# Assessing the Supplier Role of Selected Fresh Produce Value Chains in the United States

Houtian Ge[1], Patrick Canning[2], Stephan Goetz[3] and Agnes Perez[4]

[1]Department of Agricultural Economics, Sociology and Education, Pennsylvania State University, University Park, PA 16802, USA. E-mail: hug19@psu.edu

[2]Economic Research Service, United States Department of Agriculture, Washington, DC 20250, USA. E-mail: pcanning@ers.usda.gov

[3]Department of Agricultural Economics, Sociology and Education, Pennsylvania State University, University Park, PA 16802, USA. E-mail: sgoetz@psu.edu

[4] Economic Research Service, United States Department of Agriculture, Washington, DC 20250, USA. E-mail: acperez@ers.usda.gov

Selected Paper prepared for presentation at the 2015 Agricultural & Applied Economics Association and Western Agricultural Economics Association Annual Meeting, San Francisco, CA, July 26-28

# Assessing the Supplier Role of Selected Fresh Produce Value Chains in the United States

Houtian Ge
Department of Agricultural Economics and Rural Sociology
Penn State University
hug19@psu.edu

Patrick Canning
Economic Research Service
U.S. Department of Agriculture
pcanning@ers.usda.gov

Stephan Goetz
Department of Agricultural Economics and Rural Sociology
Penn State University
sgoetz@psu.edu

Agnes Perez
Economic Research Service
U.S. Department of Agriculture
acperez@ers.usda.gov

*Selected Paper prepared for presentation at the Agricultural & Applied Economics Association's 2015 AAEA & WAEA Joint Annual Meeting, San Francisco, CA, July 26-28, 2015*

**Abstract:** This paper examines the roles and potential market impacts of regional food hub development through consideration of the supplier roles in selected fresh produce value chains in the United States. In contrast to much of the literature, the effects of economies of scale was embedded in the model and annual production statistics are broken out into seasonal marketing segments to more accurately account for the highly variable geographic disposition of annual fresh produce production. The hub optimization problem was mathematically formulated as a mixed integer linear programming model with an objective to minimize the total costs associated with fresh produce assembly and hub operations. Our results suggest that scale economies have significant effect on the optimal solutions of hub numbers, locations and sizes. This paper provides a replicable empirical framework to conduct impact assessments for regional and local food systems.

Key words: Operation research, facility location, optimization, simulation, fresh produce

The views expressed here are those of the authors, and may not be attributed to the Economic Research Service, the U.S. Department of Agriculture, or Penn State University. Partial funding under USDA NIFA grant no. 2011-68004-30057 is gratefully acknowledged.

## Introduction

Food security—access by all people at all times to enough food for an active and healthy life—is one of several conditions necessary for a population to be healthy and well nourished. As a recent nationwide investigation shows, food insecurity is a serious challenge facing millions of Americans (Coleman-Jensen, Nord and Singh 2013). As domestic and worldwide population grows, there is increased interest among consumers, private and public decision makers regarding the sustainability of food supply chains (Nicholson, Gómez and Gao 2011). Consequently, the structure and optimization of key agricultural supply chains is of growing importance.

To this end, the U.S. Department of Agriculture (USDA) administers or is examining a variety of policies and programs to (i) accommodate the growing demand for food that accompanies domestic and worldwide population growth, (ii) enhance food access and affordability for disadvantaged communities, and (iii) encourage sustainable growth of the food system. To accommodate the growing demand in food and enhance the food access and affordability for disadvantages communities while also benefiting farmers, food suppliers and others in the food system, USDA has been actively encouraging a strengthening of regional and local food systems as a focal point of this effort and the development of regional food hubs is an important component of this strategy. In 2009, the USDA launched an initiative “Know Your Farmer, Know Your Food” with an aim to strengthen the critical connection between farmers and consumers and support local and regional food systems.

In the past 10 years, demand for locally grown food has increased dramatically (Jablonski, et al. 2011). As policymakers, researchers, and practitioners seek new opportunities to support food security and rural development, interest in regional and local food systems continues to grow (Boys and Hughes 2013; Clancy 2010; King et al. 2010; Martinez et al. 2010; O'Hara and Pirog 2013). Especially, the role of small- and medium-scale producers in developing local and regional food systems has attracted renewed attention, as their importance in supplying food markets gains recognition (Low and Vogel 2011). Despite local food systems’ purported potential to increase farm sales, particularly for small and mid-scale producers, and support rural economic development, one of the recurring impediments faced by producers is the lack of distribution infrastructure and services and limited marketing capacity. Many farmers and ranchers—especially smaller operations—are challenged by the lack of distribution and processing infrastructure of appropriate scale that would give them wider access to retail, institutional, and commercial foodservice markets (Martinez et al. 2010). These small- and medium-scale producers are too small to compete effectively in traditional wholesale supply chains. They often lack the volume and consistent supply necessary to attract retail and foodservice customers. Furthermore, due to limited staff or lack of experience, they are not always able to devote the attention necessary to develop successful business relationships and linkages with key wholesale buyers or have the resources to develop an effective marketing strategy.

Regional food hubs have emerged as an effective way to overcome these infrastructural and market barriers. USDA’s Agricultural Marketing Service (AMS) has as its working definition of food hubs “A centrally located facility with a business management structure facilitating the aggregation, storage, processing, distribution, and/or marketing of locally/regionally produced food products.” (www.ams.usda.gov/AMSv1.0/FoodHubs). AMS has published a Regional Food Hub Resource Guide (Barham, et.al. 2012) that highlights programs at USDA and elsewhere for supporting regional food hub development. For those smaller and mid-sized producers who wish to scale up their operations or diversify their market channels, food hubs offer a combination of production, distribution, and marketing services that allows them to take greater advantage of the growing demand for locally and regionally grown food in large volume markets that would be difficult or impossible to access on their own. From this point of view, food hubs create new marketing opportunities for farmers and ranchers to expand the scope of their consumer market, providing a critical supply chain link for rural communities and farmers to reach consumers interested in purchasing local products.

This paper examines the roles and potential market impacts of regional food hub development through consideration of the supplier roles in selected fresh produce value chains in the United States. Fresh produce suppliers--comprising shippers, importers, wholesalers, distributors, and brokers--provide a range of marketing services for domestic and international growers supplying fresh produce to U.S. markets, and to their retail and foodservice customers serving the U.S. market (Cook, 2011). The annual U.S. retail value of fresh produce reflects costs distributed roughly equally between the farm value, the value of supplier services, and the value of services from retailing and foodservice establishments. Whereas research and analysis on the production and retailing segments of fresh produce value chains is extensive and well developed, produce supplier studies have been limited and have typically been narrow in scope, leaving a gap in our understanding of these important value chains.

To narrow down this gap, this research seeks to characterize and model transportation and supplier logistic (TSL) services in selected U.S. fresh produce markets, focusing on the subset of fresh produce commodities that are highly perishable (e.g., excluding commodities that can keep fresh in long term cold storage such as apples and cabbage). Both the scale and locations of these TSL hubs are designed to minimize the total costs of assembling U.S. production and handling the supplier logistics of distributing this supply to all final markets. Important outputs produced by this model include supplier cost functions that characterize the extent of scale economies across fresh produce TSL hubs, and shipping cost functions that characterize the impedance based transportation costs for shipments between production nodes and supplier hubs.

## Literature Review

Hub-and-spoke networks became an important field of discrete location research (Camargo and Miranda 2012). In hub-and-spoke networks, direct transportation of flows between pairs of origin-destination nodes is usually extremely costly. As an alternative, flows from different origins but addressed to the same destination can be consolidated at transshipment nodes, known as hubs, prior to be routed, sometimes via other hubs, towards their destinations. Hubs are then responsible for flow aggregation and redistribution. Hub location modeling is largely explained by its widespread use in air transportation, telecommunication, ground freight transportation and other transportation scenarios (Aros-Vera, Marianov and Mitchel 2013; Horner and O’Kelly 2001; Dantrakul, Likasiri and Pongvuthithum 2014; Jouzdani, Sadjadi and Fathian 2013). These applications of hub location modeling in the literate have shed light on network optimization and helped pave the way for more complete methodological framework to study hub network design.

In the past decade, growing attention has also been given to the need for and importance of conducting more empirical studies related to the supply chain for local food products

(Abatekassa and Peterson 2011). In this area, the regional food hub concept has sparked strong interest from a wide array of food system planner, researchers and policy makers. There exists a substantial discussion in the literature regarding the role of regional and local food hubs in improving market access for producers along with their potential for expanding the availability of healthy, fresh food in communities, including underserved communities (Alumur and Kara 2008; Campbel, Ernst and Krishnaoorthy 2002; Feenstra et al. 2011; Jablonskim, Schmit and Kay 2015; O'Hara and Pirog 2013). These data driven models were built with the goal to look for patterns and practices that are consistent enough to be used as viable regional distribution solutions for local food marketing.

Two recent studies by Etemadnia et al. (2013, 2015) formulate hub location problems as mathematical programming models that minimize total network costs which include costs of transporting goods and locating facilities. Computational experiments were conducted to identify the optimal hub numbers and locations in food supply chain systems. While the methodology and analysis contribute to the analysis of optimization of local food systems, these studies impose strong simplifications on the operational level which are directly related to the solution of facility location. First, they use annual production data and neglect the seasonal differentials in production which can affect the hub operational strategies timely and generate heterogeneous costs across marketing seasons. To effectively formulate the system patterns and structure, annual network costs should be derived from the summation of components of seasonal costs. Second, they assumed homogeneous operation costs across hubs with varying handling capacities. Actually, as an economic phenomenon commonly existing in fresh supply chain system, the scale effect of economies plays an essential role to shape the optimal network configuration. The lack of scale effects in the models means generated solutions are likely to deviate from representative experimental conditions that ought to be used to reach the optimality. This study will remove these limitations and attempt to lead to more realistic models that fit into the supply chain context.

## The Scale Effects of Economy

The economies of scale play a fundamental role in the network design (Camargo, Miranda and Luna, 2009; Horner and O'Kelly, 2001). To understand the actual operation cost patterns and identify empirical evidence of scale effect inherent in hub operations, we collect and analyze data on the scope and scale of food hub operations. Based on Economic Census  data (2007) regarding county and state level fresh produce sale values and corresponding operation cost,  we queried data for geotype=02 (state or equivalent)  and 03 (county ) and obtained 108 observations. Four hierarchy quartiles are defined based on sale values of hubs: 0.05-0.20 billion dollars, 0.2-0.5 billion, 0.5-1billion and 1 billion above. For each quartile, relationship between the operation cost of hub (independent variable (X)) and the sale vale of hub (dependent variable) is regressed and shown as follows,

Table 1.  Regression Results for Scale Effect of Economies

| Quartiles | Variables | Coefficients | *t* Stat | *P*-value | *F* | R Square | obs. |
|---|---|---|---|---|---|---|---|
| Quartile 1 | Intercept | 1068.4453 | 0.2350 | 0.8161 | 48.35 | 0.6592 | 27 |
| | X Variable | $0.2169^{**}$ | 6.9537 | 0.0000 | | | |
| Quartile 2 | Intercept | 2424.2869 | 0.1762 | 0.8612 | 22.36 | 0.4114 | 34 |
| | X Variable | $0.2060^{**}$ | 4.7288 | 0.0000 | | | |
| Quartile 3 | Intercept | 11272.4243 | 0.3470 | 0.7316 | 17.33 | 0.4193 | 26 |
| | X Variable | $0.1801^{**}$ | 4.1631 | 0.0003 | | | |
| Quartile 4 | Intercept | 25760.7852 | 1.0405 | 0.3112 | 1191.31 | 0.9843 | 21 |
| | X Variable | $0.1763^{**}$ | 34.5154 | 0.0000 | | | |

Note: $^{**}$significant at 5% level

The operation costs are broken into two main categories: fixed costs and variable (or marginal) costs. Fixed costs include hub establishing and maintenance, machinery, equipment and so on that are independent of volume of products handled. They remain constant in a quartile. Variable costs include wages, utilities and other sources used in handling products. In regressions, the fixed costs are indicated by the coefficients of intercepts and the variable costs are indicated by the coefficients of variables in regressions. Regression results show fixed costs(intercept) increase with the scale of hub, and on the contrary, the marginal costs decrease with the scale of hub, i.e., the more the products handled, the less the marginal cost for one unit increase in the volume handled.  The operation cost for per dollar sale value handled is \$0.2169, \$0.2060,

$0.1801 and $0.1763 from lower quartile 1 (with smaller sale value) to upper quartile 4 (with larger sale value). In this way, handling cost of a shipper hub is based on the amount of product flow carried by the hub and is endogenously responsive to flow by rewarding the shipper for greater volumes shipped. Under cost minimization principle, the number and scale of facilities to be established typically becomes an endogenous decision. Based on these regressed results, this study will embody scale effect of economies into models and demonstrate how the scale effect leads to differing spatial network structure.

## The Problem Statement and Methodology

The objective of this study is to identify and profile supplier roles of fresh produce value chains and collect and analyze data on the scope and scale of supplier operations in order to more clearly understand their potential role and impact in the U.S. food system. For doing this, a model is built up in order to identify the optimal number, scale and locations for TSL hubs for fresh produce sourced from growers located in multiple U. S. counties. In this study, the hub optimization problem is mathematically formulated as a mixed integer linear programming model with an objective to minimize the total costs associated with product assembly and hub operations. The mathematical optimization problem is subject to a number of constraints that ensure observed market outcomes, such as total production by region and average per unit supplier and shipping costs meet observed statistics. The following notations are introduced for formulating the mathematical models,

$I$={1,2,3,4} denotes four marketing seasons in a year;

$F$={1,2,3…,$f$} denotes a set of production locations;

$S$={1,2,3,…,$s$} denotes a set of hub candidate locations;

$c \in C$ denotes capacity level of assembly hubs; each capacity level has an interval span;

$x_{i,f,s}$ denotes quantity shipped from production location f to hub location $f$ in marketing season $i$;

$p_{i,f}$ denotes production at production location $f$ in marketing season $i$;

$z_{c,s}$ denotes integer variable: $z_{c,s} = 1$ if location s is a hub with capacity c, and 0 otherwise;

$d_{f,s}$ denotes distance between production location *f* to hub location *s* (impedance miles) ;

*t* denotes fixed transportation cost ($ per thousand pound impedance mile);

*DT* denotes distance threshold between production locations and hub locations (impedance miles);

$U^1_{c,s}$ denotes maximum capacity of *c* level hub in location *s*;

$V^1_{c,s}$ denotes minimum capacity of *c* level hub in location *s*;

$U^2_{c,s}$ denotes maximum quantity of products handled at *c* level hub in location *s* during all marketing seasons;

$V^2_{c,s}$ denotes minimum quantity of products handled at *c* level hub in location *s* during all marketing seasons;

$TC_s$ denotes total annual hub s operation costs

$TC = \sum_s TC_s$ denotes annual hub operation costs nationwide

A TSL hub facility of capacity *c* is costly to build and maintain. The annual opportunity costs of equity financing, interest costs of debt financing, and replacement costs of physical and economic depreciation are all born by ownership regardless of its use in producing TSL services; denote these as fixed costs called setup and maintenance costs, $h_c^0$ . In addition, for each unit of produce handled up to the capacity to which hub facility is built, a per unit handling cost is incurred; denote these as marginal costs, $h_c^1$. For any region *s* where a TSL produce hub facility of capacity *c* is located, total handling charges from assembling commodities for distribution are:

1) $TC_s = h^0_{c,s} + h^1_{c.s} \cdot Q_s$

where $Q_{,s} = \sum_{i\in I} \sum_{f\in F} x_{i,f,s}$ , and $x_{i,f,s}$ denotes the quantity of produce grown in production node location *f* and shipped to hub *s* in marketing season *i*. It is assumed that produce hub investors will choose form a finite number, *C*, of possible hub capacities to build a TSL produce hub. Each hub size has a maximum capacity constraint, $h_c^{max}$, which determines the maximum quantity of production that can be transported to the hub across four marketing seasons.

Produce assembly at hub location s involves the shipment from surrounding production and import regions (or nodes) to the region s hub via a domestic freight network that connects all nodes and hubs. It is assumed all shipments are transferred through land transportation by using trucks, and the transportation cost is a function of travel distance (impedance miles). This cost may be paid by the grower/importer or the supplier, but either way it is reflected in the fob supplier price[1],

2) $$\sum_{i\in I}\sum_{f\in F}\sum_{s\in S}\{t\cdot d_{f,s}\cdot x_{i,f,s}\},$$

where $t\cdot d_{f,s}$ is the unit transportation cost for shipments between growing/import node $f$ and hub location *s*.

Transportation costs between node pairs are usually defined to be proportional to the distance between nodes. However, transportation costs are also subjected to some constraints, e.g. road condition, speed reduced on evening commute, traffic jams and congestion which influence the travel time and speed (Novaco, Stokols and Milanesi 1990). Our understanding of traffic congestion involves the concept of impedance. To evaluate the transportation costs between production nodes and hubs, each link in the network is assigned an impedance value other than actual mileage. Impedance represents a measure of the amount of resistance, or cost, required to traverse a path in a network, or to move from one element in the network to another. High impedance values indicate more resistance to movement. For this study, version 3 national multi-modal impedance network database is used for this purpose (http://cta.ornl.gov/transnet/SkimTree.htm).

For a national fresh produce transportation and supplier logistics system, optimal TSL hub scales and locations are determined by minimizing total costs of TSL hub operations plus shipping costs of moving all domestically grown fresh produce to a TSL hub. The objective function and system constraints of a model to solve this problem are given in equations 3 to 9:

**Minimize**

3) $$TC = \sum_{c\in C}\sum_{s\in S}\{(h^0_{c,s} + h^1_{c,s}\cdot Q_s)\cdot Z_{c,s}\} + \sum_{i\in I}\sum_{f\in F}\sum_{s\in S}\{x_{i,f,s}\cdot d_{f,s}\cdot t\}$$

Subject to:

4) $$\sum_{s\in S} x_{i,f,s} = p_{i,f} \qquad \forall\, i, f$$

---

[1] Free-on-board (fob) supplier price is the unit cost that buyers of fresh produce from hub *o* are charged assuming the buyer pays (or is charged separately) all freight costs to ship their purchase to the buyers location.

5) $Z_{c,s} \in \{1,0\}$ $\forall\, c, s$

6) $\sum_{c \in C} Z_{c,s} \leq 1$ $\forall\, s$

7) $\sum_{f \in F} x_{i,f,s} \leq \sum_{c} Z_{c,s} U_{c,s}^{1}$ $\forall i, s$

8) $\sum_{f \in F} x_{i,f,s} \geq \sum_{c} Z_{c,s} V_{c,s}^{1}$ $\forall\, i, c, s$

9) $\sum_{i \in I} \sum_{f \in F}\ x_{i,f,s} \leq Z_{c,s} U_{c,s}^{2}$ $\forall\, c, s$

10) $\sum_{i \in I} \sum_{f \in F}\ x_{i,f,s} \geq Z_{c,s} V_{c,s}^{2}$ $\forall\, c, s$

11) $x_{i,f,s} \geq 0$ $\forall\, i, f, s$

The objective function (3) is to minimize the total cost. Equation (4) ensures that in each marketing season the total quantity transported from production region *f* to all hubs *S* are equal to total quantity produced in region *f* in the season. That is, all products must be assembled into hubs across marketing seasons. Equation (5) provides the binary condition—build/non-build, while equation (6) guarantees that the total number of hubs built in hub location *s* is not more than 1. Equations (7) and (8) ensure that products will not be transported to any hub location *s* unless a hub is installed and there is sufficient capacity to handle all products transported to hub s in marketing season *i*. Equations (9) and (10) define the quantity handling threshold for a hub to enjoy a certain level of scale effect. Equation (11) ensures that shipments only flow from farms to hubs and not vice versa.

## Data

The National Agricultural Statistics Service (NASS) of the U.S. Department of Agriculture (USDA) reports State level annual production statistics for 21 different fresh market vegetable crops across 37 States (USDA/NASS, 2009). The 2007 statistics are allocated down to counties in these 37 States based on harvested county acreage for these same 21 fresh market vegetable commodities, using table 30 in the 2007 Census of Agriculture (USDA/NASS, 2009b). For States not covered in the annual NASS reports but that are covered in the Census, harvested acreage data from the Census are multiplied by yield per acre data in nearby States to impute production. To overcome data suppressions in the Census data, a constrained maximum likelihood mathematical programming model was used. The model minimized adjustments to the variance weighted initial estimates of suppressed data to align these initial estimates with the adding up requirements of the hierarchical table 30 data. This hierarchy includes requirements that the sum of harvested crop acreage across all commodities equal the published county-wide total harvested vegetable acreage, and that the sum of harvested acreage for specific crops across all counties in a State equal the published statewide total for each crop. By simultaneously solving a State/County model and a National/State model for the same commodities, the model produces maximum likelihood estimates of all suppressed county harvested acreage statistics. The same approach was used for estimates of fresh market fruit production in U.S. counties, based on annual production statistics of 34 different fresh market fruit and berry crops across 43 States (U.S. Department of Agriculture, 2009c, 2009d), and tables 31 and 32 of the Census data (USDA/NASS, 2009b). Combined production for the subset of these 55 crops produced in each county are converted to a common unit (1,000 lbs) and summed to a single production statistic per county. (County-to-county shipping costs will be based on Oak Ridge National Laboratory multi-modal impedance network data product, and average shipping costs statistics published by AMS. Empirical estimates of fixed and per unit marginal handling costs for each capacity choice are obtained from regression analysis of data in the Economic Census of Wholesale Trade.

Based on 2007 USDA/NASS data, there were 2624 counties in the contiguous U.S. grow vegetable or fruit farms. Annual production statistics will be broken out into seasonal marketing segments to more accurately account for highly variable geographic disposition of annual fresh produce production. Monthly fresh produce import data by county of unlading and export data by county of departure will also be compiled from Census Bureau sources. To anticipate this forthcoming data we partition annual domestic production data into four unequal and arbitrary groupings that comprise 12.5, 37.5, 37.5, and 12.5 percent of annual production respectively. Figure 1 show the continental U.S. fresh produce production map across 4 seasons. The production is not evenly distributed across seasons. Main production areas are located in the West Coast of the U.S. Among them, California has the highest production levels. The second highest production state is Florida in the East Coast. The Northeast also enjoys a high production level. Comparatively, production in some Rocky Mountain States and Plain States are at very low levels.

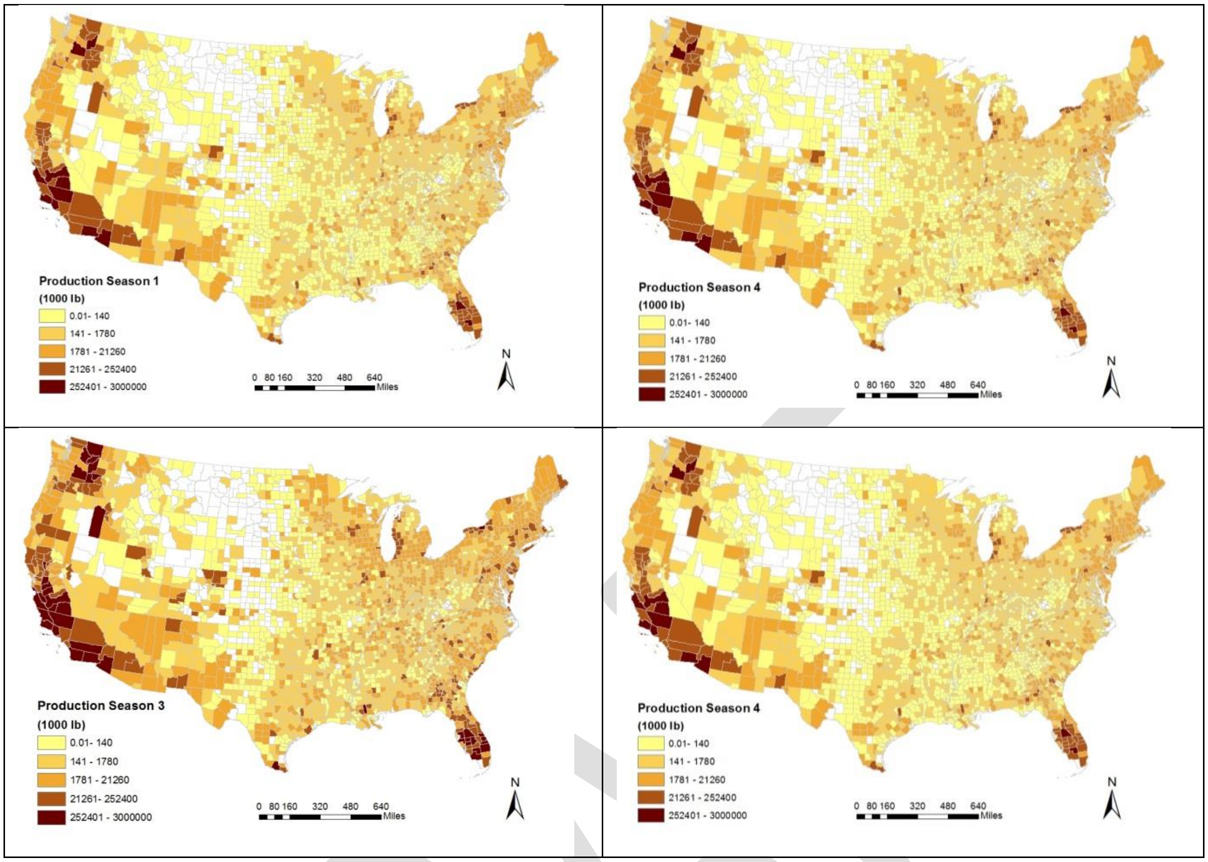

Figure 1: Distribution of Fresh Produce Production in the U.S. across Seasons

## Experimental Setting, Results and Analysis

The TSL produce assembly hub optimization problem is formulated as a mixed integer programming model with an objective to minimize the total costs associated with product assembly and hub operations. Seasonal fresh produce data across U.S. counties are applied to the model.

The models set different costs for establishing and maintaining hubs with different capacities and set different unit costs for handling products in those hubs. The capacity of a hub defines the maximum quantity that is allowed to be handled in the hub in either of four marketing seasons. Hubs with larger capacity can handle larger quantity of products than hubs with smaller capacity during marketing seasons. It is assumed that a hub must handle a certain level of product quantity for achieving a certain level of scale effect. Four different thresholds for quantity of products handled at hubs during four marketing seasons are defined and named as Quartiles 1-4 (unit: thousand pounds),

Quartile 1: 100,000 and below

Quartile 2: 100,001-200,000

Quartile 3: 200,001-2,000,000

Quartile 4: 2,000,000 and above

Regression analysis reported in table 1 will be incorporated into forthcoming model simulations. Here they inform alternative hypothetical hub cost parameters. Thus, annually attributed hub fixed costs ($h^0$) for handling quantity defined in four quartiles are \$80,000, \$120,000, \$240,000 and \$400,000 respectively for models allowing for varying scales of hubs across quartiles. Following the regressed pattern of operation cost in the previous section, a hub handling higher quantity of products enjoys lower variable handling costs. The marginal handling costs ($h^1$) used in this model are calibrated as a ratio of about 10% of their counterparts in regressions (table 1). Our focus is to test the sensitivity of model solution to the scale effect. But we aware that the magnitude of the marginal handling cost significantly influences the model solutions, and these costs will be calibrated in a future iteration of this research.

In order to identify the effect of scale economies on the optimal solutions of the problem, five experimental models are designed. We assume different values of the marginal handling costs for each of five models:

Model 1: marginal handling cost is \$22 across hubs in four quartiles (for every 1000 lb production handling, the same hereafter). Scale effect is not considered in this model. Total fixed costs for each hub is \$131,800 across quartiles. Such a cost value is averaged from Models 2-5.

Model 2: marginal handling costs are \$22, \$21.5, \$21 and \$20.5 respectively for Quartiles 1-4 hubs, i.e. the marginal cost margin between each two adjacent quartiles is \$0.5. Total fixed costs for each hub are \$80,000, \$120,000, \$240,000, and \$400,000, respectively.

Model 3: marginal handling costs are \$22, \$21, \$20 and \$19 respectively for Quartiles 1-4 hubs, i.e. the marginal cost margin between each two adjacent quartiles is \$1. Total fixed costs for each hub are \$80,000, \$120,000, \$240,000, and \$400,000, respectively.

Model 4: marginal handling costs are \$22, \$20.5, \$19 and \$17.5 respectively for Quartiles 1-4 hubs, i.e. the marginal cost margin between each two adjacent quartiles is \$1.5. Total fixed costs for each hub are \$80,000, \$120,000, \$240,000, and \$400,000, respectively.

Model 5: marginal handling costs are \$ 22, \$20, \$18 and \$16 respectively for Quartiles 1-4 hubs, i.e. the marginal cost margin between each two adjacent quartiles is \$2. Total fixed costs for each hub are \$80,000, \$120,000, \$240,000, and \$400,000, respectively.

Using the model reported in equations (3) to (9) and the set of parameter values and data, model simulations for models 1 to 5 generate data for the hub location problem. The optimization problem is compiled using GAMS and solved by using solver CPLEX. All computational executions were performed on High Performance Computing System. Next we present the results of computational experiments and conduct analysis.

There is a significant change in the hub numbers at each quartile across models due to the scale effect. As shown in Table 2, the hub numbers in quartile 1 is negatively related to the marginal cost margin and, on the contrary, the hub numbers in Quartiles 2, 3 and 4 are positively related to the marginal cost margin. The locations and scales (capacity) of these hubs in each model are shown in Figures 1-5. The hub capacity is presented by the maximum quantity of production handled among the four marketing seasons.

**Table 2. The Number of hubs at Each Quartile across Models**

| | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 |
|---|---|---|---|---|---|
| Quartile 1 | 12 | 18 | 19 | 19 | 21 |
| Quartile 2 | 75 | 73 | 77 | 115 | 122 |
| Quartile 3 | 53 | 75 | 92 | 69 | 63 |
| Quartile 4 | 540 | 422 | 338 | 257 | 196 |
| Total | 680 | 588 | 526 | 460 | 402 |

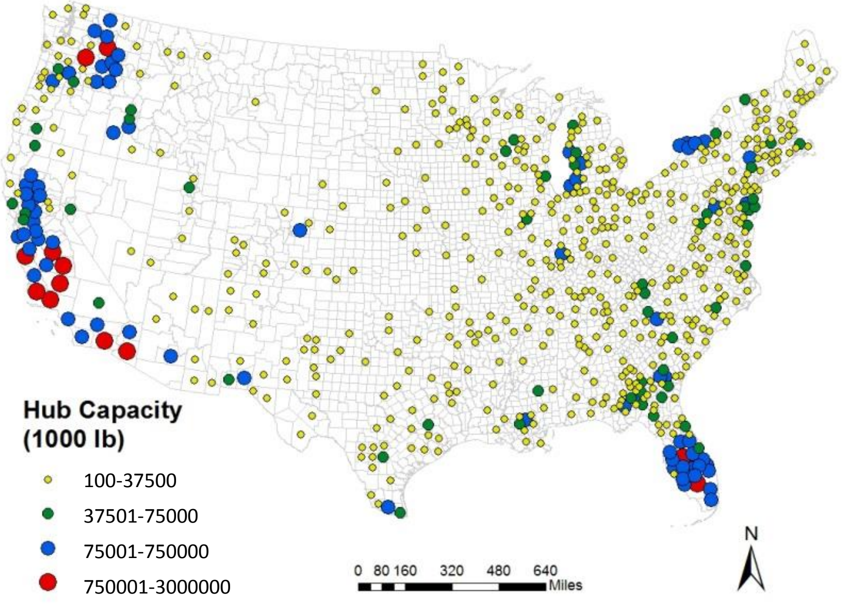

Figure 2. Hub Location for Model 1

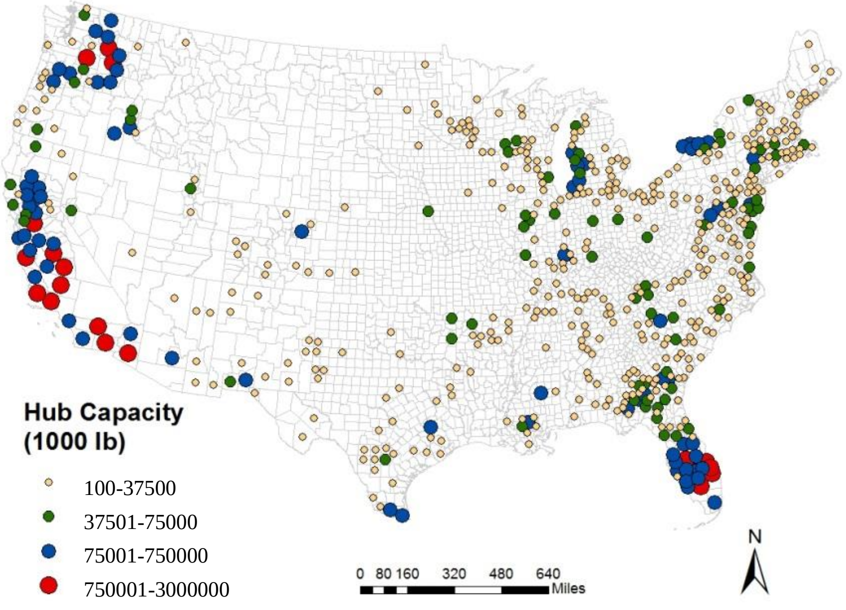

Figure 3. Hub Location for Model 2

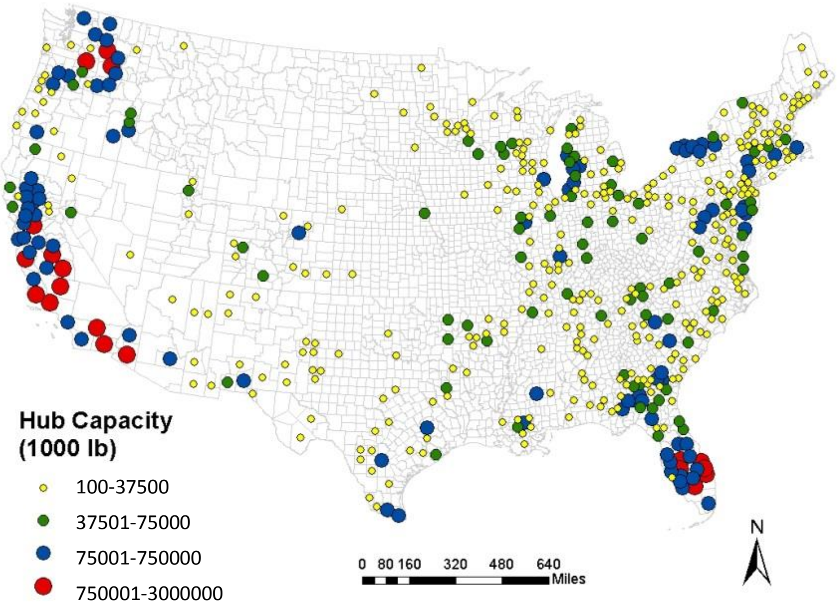

Figure 4. Hub Location for Model 3

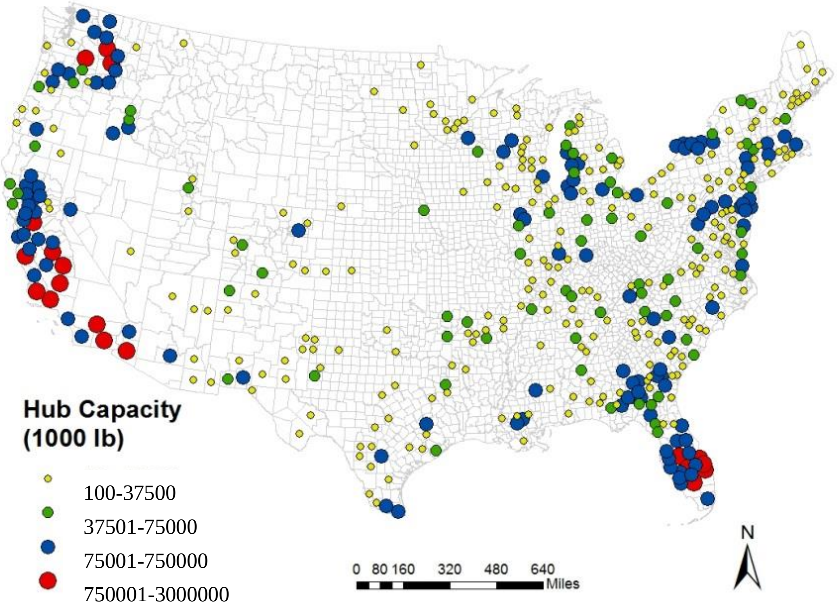

Figure 5. Hub Location for Model 4

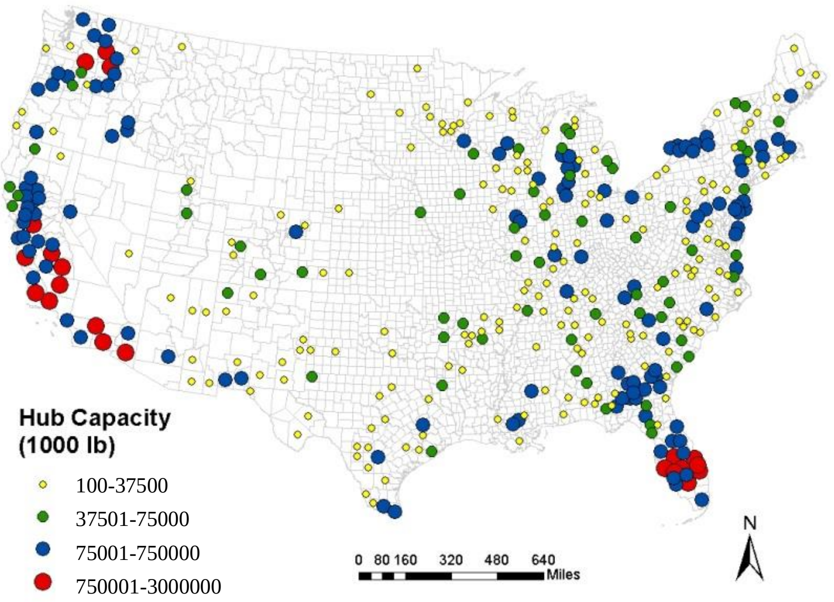

Figure 6. Hub Location for Model 5

When the scale effect is present, hubs handling larger quantity of products enjoy lower marginal handling costs. The merging of smaller hubs into larger hubs facilitates handling cost savings. As shown in Figure 7, the total quantity handled in Quartile 2 hubs decreases with the magnitude of the marginal cost margin and meanwhile the total quantities handled in Quartiles 2-4 increase with marginal cost margin. It is not surprising that the average quantity handled in each quartile continuously decreases as the marginal cost margin widens from Model 2 to Model 5 (Figure 8). In this study, there are minimum and maximum capacity constraints set for each quartile. When the marginal cost margin between quartiles increases, the incentives to establish more hubs with larger capacity become stronger. To harvest the benefit of economies of scale, distribution of products among hubs becomes strategic in order to reduce the handling costs. Figures 9-12 show the quantity handled at each individual hub in each quartile across models (in ascendant order). As Figures 10-12 show, some hubs in Quartiles 2-4 just meet the minimum threshold to enter these quartiles. As a result, the number of upper quartile hubs (with higher capacity) in higher marginal cost margin models, e.g. Models 4 and 5, is more than that in lower marginal cost margin models, e.g. Models 2 and 3. Comparatively, the total quantity handled in these upper quartile hubs in the higher margin model increases at a lower rate than the increase in the number of upper quartile hubs in the same model. As a result, the average quantity handled in upper quartile hubs drops down when the marginal handling cost margin increases.

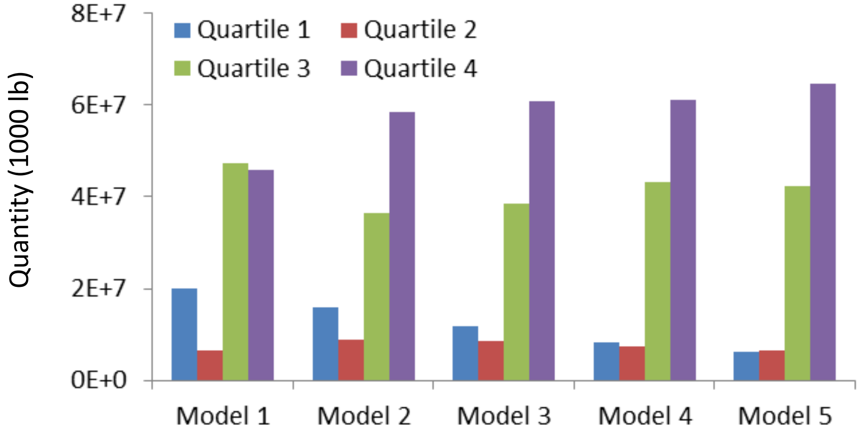

Figure 7. Total Quantity Handled at Each Quartile across Models

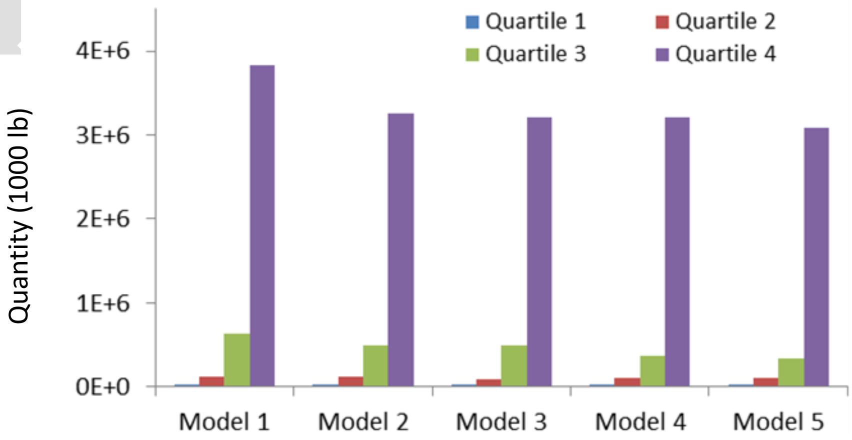

Figure 8. Average Quantity Handled at Each Quartiles across Models

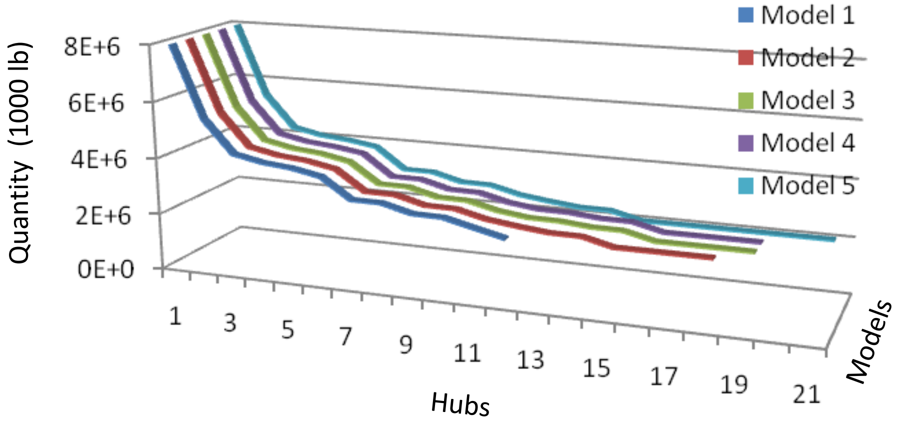

Figure 9. Quantity Handled at Quartile 1 Hubs across Models

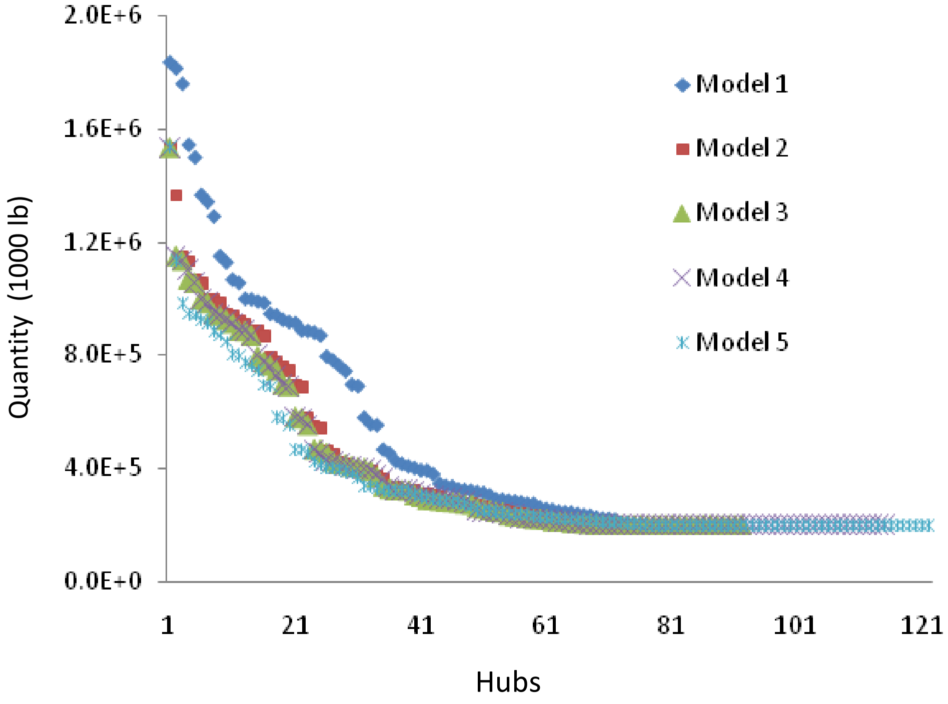

Figure 10. Quantity Handled at Quartile 2 Hubs across Models

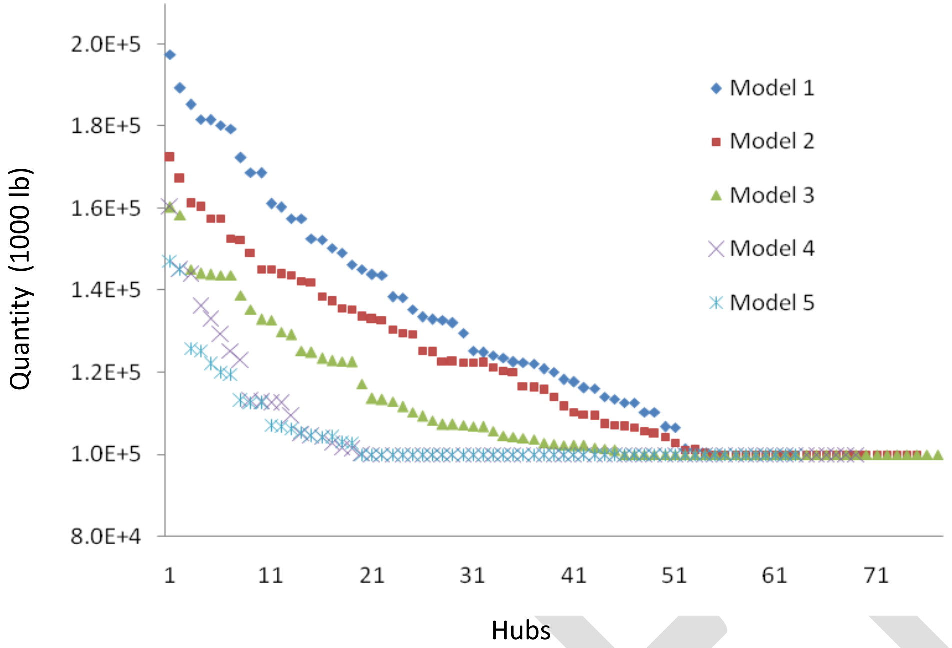

Figure 11. Quantity Handled at Quartile 3 Hubs across Models

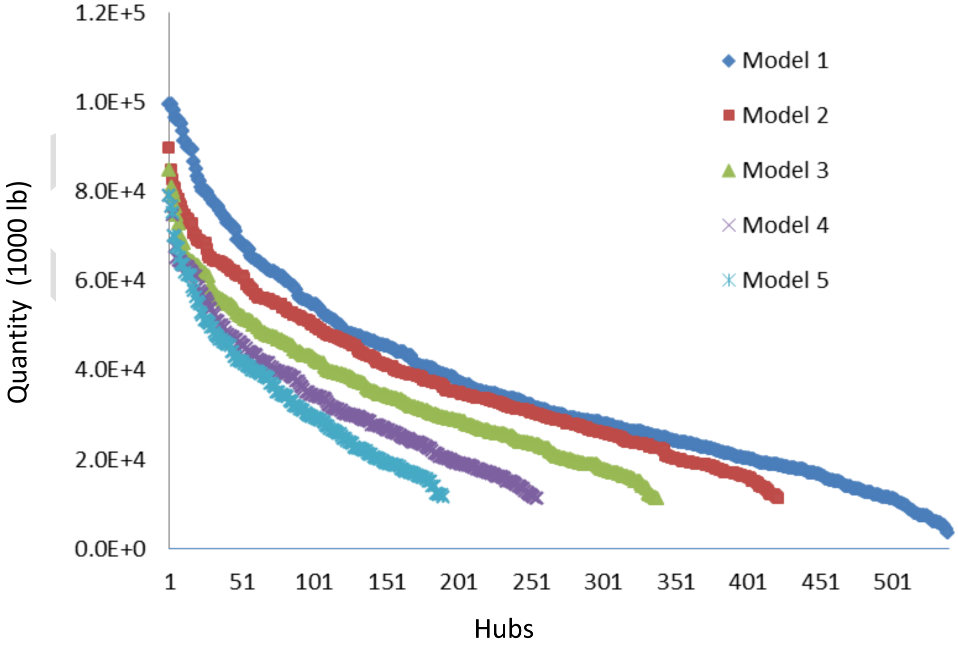

Figure 12. Quantity Handled at Quartile 4 Hubs across

The marginal cost between quartiles provided incentives for transporting products longer in order to build upper quartile hubs which take advantage of the scale effect. Just as shown in Table 3, the transportation costs increase with the marginal cost margin between quartiles. There is a tradeoff between increased transportation costs and handling costs saved from scale effect. Our model is designed to catch the optimal balance.

Table 3. Relative Costs across Models

| | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 |
|---|---:|---:|---:|---:|---:|
| Setup & Maintenance | 88,672,000 | 67,480,000 | 64,160,000 | 64,040,000 | 60,920,000 |
| Transportation | 792,026,238 | 799,020,356 | 810,559,748 | 822,420,675 | 840,224,574 |
| Handling | 2,640,038,627 | 2,511,045,387 | 2,371,308,084 | 2,224,592,197 | 2,069,377,271 |
| Total ($) | 3520736866 | 3,377,545,744 | 3,246,027,832 | 3,111,052,873 | 2,970,521,845 |

The optimal solutions of models 1 to 5 are sensitive to the change in marginal handling costs for hub quartiles. If the marginal cost margin between quartiles changes, a new solution is required to meet a new optimum. Table 4 demonstrates how the optimal solution improves the cost efficiency of each model. Cells in the 5×5 matrix in the table indicate the total costs generated in a situation in which the optimal solutions for models listed in the column are applied to models listed in the row. The total cost resulting from each unique optimal solution for each specific model is shown in the main diagonal of the matrix. Obviously, an optimal solution for a model cannot be optimal for the other, even if it was generated in a condition similar to the other.

Table 4. Comparative Cost Analysis Matrix

| | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 |
|---|---|---|---|---|---|
| Model 1 | **3,520,736,866** | 3,555,734,184 | 3,564,188,776 | 3,573,442,756 | 3,590,683,171 |
| Model 2 | 3,384,742,431 | **3,377,545,744** | 3,380,393,104 | 3,388,016,795 | 3,398,517,239 |
| Model 3 | 3,265,059,996 | 3,248,552,503 | **3,246,027,832** | 3,249,534,834 | 3,255,852,108 |
| Model 4 | 3,145,377,561 | 3,119,559,263 | 3,111,662,560 | **3,111,052,873** | 3,113,186,976 |
| Model 5 | 3,025,695,126 | 2,990,566,023 | 2,977,297,288 | 2,972,570,912 | **2,970,521,845** |

## Conclusions

Facility location is a well-established research area within operation research. The application of hub location models has long been under discussion in regional and local food system studies due to its potential contribution to the sustainability of food supply chains. This paper provides a replicable empirical framework to conduct impact assessments for regional and local food systems.

We explore the idea of endogenous hub location on the fresh produce value chain. To overcome the limitation in the literature, we integrate scale effect of economies into models. By collecting detailed expenditure and sales information from county statistics in the U.S., we identify the pattern of scale effect inherent in hub operations and apply an operation cost function that rewards economies of scale on quantities handled in different hierarchy quartiles. The hub optimization problem was mathematically formulated as a mixed integer linear programming model with an objective to minimize the total costs associated with produce assembly and hub operations. We designed several experiments to assess the scale effect of economies in the network and visualized the results. Our results provide strong evidence that scale effects have significant impact on hub location solutions. Under different marginal cost assumptions, produce transportation is re-routed to take advantage of cost saving, and thus the structure of the network is reshaped. When the marginal handling cost margin becomes wide between different hierarchy of hubs, more hubs with large capacity emerges while the number of small capacity hubs diminishes.

Given the current policy environment, the research is timely and can provide valuable input into assessing the role and potential impacts of new regional food hub infrastructure investments, as well as the costs and potential market outcomes of such investments. Such information is currently lacking and is needed to help inform decisions of the various stakeholders interested in regional food hub infrastructure investment.

Our model has several limitations that suggest topics for future investigation. Addressing scale effect of economies by explicitly considering it as costs on the objective function yields a more realistic modeling approach. But specification of suitable cost function allowing for scale effect is not an easy task. In this study, empirical estimates of fixed and per unit marginal handling costs for each capacity choice are informed by regression analysis of retrieved data from limited observations and are based on rough classification. To really advance the modeling of these sorts of economic issues, more investigation should be conducted to better understand the pattern of scale effect of economies and its influence on hub operation costs.

Our model only accounts for the early portion of fresh produce of supply chain, i.e. from the point of origin to the point of assembly. Points of distribution and points of consumption are not taken into consideration in the model. In this sense, the model is still quite abstracted from the actual U.S. fresh produce value chain system and the complexity within it is reduced when compared to reality. Incorporating all these components in models helps generate hub location solutions that better mirror the actual set of decisions within the supply chain systems. Furthermore, based on the nature of this study, the extension of seasonal production data to monthly production data facilitates catching the cost dynamic of operation cycle and helps generate solutions more compatible with actual situation of fresh produce supply chain systems. Overcoming these limitations can make models cope with realistic hub location settings.

## References

Abatekassa, G., and H.C. Peterson. 2011. “Market Access for Local Food through the Conventional Food Supply Chain.” *International Food and Agribusiness Management Review* 14(1):63-82.

Alumur, S., and B.Y. Kara. 2008. “Network Hub Location Problems: The State of the Art.” *European Journal of Operational Research* 190: 1–21.

Aros-Vera, F., V. Marianov and J. E. Mitchell. 2013. “p-Hub approach for the Optimal Park-and-Ride Facility Location Problem.” *European Journal of Operational Research*, **226**(2): 277- 285.

Barham, J., D. Tropp, K. Enterline, J. Farbman, J. Fisk, and S, Kiraly. 2012. *Regional Food Hub Resource Guide*. Washington, DC: U.S. Department of Agriculture, Agricultural Marketing Service. April.

Boys, K.A., and D.W. Hughes. 2013. “A Regional Economics-Based Research Agenda for Local Food Systems.” Journal of Agriculture, Food Systems, and Community Development 3(4):145–150.

Camargo, R.S., and G. Miranda. 2012. “Addressing Congestion On Single Allocation Hub-and-Spoke Networks.” *PesquisaOperacional* 32(3):465-496.

Camargo, R.S., G. Miranda, and H.P.L. Luna. 2009. “Benders Decomposition for Hub Location Problems with Economies of Scale.” *Transportation Science* 43(1):86–97.

Campbell, J.F., A.T. Ernst, and M. Krishnamoorthy. 2002. “Hub Location Problems.” In Z. Drezner and H.W. Hamacher, 1st ed. *Facility Location: Applications and Theory*. Springer, pp.373–407.

Coleman-Jensen, A., M. Nord, and A. Singh. 2013. *Household Food Security in the United States in 2012*. Washington D.C., U.S. Department of Agriculture, Economic Research Service, Economic Research. Rep. 155, September.

Cook, R. 2011. “Fundamental Forces Affecting U.S. Fresh Produce Growers and Marketers.” *Choices* 26(4): 4th Quarter.

Dantrakul, S., C. Likasiri, and R. Pongvuthithum. 2014. “Applied p-median and p-center Algorithms for facility Location Problems.” *Expert Systems with Applications* 41(8): 3596-3604.

Etemadnia, H., S.J. Goetza, P. Canning, and M.S. Tavallali. 2015. “Optimal Wholesale Facilities Location within the Fruit and Vegetables Supply Chain With Bimodal Transportation Options: An LP-MIP Heuristic Approach.” *European Journal of Operational Research* 244(2):648-661.

Etemadnia, H., A.Hassan, S. Goetz, and K. Abdelghany. 2013. “Wholesale Hub Locations In Food Supply Chain Systems.” *Journal of Transportation Research Board*. 2379:80-89..

Feenstra, G., P. Allen, S. Hardesty, J. Ohmart, and J. Perez. 2011. “Using a Supply Chain Analysis to Assess the Sustainability of Farm-to-Institution Programs.” *Journal of Agriculture, Food Systems, and Community Development* 1(4):69-85.

Horner, M.W., and M.E. O’kelly. 2001. “Embedding Economies of Scale Concepts for Hub Network Design.” *Journal of Transport Geography* 9(4):255-265.

Jablonski, B.B.R., J. Perez-Burgos, and M.I. Gómez. 2011. “Food Value Chain Development in Central New York: CNY Bounty.” *Journal of Agriculture, Food Systems, and Community Development* 1(4):129–141.

Jablonski, B.B.R., T.M. Schmit, and D. Kay. 2015. “Assessing the Economic Impacts of Food Hubs to Regional Economies: a framework including opportunity cost.” Working Paper, Charles H. Dyson School of Applied Economics and Management, Cornell University, Ithaca, New York, USA.

Jouzdani, J., S.J. Sadjadi, and M. Fathian. 2013. "Dynamic Dairy Facility Location and Supply Chain Planning under Traffic Congestion and Demand Uncertainty: A Case Study of Tehran." *Applied Mathematical Modelling* 37(18-19):8467–8483.

King, R., M.S. Hand, G. DiGiacomo, K. Clancy, M.I. Gómez, S.D. Hardesty, L. Lev, and E.W. McLaughlin. 2010. Comparing the Structure, Size, and Performance of Local and Mainstream Food Supply Chains. Washington, D.C.: U.S. Department of Agriculture, Economic Research Service, Economic Research, Rep. 99, June.

Low, S.A., and S. Vogel. 2011. *Direct and Intermediated Marketing of Local Foods in the United States*. Washington, D.C.: U.S. Department of Agriculture, Economic Research Service, Economic Research, Rep. 128, November.

Martinez, S., M. Hand, M.D. Pra, S. Pollack, K. Ralston, T. Smith, S. Vogel, S. Clark, L. Lohr, S. Low and C. Newman. 2010. *Local Food Systems: Concepts, Impacts, and Issues*. Washington DC: U.S. Department of Agriculture, Economic Research Service, Economic Research, Rep. 97, May.

Nicholson, C.F., M.I. Gómez, and O.H. Gao. 2011. "The Costs of Increased Localization for a Multiple-Product Food Supply Chain: Dairy in the United States." *Food Policy* 36(2):300–310.

Novaco, R.W., D. Stokols, and L. Milanesi. 1990. "Objective and Subjective Dimensions of Travel Impedance as Determinants Of Commuting Stress." *American Journal of Community Psychology* 18(2):231-257.

Oak Ridge National Laboratory, Center for Transportation Analysis website http://cta.ornl.gov/transnet/SkimTree.htm

O'Hara, J., and R. Pirog. 2013. "Economic Impacts of Local Food Systems: Future Research Priorities." Journal of Agriculture, Food Systems, and Community Development 3(4): 35-42.

U.S. Department of Agriculture.2009a. *Vegetables–2008 Summary*,NASS Vg 1-2(09), Washington DC, January.

U.S. Department of Agriculture. 2009b. United States Data (Chapter 2) in 2007 Census of Agriculture, NASS, Washington DC, February.

U.S. Department of Agriculture. 2009c. *Noncitrus Fruits and Nuts-2008 Summary*. NASS Fr Nt 1-3(09). Washington DC, May.

U.S. Department of Agriculture. 2009d. *Citrus Fruits–2009 Summary*. NASS Fr Nt 3-1(09), Washington DC, May.