Using Spatial Analysis for Monitoring Fraud in a Public Delivery Program

Yushim Kim

The Ohio State University

In this article, fraud in public delivery programs is approached using spatial analysis. By examining welfare recipients' consumer behaviors in a spatial framework, it is possible to identify stores that draw an anomalous clientele, such as those who travel long distances and avoid stores in their local area, indicating that these stores may be committing fraud. This article shows how this idea can be operationalized for monitoring fraud to secure public accountability.

Keywords: *store choices; vendor propensity measures; spatial analysis*

Introduction

The basic premise of this article is that, in general, individuals obtain their services from providers that are in close proximity to where they live. Using this premise as the starting point, the article shows that spatial models which go beyond the deterministic use of proximity can help to identify patterns that ensue from these habits. Deviations from such patterns could be indicative of unusual, perhaps illegal, behavior, such as fraud, which is defined as a misrepresentation of asset values (Sutherland, 1940).

Fraud in Food Assistance Programs

The Food Stamp Program and the Special Supplemental Nutrition Program for Women, Infants, and Children (WIC) are public delivery programs implemented by the U.S. Department of Agriculture (USDA). Together, these programs spend over \$35 billion per year (\$31 billion for food stamps and \$5 billion for WIC in fiscal year 2005) to provide food and nutritional supplement for the needy.

Authors' Note: The analysis was conducted when the author was working for the Ohio Women, Infants, and Children (WIC) program. An earlier draft of this paper was presented at the 27th Annual Association for Public Policy Analysis and Management Research Conference in Washington, DC, and was approved for publication without identifying the county and vendor names. It has benefited from discussions with Ohio WIC vendor specialists. Special thanks to Pam Hamilton, Todd Reeves, Robin Ridenour, Rick Wallace, and Nicole Wilson. The author also sincerely thank Anand Desai, Robert Greenbaum, Ningchuan Xiao, Yongwan Chun, and three anonymous reviewers who provided valuable comments and helped significantly improve the article.

For such large programs, even a small amount of fraud adds up to a substantial monetary loss at the aggregate level (Bolton & Hand, 2002). For example, between 1999 and 2002, trafficking (the exchange of public benefits for cash) diverted approximately \$395 million per year from food stamp benefits (Macaluso, 2003), which is a loss of 2.2% of the total costs of \$18 billion each year.

Whereas in the private sector, theft or “shrinkage” is mainly due to customers or employees acting as individuals, fraud in public programs more often requires the active participation of the program participant as well as the service provider, such as vendors. Vendors are private entities that contract with public agencies to deliver public services. In public delivery programs, participants can commit fraud, but the feasibility of fraud is enhanced by complicity on the part of vendors.

Currently, several types of vendor fraud have been identified in the WIC program. Overcharging is the main source of fraud. One study (USDA, 2001) showed that, in three separate investigations of WIC retail vendors, 18.1% were involved in overcharging¹ violations one or more times, including 12.4% that overcharged once, 4.2% that overcharged twice, and 1.5% that overcharged all three times. Other illegal activities of vendors include forcing a participant into unwanted purchases, allowing substitutions for prohibited items, and involvement in trafficking.

Because vendor violations are more prevalent and substantial than participant or employee violations in WIC (General Accounting Office, 1999), the focus of traditional fraud detection methods has been primarily on identifying “high-risk” vendors who have a high probability of committing vendor violations, using criteria such as store characteristics, sales volume, participant complaints, and field investigations (USDA, 2000). Although these criteria have been widely used, they are fairly simplistic and focus on small vendors, large WIC sales volume, or vendors against whom complaints have been received. Field investigations are time-consuming, expensive, and easy to avoid.

More importantly, these criteria are less effective in uncovering the interactions between corrupt vendors and participants. Attention has rarely been given to the underlying mechanisms of fraud in public delivery programs. Thus, the basic mechanisms have rarely been incorporated in designing fraud detection methods. In other words, the traditional detection methods using simple criteria treat fraud as separate individual decisions rather than viewing it as the result of complex interactions.

This article begins with a brief review of the theoretical underpinnings of store choice models and their applications. The policy context and data used to illustrate the conceptual framework are then discussed. Finally, the article shows that using spatial analysis makes it possible to identify stores, on the basis of anomalous behavior, as potentially involved in fraudulent behavior. We conclude with a discussion of the implications and limitations of this approach.

Store Choice

Store Choice Models

The literature on how consumers choose stores is vast. The underlying models can be categorized into three approaches based on normative assumptions, revealed preference, and

direct utility (Craig, Ghosh, & McLafferty, 1984). The early models rely on observations and normative assumptions regarding consumer behavior. For example, the nearest center hypothesis assumes that consumers patronize the nearest store providing the required goods and services. Gravity models are developed as an analogy to Newton's theory of gravitational attraction. The degree of attraction between two stores is based on the size of the stores and the distance between them. The main criticism of this approach is that it does not take consumer behavior and patronage into consideration.

While gravity models use a standard Newtonian attraction to predict consumer store choice, the revealed preference approach uses past behaviors to explain consumer store choice. Huff's spatial interaction model (1964) occupies a unique position between gravity models and the models based on revealed preference. The Huff model (1) incorporates two important principles. First, the patronage area of the consumer is probabilistic rather than deterministic. Second, the probability of a consumer visiting a particular store is a relative measure equal to the ratio of the utility of that store to the sum of utilities of all the stores considered by the consumer. In the model, as in the gravity models, the utility of the store is based on store size (the amount of floor space) and distance from the consumer.

The *Huff spatial interaction model* may be expressed as

$$P_{ij} = \frac{U_{ij}}{\sum_{j=1}^s U_{ij}} = \frac{S_j^\alpha \cdot D_{ij}^\beta}{\sum_{j=1}^s S_j^\alpha \cdot D_{ij}^\beta} \quad (1)$$

where

p_{ij} is the probability of consumer i visiting store j ;

J is the set of competing stores in the region;

U_{ij} is the utility of store j for consumer i ;

S_j is the size of store j ;

D_{ij} is the distance between consumer i and store j ; and

α and β are the parameters.

Empirical studies have found the Huff model to be reasonably accurate in predicting the market share of shopping centers (Craig et al., 1984; Yrigoyen & Otero, 1998). However, this model has been criticized as being simplistic in its explanation of complex consumer store choice behavior.

Extensions of the Huff model have included other store attractiveness attributes, such as service or product quality, image, price, availability of parking (Craig et al., 1984; Drezner & Eiselt, 2002), loyalty (Knox & Denison, 2000), and reputation (Ou, Abratt, & Dion, 2006). Two commonly used extensions are the multiplicative competitive interaction (MCI) models and the multinomial logit (MNL) models. MCI models were introduced by Nakanishi and Cooper (1974, 1982) and Cooper and Nakanishi (1983) to study retail location in a competitive environment (Ghosh & Craig, 1983). MNL models have been used to estimate store choice probabilities using discrete choice data (Gaver, 1980; Hauser, 1980; McFadden, 1980; Suárez, del Bosque, Rodríguez-Poo, & Moral, 2004).

Many of the early applications of the model stayed close to the physical science interpretations, including the Newtonian analogy where the square of distance is the appropriate power function. Larger exponents were introduced to indicate that the friction of distance becomes increasingly important in reducing the expected level of interaction between two locations (Haynes & Fotheringham, 1984). However, a later study of the Huff model found that store size has a greater influence on store choice than distance (Eppli & Shilling, 1996).

Models that incorporate the revealed preference seem to perform better than those based on a normative approach. However, some problems still remain in that consumers reject stores beyond a certain distance, and parameters estimated in the models are context dependent. Direct utility was introduced as an approach to overcome context dependency by calibrating parameters from simulated choice data using hypothetical store descriptions (Craig et al., 1984).

Applications of Store Choice Models

The traditional uses of store choice models also vary. They include not only trading area analyses, as discussed above, but also location allocation decisions (Benoit & Clarke, 1997; Drezner & Hamacher, 2002; Serra & Colomé, 2001) and spatial accessibility analyses (Fortney, Rost, & Warren, 2000; Guagliardo, 2004; Kwan & Weber, 2003). In other words, these store choice models have been frequently used to locate businesses and other social infrastructure amenities.

The main difference between public and private sector applications for location decisions is to be found in the objective function of the models (Revelle, Marks, & Liebman, 1970). The objective in public sector applications is to maximize social utility or to minimize social costs (Marianov & Serra, 2002). On the other hand, the objectives of location decisions in the private sector applications are to minimize the total costs of transportation, or to maximize profits or market share.

For example, public facility location models have focused on building social infrastructure, such as schools, emergency centers, and health clinics, to meet the basic needs of the population (Dear, 1978; Hansen, Peeters, & Thisse, 1983; Rosenberg, 1984). These models help identify, within a set of policy constraints, better location choices for providing program benefits and goods to spatially dispersed populations. Private sector location models have been developed to choose locations in competitive markets for franchises, grocery stores, and convenience stores (Drezner & Eiselt, 2002; Houston & Stanton, 1984; Leszczyc, Sinha, & Timmermans, 2000). Spatial models have been used to secure the spatial accessibility of consumers in the competitive environment.

Applications of these models to a policy context raise a number of issues (Moore, 1983, 2002; Murray, 1983) regarding the suitability of store choice models in contexts where poor people are seeking services. Calibration of such models needs to be context dependent, particularly for policy analysts whose goals are very different from wanting to maximize profit or market share. Although it may seem that calibration of these models is best done with real data, that approach also raises a number of other issues pertaining to the quality of the data² and their availability (Huff, 2003; Tilly, 2006).

In this article, a different use of the models is examined in the context of public service delivery programs. In spite of the limitations of the store choice models, they provide a

simple starting point for examining the differences between the predicted and actual choices of participants selecting vendors that provide public services. After calculating predicted store choice probabilities of all participants for all vendors, the statistics can be summarized and compared with the observed store choices of participants at each vendor. This exercise would capture the general propensity of vendors in terms of the kinds of participants they interact with at a certain point in time. Therefore, a measure can be constructed to show each vendor's propensity based on the spatial interaction with program participants. This measure can be utilized for identifying anomalies which could imply fraud in the delivery of public services.

This exercise serves two purposes. First, it offers a practical method for identifying high-risk vendors, thus providing a resource-effective tool for public accountability and performance management. Second, it helps identify the spatial patterns that may help discriminate among different types of interactions between vendors and participants to provide a more focused subsequent investigation.

One possible explanation of abnormal spatial patterns could be fraud, because decision factors that explain general consumer shopping behavior under normal circumstances failed to predict participant behavior. It can be hypothesized that the possibility of illegal benefit exchanges at a vendor attracts certain participants who are more inclined to participate in fraud.

On the other hand, the marketing literature suggests another explanation, which is that the anomalous pattern may occur due to cultural coherence between stores and consumers (Eckman, Kotsiopoulos, & Bickle, 1997; Ogden, Ogden, & Schau, 2004; Wang, 2004). If a vendor with a certain cultural characteristic, such as country of origin, attracts participants who have the same characteristic, the model that does not include such a factor may also not be able to explain the anomaly. These ideas are examined in the context of Ohio WIC.

Policy Context

Ohio WIC aims to safeguard the health of low-income women, infants, and children up to age 5 who are at nutritional or health risk. The program provides nutritious supplemental foods, nutritional education, and referrals to health care and other social services. In terms of participation, Ohio WIC is the eighth largest program in the United States as of July 2006. Currently, Ohio WIC serves approximately 277,000 participants each month, with a budget of over \$150 million each year (\$110 million for food costs and \$45 million for nutrition services and administrative costs in 2005).³ Ohio WIC has contracts with over 200 local clinics and 1,400 vendors in all 88 Ohio counties.

Each month, participants receive three or four vouchers at local clinics for food benefits. These participants are expected to use their benefits at WIC vendors within a specified period. Each voucher specifies what products and quantities the participant can purchase, as well as the maximum prices that the state will pay for an allowable food. After participants use the vouchers at vendors, WIC vendors submit the vouchers to the state for reimbursement. The state pays the vendor for the voucher after a validation process. The state also monitors the overall flow of all transactions in the WIC system. Each month, approximately a million vouchers are processed by the Ohio WIC program.

Data

Data were compiled from three different sources in Ohio WIC. The WIC certification system provided detailed information on program participants. The vendor management system (VMS) collects demographic information on WIC vendors that contract with the state to provide food to participants. Finally, one month (April 2004) of payment data were retrieved to identify participants who used their vouchers along with the vendors where these benefits were redeemed. In addition to the three data sources, snapshot compliance investigation data were later utilized to overcome the limitations of routine monitoring of payment data.

Given the size of the program and voucher transactions, we limit our analysis for illustrative purposes to an urban Ohio county. A county is an administrative boundary of WIC at the local level. We analyzed 78,236 voucher transactions from 22,874 participants at 78 vendors in the county. It shows that a participant used approximately 3.4 food vouchers that were issued for use during April 2004. Voucher transactions from participants with missing home addresses were excluded because measuring distance between participants' home addresses and vendors' locations is crucial to this study.

Although the unit of analysis is the interaction between vendors and participants, information is summarized by vendor because they are the main focus of monitoring for fraud by the state. There appears to be considerable variety among the vendors licensed by the state to provide food for participants in the WIC program. The average number of checkout lanes per vendor is 9, ranging from 1 to 37. The size of store implicitly reflects the type of business. Whereas small vendors with one or two lanes are family owned, larger vendors belong to national chains. In the month of study, the average number of voucher transactions at each vendor was approximately 1,000, ranging from a minimum of 12 to a maximum of 4,520 transactions.

Vendor Propensity Measures

In an effort to construct a measure of the propensity of a vendor to commit fraud, the Huff model was used with standard parameter values, that is, $\alpha = 1$ for size and $\beta = -2$ for distance. It is assumed that a participant's store choice is mainly based on whether she can save money while shopping and whether the store is close to the participant's home. There is a higher chance of saving money on grocery items, other than WIC foods, in larger stores than in convenience marts. However, distance is believed to be more crucial for women with children at health risk because their mobility may be limited. Hence, the two criteria incorporated in the model are size and distance.

The number of checkout lanes serves as a proxy measure for store size. The vendors' and participants' addresses were geocoded using ArcGIS and a census street file of the county. Euclidean distances (not travel distance) between the vendors' and participants' locations were calculated using Matlab. After two critical factors in the Huff model, store size and distance, were obtained, participants' vendor selection probabilities were calculated. This computation allowed the creation of a probability matrix among all participants and vendors. Hence, both the model's estimated probability and the actual behavior of the participants in their visits to the vendors were identified.

The steps involved in obtaining the propensity measures and spatial patterns are as follows:

- Step 1:* Collect data on the location of vendors and participants and other relevant information regarding vendor and participant characteristics to be included by selecting a relevant store choice model (i.e., the Huff model).
- Step 2:* Compute the store choice probability, which is the theoretical probability (p_{ij}) of participant i shopping at vendor j based on the characteristics.
- Step 3:* Create a p_{ij} matrix for all vendors and participants.
- Step 4:* Group participants' p_{ij} 's by vendor where the participants actually used the vouchers.
- Step 5:* Calculate statistics (minimum [MIN], maximum [MAX], and average [MEAN] values) of p_{ij} 's for each vendor to use as vendor propensity measures.
- Step 6:* Examine the spatial patterns of the vendors.
- Step 7:* Explore the explanations of abnormal spatial patterns.

Propensity measures consist of values of the mean (MEAN) and maximum (MAX) of participants' vendor selection probabilities at a vendor grouped by observed participant behaviors. For any vendor, MEAN is the average propensity that a participant will visit that vendor and MAX is the highest probability of a participant visiting that vendor among alternatives. Ideally, one would expect participants to visit vendors with high MEAN and MAX values if the store choice model provides a reasonable explanation of participants' store choice behavior.

Results

Distribution of Vendor Propensity Measures

The interpretation of vendor propensity measures is straightforward. For example, a MEAN of 0.40 at a hypothetical Vendor X indicates that, on average, Vendor X attracted participants who have approximately a 40% chance of visiting the vendor among alternatives. According to the store choice model, a high MEAN indicates that, on average, the vendor attracts participants who have higher probabilities of visiting the vendor. On the other hand, a low MEAN shows that the vendor attracts participants who have a low chance of visiting the vendor. The other measure at a vendor, MAX, also provides useful information. Vendors with low MAX are losing participants who have a theoretically high chance of visiting the vendor. Therefore, vendors with low MEAN and/or low MAX are potentially attracting participants who would not visit them under normal circumstances. Alternatively, they are losing participants who would normally visit them. This is anomalous behavior from the perspective of standard store choice models.

In the data set consisting of 78 vendors in the county, MEAN was 0.19, ranging from a minimum of 0.01 to a maximum of 0.47. On average, MAX was 0.77, ranging from a minimum of 0.01 to a maximum of 0.99. Figure 1 presents the distribution of vendors by MEAN and MAX. A cluster of vendors with extremely low MEAN and MAX is circled in the lower left corner. For example, a vendor with MEAN and MAX that are close to zero implies that the vendor was not attracting those who were likely to visit the vendor

Figure 1
Scatterplot of Propensity Measures (MEAN and MAX values) for 78 Vendors

(extremely low MAX) and was attracting those who were not likely to visit the vendor (extremely low MEAN). This provides a starting point for further investigation.

Explaining the Anomalies

If a vendor attracts participants by providing illegal benefits and certain participants respond to this offer, decision factors in the model, such as size and distance, may not play a critical role in participants’ vendor selection decisions. On the other hand, if a vendor is known to commit fraud, it is highly likely that some participants may be attracted to the vendor or repelled by the illegal activity. In such situations, the model prediction will be different from actual behavior.

To test this hypothesis, vendors were grouped by their level of risk based on two methods: (a) the state’s routine High-Risk Vendor Identification Index based on monthly payment activities, and (b) actual compliance investigations in the county. Neither of the methods is perfect for grouping the vendors by risk level. The High-Risk Vendor Identification Index is based on practical experience and is used as a routine monitoring

procedure to identify outliers. The compliance investigations serve as a check, and they complement the first method.

First, the 78 vendors were grouped by the High-Risk Vendor Identification Index. The index consists of three indicators to assess WIC sales volume at the vendor. The indicators are (a) vendor redemption ratios based on the fact that not all benefits in food vouchers are redeemed by participants, (b) WIC sales per lane, and (c) food package costs per participant at the vendor. Indicators (b) and (c) are used for identifying outliers compared to their peers in terms of store size and participants' food purchase. The vendor is considered to be at risk when it exceeds a certain threshold for the indicators. When the vendor exceeds the threshold, it is assigned 1; it is assigned 0 otherwise.⁴ The state High-Risk Vendor Identification Index is the sum of the three indicators. Of the 78 vendors, 8 vendors (10%) met two or three indicators at the same time. These 10% of the vendors were categorized as high risk, whereas the remaining 90% were classified as not high risk.

Second, vendors in the county were grouped by the results of compliance investigations that were implemented from April to June 2004. Of the 78 vendors, 51 vendors were selected for compliance investigations due to resource constraints. Three methods were used to select these vendors: (a) those that had been previously identified as high risk by the High-Risk Vendor Identification Index, (b) vendors suggested by specialists who had field experience, and (c) random sampling from those who were not identified by either method. This procedure overestimates the likelihood of a vendor committing fraud and therefore the potential number of vendors who might violate their contract. Among these 51 vendors, 12 vendors (24%) were found to have committed vendor violations in the investigation data. The results showed that 6 high-risk vendors identified using the High-Risk Vendor Identification Index were also found to commit vendor violations in the field investigation. The remaining 6 vendors were those who were not categorized as high risk by the High-Risk Vendor Identification Index but were identified as committing vendor violations in the field investigation.

Figure 2 presents MEAN box plots by risk group in both grouping methods. In the left figure, vendors were grouped by the High-Risk Vendor Identification Index. MEAN was significantly lower (0.05) in the high-risk group compared to the non-high-risk group (0.20). The right figure shows MEAN for the high-risk group (0.14) and non-high-risk group (0.21) when the vendors were grouped by the actual compliance investigation. In both figures, the high-risk vendors attracted those who have a relatively low probability of visiting the vendors compared to the non-high-risk vendors. In other words, the high-risk vendors attracted substantial numbers of participants who were not expected to visit. The basic statistics of MEAN by the two different grouping methods are reported in Table 1.

The field investigations and demographic information were further analyzed to examine whether cultural factors could explain these differences. The county from which these data were obtained has a substantial population of poor immigrants. The clustered stores in Figure 2 were owned by people of the same country of origin as the immigrants. We refer to them here as *specialty vendors*. In terms of interactions, seven of the eight specialty vendors in the county showed low MEAN (0.01~0.03) and low MAX (0.01~0.63). Of the seven high-risk vendors, four vendors with low MEAN were reported to have had overcharge violations, and two vendors had other problems, such as not having price tags for food on their shelves or not returning food vouchers. The remaining vendor had substitution violations. Interestingly, a

Figure 2
Comparison of Vendor Propensity Measures by Risk Groups
Categorized by Two Grouping Methods

Note: The left-hand figure shows the distribution of MEAN when vendors are grouped by the High-Risk Vendor Identification Index. The right-hand figure shows the distribution when vendors are grouped by the results of actual field investigation. The white lines in the box plots show medians. Note that the total number of vendors included is not very large. Therefore, these figures should be interpreted with caution. See Table 1 for statistics on this comparison.

Table 1
Comparison of MEAN by Two Vendor-Grouping Methods for the
Purpose of Evaluating Vendor Propensity Measures

MEAN	The High-Risk Vendor Identification Index (<i>N</i> = 78)			Actual Investigations (<i>N</i> = 51)		
	Mean	Std. Dev.	Freq.	Mean	Std. Dev.	Freq.
High-risk	0.09	0.15	8	0.14	0.16	12
Non-high-risk	0.20	0.11	70	0.21	0.09	39

specialty vendor was reported to commit overcharge violations in the compliance investigation, although it seemed to have relatively high propensity measures (MEAN: 0.23; MAX: 0.80) compared to the other seven high-risk specialty vendors. Therefore, cultural coherence could explain the anomaly, but it does not refute fraud in this context.

An explanation for why the last specialty vendor was different from the others in terms of spatial interactions was also needed. Among the eight specialty vendors, four vendors participated in the program since August 2002, and the other four vendors participated since August 2003. The vendor belonged to the August 2002 cohort. When the historical WIC sales activities were compared (from October 2002 to April 2004), among the August 2002 cohort, the vendor was the only one that was not categorized as a monetary high-risk vendor until October 2003. After October 2003, the vendor started meeting one of the three indicators by violating the vendor redemption ratio. It was also noticed that the location of the vendor geographically appeared out of place compared to other specialty vendors.

Discussion

In this article, the utility of store choice models was tested by constructing vendor propensity measures for monitoring fraud in the context of public delivery programs. This approach provides an opportunity to test whether traditional economic models based on rationality can be utilized in such a context. The abnormal spatial patterns associated with high-risk vendors were also identified. The study shows that these measures can be effective tools for monitoring vendor performance. However, it is more important to realize that the context should be carefully understood to reduce the possibility of alternative explanations for the observed results.

The current practice of fraud prevention and detection treats and focuses on the players in the program separately. For example, current policy is developing by focusing independently on vendor schemes, internal employee schemes, and client schemes. This approach does not consider fraud in public programs as an outcome of interactions among the different entities. Even with all of the obvious limitations of the theoretical model, this research shows that capturing static interactions using a store choice model can be a useful tool for identifying abnormal spatial patterns of fraudulent behavior among vendors and participants.

The State of Louisiana built an excellent geographic information systems (GIS) and business intelligence (BI) framework for discovering fraud (Frontenot, 2004). Louisiana's

fraud detection application shows how the combination of GIS and BI can help detect food stamp fraud. Missouri has followed in Louisiana's footsteps to develop a fraud detection program (Douglas & Matlack, 2005; Perlman, 2005). In these examples, however, GIS maps simply function as a visualization tool to illustrate the analyses of business intelligence. In other words, once the basic data and information were in place, the State of Louisiana began extending its fraud detection capabilities by incorporating GIS into the BI (Environmental Systems Research Institute, 2004).

Here, a different way of utilizing GIS and spatial analysis was proposed for monitoring criminal activity in public delivery programs. GIS can not only present complex information in a visual format, but also be used as a modeling tool to incorporate spatial interaction based on proximity and other attributes in a probabilistic manner for monitoring fraud. This approach using propensity measures addresses the limitation of the Louisiana system and also improves the utility of GIS for policy analysis. The method can be transferred to similar public delivery programs in order to ensure public accountability.

This work can also be naturally extended to the emerging areas of national security and emergency management. For example, Nunn (2005) highlighted the problem of attempting to find something whose basic nature is unknown. He stated that

without knowing what to expect, though, it's hard to know what clues are even relevant. You cannot connect the clues because you don't know the plot, but you can't know the plot unless you figure out the clues. What is missing is an "attack hypothesis" setting out what to expect to find in the data patterns. (p. 10)

We have proposed here an approach to an attack hypothesis using currently available information.

Policy Implications

States need to send the right signals to vendors and participants to make them aware that their activities are monitored periodically. Identifying and reporting even minor violations can serve as a deterrence for potentially more serious violations in the future. Continuous training, education, and communication may prevent vendors from becoming high-risk vendors. Vendor propensity measures can help identify those potential violators using spatial analysis. By identifying would-be violators using this type of tool, the state can intervene before the vendor is fully developed as a violator. However, one should also be careful in interpreting the results of such an analysis outside of the context in which they occur. Finally, states need to invest resources to understand participants' benefit redemption patterns and interactions with vendors because the target group can be very different from the general consumers that most academic literature has been focused on. This effort will pay dividends in the form of highly effective vendor management practices and a program with integrity.

Research Limitations

Although there are advanced models that can be used to explain consumers' store choices, as discussed earlier, this study used the Huff model for two reasons. The first was data availability. Without extra resources and support, it is not easy to design or collect

other variables through the existing information system in practice. The second reason was the purpose of utilizing social science models for policy analysis. The Huff model was used not only because it is a foundation of other extensions but also because it provides sufficient utility for the purpose of monitoring fraud in the given context.

In the Huff model, both decision factors could have a nonlinear nature. As criticized, people may not consider stores beyond a certain distance. In Ohio WIC, the basic unit of operation is at the county level. Participants are informed of all vendors within the administrative boundary. However, there could still be a certain threshold of distance, especially in metropolitan areas with a relatively large number of vendors. Size measure, based on the number of lanes, may also not have a linear influence on participants' store choice as used in the model.

The simplicity and elegance of utility-maximizing rational choice models can be very attractive. However, these models do not fully capture the complexities of vendor-participant interactions (Drezner & Eiselt, 2002). The static approach of such economic models does not leave much room for randomness, adaptive behavior, or learning, which are critical aspects of human decision making (Fawcett & Provost, 1997; March, 1978; Simon, 1996; Wilhelm, 2004). These analyses are beyond the scope of the current focus. Further studies are needed to consider such complexity of human behavior.

Notes

1. *Overcharging* means a charge by a WIC-authorized store to the program through redemption of a WIC voucher for an allowable food in excess of the store's shelf price for that food or in excess of the price charged a non-WIC participant for that food.

2. The quality of data implies not only the messiness of data but also the fundamental nature of data in social science. Tilley (2006) argued that people negotiated reasons when they were asked to provide the reason of certain behavior. Therefore, it may be necessary to ask a fundamental question on the nature of data collected and analyzed by social scientists asking people for reasons behind their choices.

3. Data were retrieved from the USDA website (November 7, 2006).

4. The vendor redemption ratio of over 90% is considered suspect. The vendor redemption ratio is based on two facts. Not all food benefits in food vouchers are redeemed by participants. The state overissues approximately 10% of the total price of products in food vouchers due to price variance or fluctuations in different regions or over time. The other two measures are based on absolute values in the given context (sales per lane $\geq$ \$5,000 and food package costs per participant $\geq$ \$50). Although the cutoffs can be transformed to percentile, the state's cutoffs were used in this article in order to be based on consistent criteria and results that were used for explaining the anomaly in the state.

References

- Benoit, D., & Clarke, G. P. (1997). Assessing GIS for retail location planning. *Journal of Retailing and Consumer Services*, 4(4), 239-258.
- Bolton, R. J., & Hand, D. J. (2002). Statistical fraud detection: A review. *Statistical Science*, 17(3), 235-255.
- Cooper, L. G., & Nakanishi, M. (1983). Standardizing variables in multiplicative choice models. *Journal of Consumer Research*, 10(1), 96-108.
- Craig, C. S., Ghosh, A., & McLafferty, S. (1984). Models of the retail location process: A review. *Journal of Retailing*, 60(1), 5-36.

- Dear, M. (1978). Planning for mental health care: A reconsideration of public facility location theory. *International Regional Science Review*, 3(2), 93–111.
- Douglas, B., & Matlack, E. (2005, September). The road to GIS at Mississippi's Department of Human Services. *Direction Magazine*. Retrieved February 26, 2007, from <http://www.directionsmag.com>.
- Drezner, T., & Eiselt, H. A. (2002). Consumers in competitive location models. In Z. Drezner, & H. W. Hamacher (Eds.), *Facility location: Application and theory* (pp. 151–178). Berlin: Springer.
- Drezner, Z., & Hamacher, H. W. (Eds.). (2002). *Facility location: Application and theory*. Berlin: Springer.
- Eckman, M., Kotsiopoulos, A., & Bickle, M. C. (1997). Store patronage behavior of Hispanic versus non-Hispanic consumers: Comparative analyses of demographics, psychographics, store attributes, and information sources. *Hispanic Journal of Behavioral Sciences*, 19(1), 69–83.
- Eppli, M. J., & Shilling, J. D. (1996). How critical is a good location to a regional shopping center? *The Journal of Real Estate Research*, 12(3), 459–468.
- Environmental Systems Research Institute. (2004, July–September). Taking a bit out of illegal food stamp traffic. *ArcUser*, from <http://www.esri.com/news/arcuser/0704/foodstampfraud.html>.
- Fawcett, T., & Provost, F. (1997). Adaptive fraud detection. *Data Mining and Knowledge Discovery*, 1, 291–316.
- Fortney, J., Rost, K., & Warren, J. (2000). Comparing alternative methods of measuring geographic access to health service. *Health Services & Outcomes Research Methodology*, 1(2), 173–184.
- Frontenot, D. (2004, July). The State of Louisiana takes a bite out of illegal food stamp traffic. *Information Builders*. Retrieved February 26, 2007, from <http://www.dmreview.com>.
- General Accounting Office. (1999). *Food assistance: Efforts to control fraud and abuse in the WIC program can be strengthened. Report to Congressional Committee* (GAO/RCED-99-224). Washington, DC: Author.
- Gaver, K. M. (1980). Comments on “econometric models for probabilistic choice among products.” *The Journal of Business*, 53(3), S35–S36.
- Ghosh, A., & Craig, S. (1983). Formulating retail location strategies in a changing environment. *Journal of Marketing*, 47(3), 56–68.
- Guagliardo, M. F. (2004). Spatial accessibility of primary care: Concepts, methods, and challenges. *International Journal of Health Geographics*, 3(3), from <http://www.ijhealthgeographics.com/content/3/1/3>.
- Hansen, P., Peeters, D., & Thisse, J. F. (1983). Public facility location models: A selective survey. In J. F. Thisse and H. G. Zoller (Eds.), *Locational analysis of public facilities* (pp. 223–255). Amsterdam: North-Holland Publishing.
- Hauser, J. (1980). Comments on “econometric models for probabilistic choice among products.” *The Journal of Business*, 53(3), S31–S34.
- Haynes, K. E., & Fotheringham, A. S. (1984). *Gravity and spatial interaction models*. Beverly Hills, CA: Sage.
- Houston, F. S., & Stanton, J. (1984). Evaluating retail trade areas for convenience stores. *Journal of Retailing*, 60(1), 124–136.
- Huff, D. L. (1964). Defining and estimating a trading area. *Journal of Marketing*, 28, 34–38.
- Huff, D. L. (2003, October–December). Parameter estimation in the Huff model. *ArcUser*, 34–36.
- Knox, S. D., & Denison, T. J. (2000). Store loyalty—its impact on retail revenue: An empirical study of purchasing behavior in the UK. *Journal of Retailing and Consumer Services*, 7, 33–45.
- Kwan, M., & Weber, J. (2003). Individual accessibility revisited: Implications for geographical analysis in the twenty-first century. *Geographical Analysis*, 35(4), 341–353.
- Leszczyc, P. T. L. P., Sinha, A., & Timmermans, H. J. P. (2000). Consumer store choice dynamics: An analysis of the competitive market structure for grocery stores. *Journal of Retailing*, 76(3), 323–345.
- Macaluso, T. (2003). *The extent of trafficking in the food stamp program: 1999–2002*. Alexandria, VA: Food and Nutrition Service.
- March, J. G. (1978). Bounded rationality, ambiguity, and the engineering of choice. *The Bell Journal of Economics*, 9, 587–608.
- Marianov, V., & Serra, D. (2002). Location problems in the public sector. In Z. Drezner & H. W. Hamacher (Eds.), *Facility location: application and theory* (pp. 119–149). Berlin: Springer.
- McFadden, D. (1980). Econometric models for probabilistic choice among products. *The Journal of Business*, 53(3), S13–S29.
- Moore, M. H. (1983). Social science and policy analysis. In D. Callahan and B. Jennings (Eds.), *Ethics, the social sciences, and policy analysis* (pp. 271–291). New York: Plenum.

- Moore, M. H. (2002). The limits of social science in guiding policy. *Criminology and Public Policy*, 2(1), 33–42.
- Murray, T. (1983). Partial knowledge. In D. Callahan & B. Jennings (Eds.), *Ethics, the social sciences, and policy analysis* (pp. 305–331). New York: Plenum.
- Nakanishi, M., & Cooper, L. G. (1974). Parameter estimate for multiplicative interactive choice model: Least squares approach. *Journal of Marketing Research*, 11, 303–311.
- Nakanishi, M., & Cooper, L. G. (1982). Simplified estimation procedures for MCI models. *Marketing Science*, 1(3), 314–322.
- Nunn, S. (2005). Preventing the next terrorist attack: The theory and practice of homeland security information systems. *Journal of Homeland Security and Emergency Management*, 2(3), 1–28.
- Ogden, D. T., Ogden, J. R., & Schau, H. J. (2004). Exploring the impact of culture and acculturation on consumer purchase decisions: Toward a microcultural perspective. *Academy of Marketing Science Review*, 3, from http://www.amsreview.org/articles/ogden03_2004.pdf.
- Ou, W., Abratt, R., & Dion, P. (2006). The influence of retailer reputation on store patronage. *Journal of Retailing and Consumer Services*, 13, 221–230.
- Perlman, E. (2005, December). Pairing up: Mississippi is moving ahead by marrying GIS data to a host of state and local social indicators. *Governing*. Retrieved December 1, 2006, from <http://governing.com/articles/12talk.htm>.
- Revelle, C., Marks, D., & Liebman, J. C. (1970). An analysis of private and public sector location models. *Management Science*, 16(11), 693–707.
- Rosenberg, M. W. (1984). Locating small scale public facilities. *Socio-economic planning sciences*, 18(3), 151–158.
- Serra, D., & Colome, R. (2001). Consumer choice and optimal location models: Formulations and heuristics. *Regional Science*, 80, 439–464.
- Simon, H. A. (1996). *The sciences of the artificial* (3rd ed.). Cambridge, MA: MIT Press.
- Suárez, A., del Bosque, I. R., Rodríguez-Poo, J. M., & Moral, I. (2004). Accounting for heterogeneity in shopping center choice models. *Journal of Retailing and Consumer Services*, 11, 119–129.
- Sutherland, E. H. (1940). White-collar criminality. *American Sociological Review*, 5(1), 1–12.
- Tilly, C. (2006). *Why? What happens when people give reasons . . . and why*. Princeton, NJ: Princeton University Press.
- U.S. Department of Agriculture (USDA). (2000). *Federal register* (Part VI, CFR Part 246). Washington, DC: Author.
- U.S. Department of Agriculture (USDA). (2001). *WIC vendor management study 1998*. (Report No. WIC-01-WICVM). Washington, DC: Author.
- U.S. Department of Agriculture (USDA). (N.d.). Program data. Retrieved November 7, 2006, from <http://www.usda.gov>.
- Wang, L. (2004). An investigation of Chinese immigrant consumer behavior in Toronto, Canada. *Journal of Retailing and Consumer Services*, 11, 307–320.
- Wilhelm, W. K. (2004). The fraud management lifecycle theory: a holistic approach to fraud management. *Journal of Economic Crime Management*, 2(2), from <http://www.utica.edu/academic/institutes/ecii/jecm/>.
- Yrigoyen, C. C., & Otero, J. V. (1998). Spatial interaction models applied to the design of retail trade areas. In *38th Congress of the European Regional Science Association*. Vienna, AT. August 28–September 1.

Yushim Kim, Ph.D., is a postdoctoral researcher in the John Glenn School of Public Affairs at The Ohio State University. She worked as a researcher in the Ohio Women, Infants, and Children program for 5 years. Her research interests include spatial analysis and the use of social simulation techniques, such as agent-based models for policy analysis, focusing on public service delivery and crime. E-mail: kim.1031@osu.edu.