impact the SPSC research. The optimal solutions to the LM can be efficiently obtained by Gurobi which is shown by our computational experiments. They provide lower and upper bounds for optimal solutions of the GM which can be viewed as a stochastic multi-echelon location – allocation problem. The bounds and their corresponding solutions may then be used to speed up optimization algorithms for the GM which is another promising path for further research. So are heuristics and metaheuristics for the GM; to our knowledge neither of them has been proposed for the GM or other general models in the literature.

The paper considers two-objectives yet it focuses on the expected annual profit as the primary objective and uses optimal solutions for that objective to evaluate the secondary objective which is the expected number of jobs created during the life-time of the project. Other approaches to dealing with multiple objectives include an objective which is a convex combination of the two or the construction and analysis of the Pareto frontier. Either creates an interesting avenue for research aimed at providing further insights into the expected profit – expected number of jobs created trade-off. However, one needs to keep in mind that those approaches are typically used for a single instance, rather than the thousands of them required to study policy impact and this may further increase complexity. All these provide many opportunities for further research on the SPSC.

ACKNOWLEDGEMENT

REFERENCES

Agarwal, A.K., 2007. Biofuels (alcohols and biodiesel) applications as fuels for internal combustion engines. *Progress in Energy and Combustion Science*, 33(3), pp. 233–271.

Aguilar, F.X., Cai, Z., Mohebalian, P., Thompson, W., 2015. Exploring the drivers' side of the "blend wall": U.S. consumer preferences for ethanol blend fuels. *Energy Economics*, 49, pp. 217–226.

Andersen, F.E., Díaz, M.S., Grossmann, I.E., 2013. Multiscale strategic planning model for the design of integrated ethanol and gasoline supply chain. *AIChE Journal*, 59(12), pp. 4655–4672.

Awudu, I., Zhang, J., 2012. Uncertainties and sustainability concepts in biofuel supply chain management: A review. *Renewable and Sustainable Energy Reviews*, 16(2), pp. 1359–1368.

Azarmand, Z., Neishabouri, E., 2009. *Location Allocation Problem*, pp. 93–109. Heidelberg: Physica-Verlag HD.

Ba, B.H., Prins, C., Prodhon, C., 2016. Models for optimization and performance evaluation of biomass supply chains: An operations research perspective. *Renewable Energy*, 87(2), pp. 977–989.

Babcock, B.A., 2012. The impact of US biofuel policies on agricultural price levels and volatility. *China Agricultural Economic Review*, 4(4), pp. 407–426.

Baeyens, J., Kang, Q., Appels, L., Dewil, R., Lv, Y., Tan, T., 2015. Challenges and opportunities in improving the production of bio-ethanol. *Progress in Energy and Combustion Science*, 47, pp. 60–88.

BETO, U., 2013. Replacing the whole barrel-to reduce U.S. dependence on oil. *Bioenergy Technologies Office, US (ed.). Office of Energy Efficiency and Renewable Energy.*

Bischoff, M., Dächert, K., 2009. Allocation search methods for a generalized class of location – allocation problems. *European Journal of Operational Research*, 192(3), pp. 793–807.

Chen, C.-W., Fan, Y., 2012. Bioethanol supply chain system planning under supply and demand uncertainties. *Transportation Research Part E: Logistics and Transportation Review*, 48(1), pp. 150–164.

Chukwuma, E., 2019. Facility location allocation modelling for bio-energy system in Anambra State of Nigeria: Integration of GIS and location model. *Renewable Energy*, 141, pp. 460–467.

Cooper, L., 1963, June. Location – allocation problems. *Operations Research*, 11(3), pp. 331–343.

DiNapoli, J. Renshaw J., 2018. Exclusive: Philadelphia Energy Solutions to file for bankruptcy – memo https://www.reuters.com/article/us-philadelphiaenergysolutions-bankruptc/exclusive-philadelphia-energy-solutions-to-file-for-bankruptcy-memo-idUSKBN1FA18P [Accessed on January 21, 2018].

Dios Ortúzar, J. de, Willumsen L.G., 2011. *References*, pp. 551–579. John Wiley & Sons, Ltd.

Duffield, J.A., Xiarchos, I.M., Halbrook, S.A., 2008. Ethanol policy: past, present, and future. *South Dakota Law Review*, 53, 425.

El-Naggar, N.E.-A., Deraz, S., Khalil, A., 2014. Bioethanol production from lignocellulosic feedstocks based on enzymatic hydrolysis: Current status and recent developments. *Biotechnology*, 13(1), pp 1–21.

Environmental Protection Agency, U. S. (2017). Overview for renewable fuel standard – renewable fuel standard program – US EPA. https://www.epa.gov/renewable-fuel-standard-program/overview-renewable-fuel-standard [accessed on June 17, 2017].

Gebreslassie, B.H., Yao, Y., You, F., 2012. Design under uncertainty of hydrocarbon biorefinery supply chains: Multiobjective stochastic programming models, decomposition algorithm, and a comparison between CVaR and downside risk. *AIChE Journal*, 58(7), pp. 2155–2179.

Ghaderi, H., Pishvaee, M.S., Moini, A., 2016. Biomass supply chain network design: An optimization-oriented review and analysis. *Industrial Crops and Products*, 94, pp. 972–1000.

Ghahremanlou, D., Kubiak, W., 2020. Impact of government policies on sustainable petroleum supply chain (spsc): A case study - Part II (The State of Nebraska). Submitted. *Decision Making in Manufacturing and Services*, 14(1).

Gupta, A., Verma, J.P., 2015. Sustainable bio-ethanol production from agro-residues: A review. *Renewable and Sustainable Energy Reviews*, 41, pp. 550–567.

Humbird, D., Davis, R., Tao, L., Kinchin, C., Hsu, D., Aden, A., Schoen, P., Lukas, J., Olthof, B., Worley, M., Sexton, D., Dudgeon, D., 2011. Process design and economics for biochemical conversion of lignocellulosic biomass to ethanol: dilute-acid pretreatment and enzymatic hydrolysis of corn stover. Technical Report, National Renewable Energy Laboratory (NREL), Golden, CO.

Independent Statistics & Analysis, U.S. Energy Information Administration, 2015, November 19. Passenger travel accounts for most of world transportation energy use. Today in Energy , https://www.eia.gov/todayinenergy/detail.php?id=23832 [accessed on February 12, 2016].

Kazemzadeh, N., Hu, G., 2015. Evaluation of the impacts of governmental policies on the biofuels supply chain design under uncertainty. *International Journal of Sustainable Economy*, 7(3), pp. 203–219.

Kessel, D.G., 2000. Global warming – facts, assessment, countermeasures. *Journal of Petroleum Science and Engineering*, 26(1), pp. 157–168.

Kuenne, R.E., Soland, R.M., 1972, December. Exact and approximate solutions to the multisource weber problem. *Mathematical Programming*, 3(1), 193–209.

Legal Information Institute, 2010. 40 CFR § 80.1406 – Who is an obligated party under the RFS program?, https://www.law.cornell.edu/cfr/text/40/80.1406 [accessed on March 20, 2017].

Lima, C., Relvas, S., Barbosa-Póvoa, A.P.F.D., 2016. Downstream oil supply chain management: A critical review and future directions. *Computers & Chemical Engineering*, 92, pp. 78–92.

Liu, P., Pistikopoulos, E.N., Li, Z., 2010. Decomposition based stochastic programming approach for polygeneration energy systems design under uncertainty. *Industrial & Engineering Chemistry Research*, 49(7), pp. 3295–3305.

Mafakheri, F., Nasiri, F., 2014. Modeling of biomass-to-energy supply chain operations: applications, challenges and research directions. *Energy Policy*, 67, pp. 116–126.

IHS Markit, 2019. Gasoline octane improvers/oxygenates. Chemical Economics Handbook, https://ihsmarkit.com/products/gasoline-octane-improvers-chemical-economics-handbook.html [accessed on October 20, 2019].

McPhail, L., Westcott, P., Lutman, H., 2011. The renewable identification number system and U.S. biofuel mandates. Economic Research Service (ERS), United States Department of Agriculture (USDA).

Mestre, A.M., Oliveira, M.D., Barbosa-Póvoa, A.P., 2015. Location–allocation approaches for hospital network planning under uncertainty. *European Journal of Operational Research*, 240(3), pp. 791–806.

Meyer, S., 2007. U.S. biofuels analysis under uncertaintiy and volatility in the biofuel industry: FAPRI stochastic modelind. "Biofuels, Food and Feed Tradeoffs" Farm Foundation, 12–13 April 2007, Saint Louis, Missouri.

Murray, A.T., Church, R.L., 1996, June. Applying simulated annealing to location-planning models. *Journal of Heuristics*, 2(1), pp. 31–53.

Najmi, A., Hamed Shakouri G., Nazari, S., 2016. An integrated supply chain: A large scale complementarity model for the biofuel markets. *Biomass and Bioenergy*, 86, pp. 88–104.

Qiu, C., Colson, G., Wetzstein, M., 2014. An ethanol blend wall shift is prone to increase petroleum gasoline demand. *Energy Economics*, 44, pp. 160–165.

Renewable Fuels Association (RFA), 2015. Blend wall, http://ethanolrfa.org/issues/blend-wall/ [accessed on August 12, 2016].

Renshaw, J., 2018. U.S. biofuels sector blasts EPA settlement with bankrupt Philadelphia refinery, https://www.reuters.com/article/us-usa-biofuels-rins/epa-settlement-with-bankrupt-philadelphia-refinery-roils-biofuels-market-idUSKCN1GP1RC [accessed on March 13, 2018].

Sahebi, H., Nickel, S., Ashayeri, J., 2014. Strategic and tactical mathematical programming models within the crude oil supply chain context – A review. *Computers & Chemical Engineering*, 68, pp. 56–77.

Serrano-Hernandez, A., Faulin, J., Belloso, J., Sawik B., 2017. The effect of environmental criteria on locating a biorefinery: A green facility location problem. *Decision Making in Manufacturing and Services*, 11(1–2), pp. 19–30.

Shankar, B.L., Basavarajappa, S., Chen, J.C., Kadadevaramath, R.S., 2013. Location and allocation decisions for multi-echelon supply chain network – A multi-objective evolutionary approach. *Expert Systems with Applications*, 40(2), pp. 551–562.

Sharma, B., Ingalls, R., Jones, C., Khanchi, A., 2013. Biomass supply chain design and analysis: basis, overview, modeling, challenges, and future. *Renewable and Sustainable Energy Reviews*, 24, pp. 608–627.

Simeone, C., 2018. Part 1: Philadelphia energy solutions bankruptcy basics, https://kleinmanenergy.upenn.edu/blog/2018/02/02/part-1-philadelphia-energy-solutions-bankruptcy-basics [KCEP blog] [accessed on 02 May 2018].

Stein, K., 2018. An entirely predictable bankruptcy, https://www.instituteforenergyresearch.org/renewable/entirely-predictable-bankruptcy [accessed on January 23, 2018].

Thompson, W., Meyer, S., Westhoff, P., 2009. Renewable identification numbers are the tracking instrument and bellwether of U.S. biofuel mandates. Identifikationsnummern für Erneuerbare Energien als Mittel der Rückverfolgbarkeit und Indikator für US-amerikanische Biokraftstoffverpflichtungen. Les numéros d'identification des carburants renouvelables sont les instruments et baromètres de suivi des obligations légales concernant les biocarburants aux États-Unis. *EuroChoices*, 8(3), pp. 43–50.

Tong, K., Gong, J. Yue, D., You, F., 2013. Stochastic programming approach to optimal design and operations of integrated hydrocarbon biofuel and petroleum supply chains. *ACS Sustainable Chemistry & Engineering*, 2(1), pp. 49–61.

Tong, K., Gleeson, M.J., Rong, G., You, F., 2014. Optimal design of advanced drop-in hydrocarbon biofuel supply chain integrating with existing petroleum refineries under uncertainty. *Biomass and Bioenergy*, 60, pp. 108–120.

Tong, K., You, F., Rong, G., 2014. Robust design and operations of hydrocarbon biofuel supply chain integrating with existing petroleum refineries considering unit cost objective. *Computers & Chemical Engineering*, 68, pp. 128–139.

Wang, K.-J., Lee, C.-H., 2015. A revised ant algorithm for solving location–allocation problem with risky demand in a multi-echelon supply chain network. *Applied Soft Computing*, 32, pp. 311–321.

Whistance, J., Ripplinger, D., Thompson, W., 2016. Biofuel-related price transmission using renewable identifi- cation number prices to signal mandate regime. *Energy Economics*, 55, pp. 19–29.

Willette, J.K., 2018. Bankruptcy judge rules on Philadelphia Energy Solutions case, https://www.farmprogress.com/energy/bankruptcy-judge-rules-philadelphia-energy-solutions-case [accessed on April 06, 2018].

Yacobucci, B.D., 2010. Intermediate-level blends of ethanol in gasoline, and the ethanol 'blend wall'. Congressional Research Service Report (R40445).

Yan, L., 2012. Analysis of the international oil price fluctuations and its influencing factors. *American Journal of Industrial and Business Management*, 2(2), pp. 39–46.

Yue, D., You, F., Snyder, S.W., 2014. Biomass-to-bioenergy and biofuel supply chain optimization: Overview, key issues and challenges. *Computers & Chemical Engineering*, 66, pp. 36–56.

5. APPENDIX

5.1. Notations

Sets	
N	set of counties
S	set of scenarios
Indices	
i	county index, $i \in N$
j	county index, $j \in N$
m	capacity level of bio-refineries $m \in \{1, 2, 3\}$
n	capacity level of blending sites $n \in \{1, 2, 3, 4, 5, 6\}$
s	scenario index, $s \in S$

Decision variables	
Continuous non-negative variables for scenario $s \in S$	
o_{is}	amount of ethanol sold to exporter from bio-refinery in county j (gal)
e_{ijs}	amount of ethanol shipped from bio-refinery in county i to blending site in county j (gal)
f_{jis}	amount of feedstock (corn stover) shipped from harvesting site in county j to bio-refinery in county i (MT)
g_{js}	amount of petroleum gasoline purchased for blending with ethanol in blending site in county j (gal)
h_{js}	amount of ethanol purchased from other states for blending with gasoline in blending site in county j (gal)
k_{js}	amount of ethanol purchased from other countries for blending with gasoline in blending site in county j (gal)
x_{jis}	amount of fuel (ethanol-gasoline blend) shipped from blending site in county j to distribution center in county i (gal)
Binary variables	
b_{nj}	equals 1 if a blending site with capacity level n is set up in county j
r_{mi}	equals 1 if a bio-refinery with capacity level m is set up in county i

Parameters	
Harvesting sites	
A_{js}	amount of feedstock (corn stover) at county j in scenario s (MT)
F	sustainability factor for harvesting site in each county
L	feedstock loss factor due to baling and loading in each county

<div align="center">Bio-refineries and blending sites - design</div>

B	amount of loan to set up bio-refineries and blending sites in the state under study (\$)
t	loan payback period (y)
ϕ	interest rate of the loan received for establishing bio-refineries and blending sites
C_m	cost to set up a bio-refinery with capacity level m
W_n	cost to set up a blending site with capacity level n (\$)
U_m	capacity of a bio-refinery with capacity level m (MT)
H_n	capacity of a blending site with capacity level n (gal)
Q	lifetime of the bio-refineries and blending sites (y)

<div align="center">Bio-refineries and blending sites – operation</div>

C^{FE}	conversion cost per unit of ethanol produced (\$/gal)
V	conversion factor for bio-refineries (corn stover to ethanol) (gal/MT)
C^B	blending cost per unit of ethanol-gasoline blend produced (\$/gal)
E	maximum amount of ethanol can be imported from other states (gal)
J_s^{Co}	number of jobs created per dollar of expenditures on construction of bio-refineries and blending sites in scenario s ($job/\$ \cdot$ y)
J_s^{FE}	number of jobs created annually per dollar of expenditures on conversion operation in scenario s ($job/\$ \cdot$ y)
J_s^B	number of jobs created annually per dollar of expenditures on blending operation in scenario s ($job/\$ \cdot$ y)

<div align="center">Unit prices</div>

P_s^F	price of feedstock purchased in scenario s (\$/MT)
P_s^E	price of ethanol sold to the exporter in scenario s (\$/gal)
P_s^{EI}	price of ethanol purchased from other states in scenario s (\$/gal)
P_s^{EE}	price of ethanol purchased from other countries in scenario s (\$/gal)
P_s^G	price of gasoline (from crude oil) purchased in scenario s (\$/gal)
P_s	price of fuel (ethanol-gasoline blend) sold to the distribution centers in scenario s (\$/gal)
P^R	price of RIN (\$/RIN)

<div align="center">Distribution centers</div>

D_{is}	fuel (ethanol-gasoline blend) demand at county i in scenario s (gal)

<div align="center">Transportation</div>

C^{FTF}	feedstock fixed transportation cost (\$/MT)
C_s^{VTF}	feedstock variable transportation cost in scenario s (\$/$MT \cdot$ mi)
C^{FTEG}	fuel (ethanol-gasoline blend) fixed transportation cost (\$/gal)
C_s^{VTEG}	fuel (ethanol-gasoline blend) variable transportation cost in scenario s (\$/gal $\cdot$ mi)
C^{FTE}	ethanol fixed transportation cost (\$/gal)
C_s^{VTE}	ethanol variable transportation cost in scenario s (\$/gal $\cdot$ mi)
d_{ij}	direct distance from county i to county j (mi)

J_s number of jobs created for feedstock transported in scenario s $(job/MT \cdot \text{mi} \cdot \text{y})$

J_s^{TEG} number of jobs created for fuel transported in scenario s $(job/\text{gal} \cdot \text{mi} \cdot \text{y})$

J_s^{TE} number of jobs created for ethanol transported in scenario s $(job/\text{gal} \cdot \text{mi} \cdot \text{y})$

τ tortuosity factor (for converting direct distance to real distance)

<center><i>Policies</i></center>

M_s amount of ethanol mandate for the state under study in scenario s (gal)

$\overline{R}$ renewable fuel standard for first generation of ethanol

$\overline{\overline{R}}$ renewable fuel standard for second generation of ethanol

T tax credit per unit of ethanol (locally produced and/or imported from other states) blended with petroleum gasoline (\$/gal)

$\overline{T}$ tax credit per unit of ethanol (imported from other countries) blended with gasoline (coming from crude oil) (\$/gal)

α blend wall

β coefficient of current ethanol mandate

η coefficient of current tax credit for blended ethanol (that locally produced and/or imported from other states)

θ coefficient of current tax credit for blended ethanol (that imported from other countries)

RIN_s amount of RINs for scenario s

<center><i>General</i></center>

ω_s probability of scenario s

<center>**Objective function components**</center>

<center><i>Revenues</i> (\$)</center>

R^R total revenue resulting from RINs sold

R^S total revenue resulting from fuel (ethanol-gasoline blend) sold

R^{TL} total revenue resulting from tax credit for blended ethanol (locally produced)

R^{TC} total revenue resulting from tax credit for blended ethanol (imported from other countries)

R^{EE} total revenue resulting from ethanol sold to the exporter

<center><i>Costs</i> (\$)</center>

C^A total cost resulting from the annual loan payback

C^{FP} total cost resulting from feedstock purchased

C^O total cost resulting from bio-refineries and blending sites operation (conversion and blending)

C^{TF} total cost resulting from transportation of feedstock (corn stover)

C^{TE} total cost resulting from transportation of ethanol from bio-refineries to blending sites

C^{TEG} total cost resulting from transportation of fuel (ethanol-gasoline blend) from blending sites to distribution centers

C^I	total cost resulting from ethanol imported from other states and other countries
C^G	total cost resulting from gasoline (from crude oil) purchased

<div align="center">Jobs (job)</div>

J^C	total jobs resulting from construction of bio-refineries and blending sites
J^{TF}	total jobs resulting from transportation of feedstock (corn stover)
J^{TE}	total jobs resulting from transportation of ethanol from bio-refineries to blending sites
J^{TEG}	total jobs resulting from transportation of fuel (ethanol-gasoline blend) from blending sites to distribution centers
J^O	total jobs resulting from bio-refineries and blending sites operation (conversion and blending)

5.2. Aggregated Variables

The variables of the Lean Model (LM) and the General Model (GM) are related by Equations (81)–(95) listed below. The equations informally state that the value of a variable in the LM is obtained by an *aggregation* of the values of variables in the GM over all counties, or conversely the values of variables in the GM are obtained by a *disaggregation* of the value of the variable in the LM. The disaggregation is not unique.

$$\sum_i r_{mi} = r_m, \qquad \forall m \tag{81}$$

$$\sum_n b_{nj} = b_n, \qquad \forall n \tag{82}$$

$$\sum_j \sum_i x_{jis} = x_s, \qquad \forall s \in S \tag{83}$$

$$\sum_j \sum_{i=j} x_{jis} = x_s^I, \qquad \forall s \in S \tag{84}$$

$$\sum_j \sum_{i \neq j} x_{jis} = x_s^E, \qquad \forall s \in S \tag{85}$$

$$\sum_j h_{js} = h_s, \qquad \forall s \in S \tag{86}$$

$$\sum_j k_{js} = k_s, \qquad \forall s \in S \tag{87}$$

$$\sum_i \sum_j e_{ijs} = e_s, \qquad \forall s \in S \tag{88}$$

$$\sum_{i} \sum_{j=i} e_{ijs} = e_s^I, \qquad \forall s \in S \tag{89}$$

$$\sum_{i} \sum_{j \neq i} e_{ijs} = e_s^E, \qquad \forall s \in S \tag{90}$$

$$\sum_{j} g_{js} = g_s, \qquad \forall s \in S \tag{91}$$

$$\sum_{j} \sum_{i} f_{jis} = f_s, \qquad \forall s \in S \tag{92}$$

$$\sum_{j} \sum_{i=j} f_{jis} = f_s^I, \qquad \forall s \in S \tag{93}$$

$$\sum_{j} \sum_{i \neq j} f_{jis} = f_s^E, \qquad \forall s \in S \tag{94}$$

$$\sum_{i} o_{is} = o_s, \qquad \forall s \in S \tag{95}$$

The total demand in the state in scenario s.

$$\sum_{i} D_{is} = D_s, \qquad \forall s \in S \tag{96}$$

The total supply of corn stover in the state in scenario s.

$$\sum_{j} A_{js} = A_s, \qquad \forall s \in S \tag{97}$$

5.3.   Proofs of Relationship Between GM and LM

Proof. Observation 1 – We set $P_{mn}^i = 1$ if and only if $r_{mi} = 1$ and $b_{in} = 1$ for all m, n, and $i \in N$. Thus $P_{mn}^i = 1$ if and only if a bio-refinery of size U_m and a blending site of size H_n are both set up in county i. By Equations (1), (2), and (3) which are satisfied by Y, and using Definitions (81) and (82) (see Appendix 5.2) we have Constraints (34), (35), (36), (46), (47), (49), (50), (57), and (58) satisfied by X. By Definitions (88), (92), (95), and (97), the Flow Constraints (4), (5), and (6) met by Y imply that X meets (37), (38), and (39). Definitions (92), (93), and (94) imply (43) for the corn stover flow in X, (88), (89), and (90) imply (44) for the ethanol flow in X, and Definitions (83), (84), and (85) along with Constraints (46) and (47) met for Y imply (45) for the fuel flow in X. Constraints (7), (8), (9), (10), and (11) satisfied by X by Definitions (88), (86), (91), (87), (95), and (96) imply that Constraints (40), (41), and (42) are satisfied by X.

The maximum amount of feedstock shipped internally within county i with a bio-refinery of size U_m equals $B_{js}^m = \min\{A_{js} \cdot (1 - F) \cdot (1 - L), U_m\}$ for $m = 1, 2, 3$, $j = 1, ..., |N|$, and $s \in S$, thus the actual amount, f_s^I, shipped in Y meets the constraint

(48) in X by definition (93). The maximum amount of fuel shipped internally within county i with a blending site of size H_n equals $C_{js}^n = \min\{D_{js}, H_n\}$ for $n = 1, ..., 6$, $j = 1, ..., |N|$, and $s \in S$. Therefore, the actual amount, x_s^I, shipped in Y meets Constraint (51) in X by Definition (84). Finally, the maximum amount of ethanol shipped internally within county i with a bio-refinery of size U_m and a blending site of size H_n equals $E_{mn} = \min\{V \cdot U_m, \alpha \cdot H_n\}$ for $m = 1, 2, 3$, and $n = 1, ..., 6$, thus the actual amount, e_s^I, shipped in Y meets Constraint (56) in X by Definition (89).

Proof. Observation 2 – Let $X = (r_m, b_n, f_s, f_s^I, f_s^E, e_s, e_s^I, e_s^E, o_s, h_s, k_s, x_s, x_s^I, x_s^E)$ be an optimal solution to the LM. We obtain a feasible solution to the GM as follows. First, locate a bio-refinery of size U_m in county j if and only if $r_{mj} = 1$ in X, and locate a blending site of size H_n in county j if and only if $b_{nj} = 1$ in X. Constraints (46), (47), (49), and (50) of the LM guarantee that these locations satisfy Constraints (2) and (3) of the GM. Since X satisfies Budged Constraint (34) so does Y satisfies (1) in the GM. Let Bio and Bl be the sets of counties with bio-refineries and blending sites respectively in Y. Consider the flow of corn stover. By (48) we get

$$\sum_{i \in Bio} \min\{A_{is} \cdot (1 - F) \cdot (1 - L), U^i\} \geqslant f_s^I, \tag{98}$$

which guarantees that the locations in Bio ensure the internal flow f_s^I required by X, here U^i is the capacity of bio-refinery located in i. Thus, we can obtain an internal flow f_{iis} for each county i so that the total internal flow equals $f_s^I = \sum_i f_{iis}$ in scenario s. Once the internal flows of corn stover have been fixed we can calculate the external flow f_{ijs} of corn stover from county i to county j, $i \neq j$, to meet the total external flow f_s^E required by X. The flows f_{ijs} can be calculated by solving a minimum cost network flow problem N_f with a given flow f_s^E to minimize the corn stover variable transportation costs. The network node capacities (supply of feedstock in county i and capacity of bio-refinery in county i) are determined by Constraints (4) and (5) and further adjusted by the internal flows f_{iis} which are fixed before the external flows f_{ijs} for different i and j are calculated. Constraints (37), (38), and (43) guarantee that the flows f_{ijs} are feasible for Y. The feedstock flows also determine the amount of ethanol $V \cdot f_{js}$ produced in county j which will subsequently serve as the upper limit on the amount of ethanol e_{js} (produced in county j and used in the state) in the network flow problem N_e used to calculate ethanol flows consistent with the corn stover flows f_{ijs} in order to satisfy (6) in Y by (39) in X.

Now consider the flow of fuel. By (51) we get

$$\sum_{i \in Bl} \min\{D_{is}, H^i\} \geqslant x_s^I, \tag{99}$$

which guarantees that the locations in Bl ensure the internal flow x_s^I required by X, here H^i is the capacity of blending site located in i. Thus, we can obtain an internal flow x_{iis} for each county i so that the total internal flow equals $x_s^I = \sum_i x_{iis}$ in scenario s. Once the internal flows have been fixed we can calculated the external flow x_{ijs} of fuel from county i to county j, $i \neq j$, to meet the total external flow x_s^E required by X.

The flows x_{ijs} can be calculated by solving a minimum cost network flow problem N_x with a given flow x_s^E to minimize the fuel variable transportation costs. The network node capacities (fuel demand in county i) are determined by Constraint (11) and further adjusted by the internal flows x_{iis} which are fixed before the external flows x_{ijs} for different i and j are calculated. Constraint (45) guarantees that the flows x_{ijs} are feasible for Y. The fuel flows also determine the amount of fuel x_{js} produced in county j which will subsequently serve as the upper limit on the amount of ethanol e_{is} in the network flow problem N_e used to calculate ethanol flows consistent with the fuel flows x_{ijs} in order to satisfy (10) in Y by (45) in X.

Finally, consider the flow of ethanol. Since X is optimal we have $P_{mn}^j = 1$ if and only if a bio-refinery of size U^m and a blending site of size H^n are both set up in county j. Thus, the internal flow of ethanol occurs only in counties having both a bio-refinery and a blending site set up, and (56) guarantees that those counties in $Bio \cap Bl$ ensure the amount e_s^I required by X. The remaining ethanol, e_s^E, is shipped from the counties $i \in Bio \setminus Bl$ to the counties $j \in Bl \setminus Bio$. The flow e_{ijs} can be calculated, once the internal flows e_{iis} have been fixed, by solving a minimum cost network flow problem N_e with a given flow e_s^E to minimize the ethanol variable transportation costs. The network node capacities are determined by the constraint (7), and the flows f_{js} and x_{js} that have already been calculated. They are further adjusted by the internal flows e_{iis}, which are fixed before the external flows e_{ijs} are calculated. The constraint (45) guarantees that the flows e_{ijs} are feasible for Y.

At this point we have shown how to obtain feasible corn stover flows f_{ijs}, ethanol flows e_{ijs}, and fuel flows x_{ijs} for the GM. Now, we can use equations $V \cdot f_{is} = e_{i,s} + o_{i,s}$ (see the constraint (6) in the GM), where $V \cdot f_{is} \leqslant V \cdot U^i$ and $\sum_i e_{is} = e_s$ and $\sum_i o_{is} = o_s$ to calculate the amount of ethanol o_{is} exported from county i in scenario s. Thus, we get (6) satisfied by Y since X satisfies (38) and (39). Moreover, we can use equations $e_{is} + h_{is} + k_{is} + g_{is} = D_{is}$ (see Constraint (10) in the GM), and inequalities $e_{is} + h_{is} + k_{is} + g_{is} \leqslant \sum_n H_n \cdot b_{ni}$ (see Constraint (7) in the GM), $e_{is} + h_{is} + k_{is} \leqslant \alpha \cdot (e_{is} + h_{is} + k_{is} + g_{is})$ (see Constraint (8) in the GM), and $\sum_i h_{is} \leqslant E$ (see Constraint (9) in the GM), $\sum_i h_{is} = h_s$, $\sum_i e_{is} = e_s$ and $\sum_i k_{is} = k_s$ to calculate the amount of ethanol h_{is} purchased from other states, or abroad, k_{is}, by county i, and the amount of gasoline, g_{is}, purchased by county i. Clearly, these amounts can be calculated, for instance, to minimize the cost of the purchases. Thus we get (7), (8), and (9) satisfied by Y since X satisfies (40), (41) and (42). Therefore, we get solution Y that is feasible for the GM. Clearly, the solution is not unique since the flows, for instance, can be calculated differently.

Proof. Observation 3 – By Observation 1, Y can be converted into a feasible solution X to the LM. The only difference between $L_1(X)$ and $G_1(Y)$ consists in replacing the distances d_{ij} with a single distance δ, please check (69)–(71), (22)–(24), (75)–(77), and (29)–(31).

However we have

$$\delta \cdot f_s^E \leqslant \sum_j \sum_{i \neq j} f_{jis} \cdot d_{ji}, \qquad \forall s \in S,$$

$$\delta \cdot e_s^E \leqslant \sum_j \sum_{i \neq j} e_{ijs} \cdot d_{ij}, \qquad \forall s \in S,$$

$$\delta \cdot x_s^E \leqslant \sum_j \sum_{i \neq j} x_{jis} \cdot d_{ji}, \qquad \forall s \in S.$$

Therefore, $L_1(X) \geqslant G_1(Y)$. For an optimal solution X_{min} to the LM we have $L_1(X_{min}) \geqslant L_1(X)$, thus $L_1(X_{min}) \geqslant G_1(Y)$.

By Observation 2, X_{max} can be converted into a feasible solution Y' to the GM. The only difference between $L_1(X_{min})$ and $G_1(Y')$ consists in replacing the distances d_{ij} with a single distance Δ, please check (69)–(71), (22)–(24), (75)–(77), and (29)–(31). However, we have

$$\Delta \cdot f_s^E \geqslant \sum_j \sum_i f_{jis} \cdot d_{ji}, \qquad \forall s \in S,$$

$$\Delta \cdot e_s^E \geqslant \sum_j \sum_i e_{ijs} \cdot d_{ij}, \qquad \forall s \in S,$$

$$\Delta \cdot x_s^E \geqslant \sum_j \sum_i x_{jis} \cdot d_{ji}, \qquad \forall s \in S.$$

Therefore, $G_1(Y') \geqslant L_1(X_{max})$. For an optimal solution Y to the GM we have $G_1(Y) \geqslant G_1(Y')$, thus $G_1(Y) \geqslant L_1(X_{max})$. We proved that $L_1(X_{min}) \geqslant G_1(Y) \geqslant L_1(X_{max})$ as required.

Proof. Observation 4 – The proof is similar to the proof of Observation 3 thus it will be omitted.

Proof. Observation 5 – By Observation 1, Y can be converted into a feasible solution X to the LM. The only difference between $L_1^\delta(X) + L_2^\Delta(X)$ and $G_1(Y) + G_2(Y)$ consists in replacing the distances d_{ij} with a single distance δ in the objective L_1 and a single distance Δ in the objective L_2, please check (69–71), (22–24), (75–77), and (29–31). Therefore, $L_1^\delta(X) + L_2^\Delta(X) \geqslant G_1(Y) + G_2(Y)$. For an optimal solution Y_{min} we have $L_1^\delta(Y_{max}) + L_2^\Delta(Y_{max}) \geqslant L_1^\delta(X) + L_2^\Delta(X) \geqslant G_1(Y) + G_2(Y)$.

By Observation 2, Y_{min} can be converted into a feasible solution Y' to the GM. The only difference between $L_1^\Delta(X) + L_2^\delta(X)$ and $G_1(Y') + G_2(Y')$ consists in replacing the distances d_{ij} with a single distance Δ in the objective L_1 and a single distance δ in the objective L_2, please again check (69–71), (22–24), (75–77), and (29–31). Therefore, $L_1^\Delta(X) + L_2^\delta(X) \leqslant G_1(Y') + G_2(Y')$. For an optimal solution Y we have $L_1^\Delta(Y_{min}) + L_2^\delta(Y_{min}) \leqslant G_1(Y') + G_2(Y') \leqslant G_1(Y) + G_2(Y)$.