**CONTRACT NO. GS-23F-9777H**

# ANNUAL MEASURE OF ERRONEOUS PAYMENTS TO WIC VENDORS: 2011

## Methodology Report

May 31, 2012

Prepared and Submitted by:
Richard E. Mantovani, Project Director
ICF International Inc.
530 Gaither Road,
Rockville, MD 20850

# Contents

APPENDIXES

# 1. EXECUTIVE SUMMARY

## 1.1 INTRODUCTION

The generation of an improper payment estimate based on Special Supplemental Nutrition Program for Women, Infants, and Children (WIC) vendor over- and underpayments was last estimated from a nationally representative sample of WIC vendors by the 2005 WIC Vendor Management Study. Since that time, yearly updates to the estimates have been made using WIC administrative data. This report has three objectives: 1) to explain the current methodology that has been used in previous update studies and that will be used in the 2011 update; 2) to develop and test alternative models for generating over- and undercharge estimates; and 3) to provide a preliminary design for assessing the effect of changes to the WIC food package on under- and overcharges.

## 1.2 CURRENT METHODOLOGY

With regard to the first objective, approaches for developing estimates of overcharges have been based on a procedure for adjusting State-conducted investigative cases to the population of WIC vendors. Investigative case information derives from The Integrity Profile (TIP), an annual database provided by the States on all authorized WIC vendors and containing investigative activity by the State and other investigative agencies. When choosing vendors to investigate, State and other investigative agencies use a variety of selection criteria, which do not allow violation statistics to be simply translated to the WIC vendor population. The utilization of a post-stratification method such as raking allows this translation by providing weights for each investigated vendor based on the differences in vendor characteristics between the investigation sample and the population. The weights obtained from raking are applied to violating vendors to produce both vendor-based and food outlay-based estimates of overcharging.

The estimation of undercharges results from a three-step process using predictive equations derived from data supplied by the WIC Vendor Management Study. In the first step, each vendor in TIP is assigned a probability of undercharging. When summed, this provides the overall number of vendors undercharging. Second, based on a predictive equation, vendors are assigned the amount that they would undercharge, if they undercharged. Third, the probability of an undercharge is multiplied by the amount estimated in the second step to produce the undercharge amount for each vendor, which is summed to obtain a total undercharge estimate.

### 1.3 ALTERNATE METHODOLOGY

For the second objective, we developed an alternative specification for generating over- and undercharges using techniques discussed in the 2010 Methodology Report and addendums. The alternative approaches for estimating over- and undercharges consisted of six steps. First, models predicting the probability that a vendor will over/undercharge were developed using TIP investigative data for overcharges and 2005 WIC Vendor Management Report for undercharges. Second, the models were used to assign a probability to every vendor in the WIC vendor population based on the vendor's characteristics. Third, the total vendor over- and underpayment rates for the population were calculated by summing up the probabilities over the entire WIC vendor population, and divided by the population total. Fourth, we estimated a function representing the proportion of dollars over- or undercharged on any particular transaction from the 2005 WIC Vendor Management Report. We applied this proportion to vendor food outlays to obtain the amount that would have been expected to have been over/undercharged by over/undercharging vendors, and then multiplied by the probability that the vendor over/undercharged. Finally, we summed the product of the last step to obtain the overall over/underpayment amounts.

### 1.4 RESULTS

This approach used regression methods and simulation to derive the estimates and the dispersion statistics. The estimates derived from the alternative methods show the following.

- Overcharge violations and rates are lower than those estimated by raking by a substantial amount (i.e., at about 5 percent, they are roughly 3 to 4 percentage points lower). Dispersion statistics show that the values generated by the alternative approach assume a wider range of values. It is thought that the differences are due to the absence of State-level WIC vendor practice indicators in the raking approach.

- Overcharge amounts are lower than the estimates generated by the raking process by a considerable amount. Using all investigations, the alternative estimate is $12 million and the raking estimate is $30 million in FY2010. Using completed investigations, the difference is $19 million to $30 million. However, the range of acceptable values is wider than those estimated through the raking procedure. The difference again is thought to be due to the lack of State-level vendor management practice indicators in the raking approach.
- Incidents of undercharges approximated those obtained from using the standard approach used in previous studies; however, the dispersion associated with these undercharges is wider than those obtained from the standard approach used previously.
- The amount of undercharges using the alternate approach is slightly greater than those obtained from using the standard approach ($16 million to $14 million). The dispersion statistics show a wider range than those obtained through the standard approach. One issue related to these differences is the use of information from the 2005 WIC Vendor Management Report that showed underpayments exceeded overpayments—a finding that is anomalous with previous studies.
- When results for FY2009 and FY2010 are compared, the estimates show similar trends in rates between the standard approaches and the alternative approaches. Thus, the alternative approach seems to provide the same trend, if not the same values, as the standard approaches.

## 1.5 CONCLUSIONS

The alternative models were formulated to deal with some of the weaknesses of the raking approach. The efforts in this report are focused on demonstrating proof-of-concept of the alternative approach. This demonstration was successful in that it provides a more flexible and interpretable approach than the standard approaches used in previous studies. However, the alternative approaches do need to be developed along several lines including further specification of the models to: 1) incorporate other possible predictors, 2) convert the model to a true hierarchical model that models State-level differences, and 3) address concerns about changes in the model itself over several years. The models ought to undergo a further level of sensitivity testing, especially with regard to the major assumption governing the estimation procedure and to assumptions about State variations in vendor behaviors, and on how the results from the 2005 WIC Vendor Management Study ought to be used. They should also take into account the data that are currently being collected for the 2012 WIC Vendor Management Study.

# 2. BACKGROUND

The Special Supplemental Nutrition Program for Women, Infants, and Children (WIC) provides supplemental foods to participating women, infants, and children largely through transactions with authorized vendors. The WIC participant presents a food instrument that specifies the quantity and types of foods eligible for purchase to the vendor, who then rings up the purchase, collects the food instrument, and redeems the instrument with the State agency. These vendors include small and large food retailers, pharmacies, WIC-only vendors,[1] and commissaries. The number of authorized vendors in the United States and its territories and possessions average about 48,000.[2]

One of the programmatic concerns of the Food and Nutrition Service (FNS) is vendor overcharging.[3] Overcharging occurs when vendors, intentionally or not, charge the WIC participant more than a non-WIC customer for items prescribed by the food instrument. This practice results in improper payments to the vendor and reduces the funds available to serve WIC participants. Undercharging has also been a concern as a form of improper payment, although undercharging results in no apparent benefit to the vendor. The Improper Payments Information Act of 2002 (Public Law 107-300) requires FNS to report on these activities, including the absolute sum of overcharges and undercharges. [4]

About every 7 years, FNS conducts a WIC Vendor Management Study to examine improper payments, in particular over- and undercharges. These studies were conducted in 1991, 1998, and 2005,[5] and one is scheduled to be conducted in 2012. These studies use covert purchases in a

---

[1] These are stores that sell only WIC foods to WIC participants. In addition, there are WIC above-50-percent vendors, which do at least half their business with WIC participants.
[2] Some States, such as Mississippi and Vermont, operate food delivery systems that do not use retail vendors within that State.
[3] Other programmatic concerns include partial buys, substitutions, and trafficking, since these subvert the intention of the program. Substitution occurs when an item not on the food instrument is purchased, and trafficking involves the outright purchase of food instruments at a discount by the vendor, who then redeems them at full value.
[4] USDA estimates a national improper payment dollar figure for its *Performance and Accountability Report* (PAR) that differs from the estimate developed in this study. This study's improper payment estimate excludes several States and certain vendor types. The PAR's national estimate applies this study's improper payment rate to post-rebate food outlays across all States and vendors.
[5] Although the last WIC Vendor Management Study references 2005, it used data collected for vendors authorized at the end of the 2004 calendar year.

nationally representative sample of vendors to produce estimates of the proportion of stores over- or undercharging and the total dollar value of over- and undercharges.

On an annual Federal fiscal year basis, FNS receives information on the redemption activity for all WIC vendors as part of The Integrity Profile (TIP) data system. This system includes information on investigations and other vendor monitoring and training actions taken by States and other entities. Through TIP, a comprehensive, annually updated portrait of vendor erroneous payment activity and overall redemption activity is provided. Because it is an annual compilation of State investigative activity, TIP may be viewed as a base for updating overcharge estimates produced by the 2005 WIC Vendor Management Study (also called the 2005 bookend study). However, because vendors showing a high-risk profile are usually selected for investigation, the data from State investigations alone would be expected to overestimate overpayments for the population as a whole. Incidents of undercharges that occur within a covert investigation are not recorded in TIP. Thus, the basic estimates of overcharging would have to be derived from the 2005 bookend study and statistically "aged" to conform to vendor population and redemption changes.

Since 2005, FNS has generated annual estimates of the amount of dollars that are over- and undercharged and the number of vendors over- and undercharging. The methodology for generating the overcharge estimates has relied on statistical post-stratification of TIP investigative data. In general terms, post-stratification adjusts the investigative sample to the WIC vendor population, thereby making investigative data more representative of the population as a whole. The estimation of undercharges, due to the lack of undercharge information in the TIP file, uses a regression-based estimation model that was developed from data made available from the 2005 WIC Vendor Management Study.

This 2011 Methodology Report provides information on the post-stratification raking technique and the regression model that have previously been used for estimating over- and undercharges, respectively. These approaches will again be the primary approaches for generating estimates for the FY2011 improper payment estimates. Although there are other viable methodologies for generating these estimates, the use of current methodologies will produce an official estimate that is consistent, at least in terms of the underlying methods, with previous estimates. However, there are other approaches that can be used to translate the TIP information to the population. Some of these methods were discussed in the 2010 Methodology Report. In the following sections, we will first describe the current approaches and then explore an alternative methodology, first for overcharges and then for undercharges. The examination of the alternative approaches will be an extension of work presented as part of the 2010 WIC Erroneous Methodology Report and related memos.

# 3. METHODOLOGY FOR CONSTRUCTING FILES

The construction of the data files for generating the over- and undercharge estimates involves several steps to assemble information on the extent of their WIC business, their characteristics, and on their investigative behaviors. This section provides background on the approach for assembling these files.

## 3.1 OBTAINING FOOD OUTLAY ESTIMATES

The calculation of over- and undercharge estimates uses "food outlays", or the dollar amount spent by WIC on food (all food, including infant formula) after rebates have been applied. "Rebates" are contracted amounts paid by manufacturers (mostly formula manufacturers) to the WIC State agency for using their product. These rebates reduce the cost of foods to the WIC program. "Total WIC redemptions", which is the amount paid to the vendor by the State for purchases made within that store, is the only information on transaction quantity provided on The Integrity Profile (TIP) dataset. Since the WIC redemption figure in TIP includes rebates, it has to be adjusted to reflect food outlays only. To do this, we use State level food outlays figures published by FNS (http://www.fns.usda..gov/pd/24wicfood$.htm) for a particular year to adjust redemption figures reported by each vendor. This adjustment can be done within State through the following formula.

$$Food\ Outlays\ for\ vendor\ i = \frac{Food\ Outlays\ for\ State\ j * Redemptions\ for\ Vendor\ i}{Total\ Redemptions\ for\ State\ j}$$

Where food outlays for State j is the total amount reported on the FNS website for State "j', total WIC redemptions for a vendor "i" authorized by State "j" is provided by TIP, and total WIC redemptions for State "j" is the sum of all the WIC redemptions reported in TIP for vendors authorized in State "j". This formula adjusts the share of total WIC redemptions for each vendor within the State proportional to the total food outlays. Since our focus was on calculating food outlays for vendors, we were less concerned with that figures exactly matching the total overall redemption figures that are published on the FNS website. Thus a food outlay figure is calculated for each vendor reported in TIP.

The above procedure was used on the entire population of WIC vendors in the 2010 study and will be used in the 2011 study. In these studies, when particular vendors were eliminated as out-of-scope (see 3.2 below), the total food outlays associated with in-scope vendors, if summed, were reduced to levels below that which was reported on the FNS website. In studies conducted in 2009 and before, the approach was applied after the out-of-scope stores were eliminated. In that case, the total food outlays should be very close approximations to the figures reported on the FNS website. It is likely, then that in these reports, the amount of overcharges could be overstated—the amount depending on the degree to which overcharging occurred in vendors that were declared out-of-scope.

### 3.2 IDENTIFYING IN-SCOPE VENDORS

The 2005 WIC Vendor Management Study was conducted in 45 States. It excluded the territories and ITO State designated agencies. The reasons for excluding these entities related to survey cost and logistic reasons (Alaska, Hawaii, North Dakota, the ITOs, and territories) and to the type of food delivery system present in the State (Mississippi used a direct distribution vendor that was outside of the traditional retail vendor system and Vermont was a home delivery system).

Certain types of vendors were also eliminated. These included: commissaries, direct distribution vendors, and home delivery systems. In updating the over- and undercharge estimates, we attempted to replicate the 2005 WIC Vendor Management Study and thus used the criteria used in that study to select vendors. In addition to those vendors mentioned above, we also eliminated pharmacies.

Although in many cases, identification of in-scope vendors involved using the coding supplied on the TIP file, we also performed manual edits to catch instances in which a vendor was perceived to be misclassified. This was particularly apparent in the case of commissaries that were not coded as commissaries in the TIP file but were obviously commissaries as determined from the name and location information provided.

### 3.3 IDENTIFY VENDOR CHARACTERISTICS

Two vendor-level characteristics were used for generating the estimates: store type and ownership. TIP only provides information on the following vendor types that were defined as being in-scope: retailers, WIC 50 percent stores, and WIC-only stores. The original assumption was that store type and size were important from the perspective that larger stores (supermarkets) were less likely to over- or undercharge relative to smaller stores. In addition, these larger stores were less likely to be selected for investigation. Although two of the store types were fairly well defined, the retailer category was much too heterogeneous to serve any purpose for deriving the estimates. This category was split into three categories based on data available from the SNAP STARS database.

Likewise, an assumption was made that privately owned retailers would differ from publicly owned retailers in their over- and undercharge behaviors, and would be more likely to be chosen for investigation than publicly owned stores. The thought was that publicly owned retailers, who would be in general more responsible to shareholders, would have procedures in place to ensure the integrity of their transactions. This information was again available from the SNAP STARS database.

The effort to retrieve information from the SNAP STARS database was facilitated by the SNAP ID that was available on TIP. SNAP ID's from TIP were matched with STARS ID's and STARS information was merged with TIP information. Attempts were made to resolve non-matching ID's by manually attempting matches based on store name, location and other information.

After retrieving the information from STARS, the vendors were recoded in the following way.

- WIC retail vendors were designated as large retailers if they were classified as superstores, supermarkets and large groceries, or if they were medium sized groceries with over 500,000 dollars in total sales. Otherwise, if the WIC vendor was included in the STARS file, they were classified as a small retailer. If they were classified as a retailer in TIP but were not identified in STARS, their store type was classified as unknown. This third category was further examined to determine store type (for instance, if the store was within the Safeway Chain, it was classified as a large retailer).
- A similar procedure was used for determine public versus private designation, with STARS data being used to determine a publicly traded firm, while all other matching WIC retailers being assigned a privately owned designation. Those that could not be matched were assigned to an unknown category unless they could be designated through visual examination to a private or public designation. The visual examination used Web-based information to identify those chains that were publicly traded. We did not designate non-chain stores.

In updating the information on the TIP file, we used resolved information from previous efforts. In other words, if a store was classified as a large store in a previous study, it was also classified as a large store in the current study. This meant that only newly authorized stores needed to be classified. However, in 2007, FNS decided to revise the store type classification scheme previously in place, and changed categories. For the 2009 studies onward, we used the new classification scheme—previous to that time we used the old scheme. Because we separated WIC vendors into two categories, the changes made in the classification scheme would have little impact on store classifications. The only affect would be for vendors classified as SNAP medium sized groceries, which included small and medium sized groceries identified in the previous scheme.

## 3.4 IDENTIFYING VENDOR'S NEIGHBORHOOD CHARACTERISTICS

Two neighborhood indicators (poverty level and urbanization) were identified as useful both for adjusting the investigation sample to the population and as possible indicators of vendors who might be more prone toward overcharging. Neighborhoods were defined as ZIP code areas. This focus on ZIP codes provided first an easy way to locate vendors since they provide this information in TIP, and second, are more realistic than counties in that they provide a more precise demographic characterization of the area surrounding the vendor.

Census data from Census 2000 was used to provide this information; however, Census does not provide demographic information by Postal ZIP codes, but by a designation they term "ZIP Code Tabulation Area", which largely corresponds to the ZIP code designations but because they are built up from census tracts do not have the same boundaries. In addition, many ZIP codes do not have corresponding ZCTA's due to their being associated with work places, airports, or other areas that contain no residents. The transformation of TIP provided ZIP codes to ZCTA's codes was accomplished by the following rules.

- If the ZIP code matched a ZCTA code and both were in the same State then the demographic data for that ZCTA was added to the vendor's TIP record.
- If not, the vendor's full address and name were used to look up the vendor's zip code, and this was checked against existing ZCTA codes. If it was matched, then the ZCTA information was merged with that vendor's record.
- If there was no match to a ZCTA code, the ZCTA code closest to the ZIP code was selected.

Since most WIC vendors do not change addresses, previously resolved matches between ZIP codes and ZCTA were used as a basis for resolving current TIP vendors. For the 2005 study, approximately 2 percent of the WIC vendors did not submit ZIP code information that matched a ZCTA code. The procedure was to look-up the vendor provided ZIP code, gather information about surrounding ZIP codes that could be viable substitutes, and then match these to Census ZCTA codes. In selecting a substitute ZIP code, we took into account factors such as the proximity of the ZIP code and the shape of the substitute ZIP code. In each subsequent study, the number of ZIP codes that had to be manually matched to a ZCTA code was between 50 and 100. [6]

The resulting added fields were resolved to generate the following measures.

- Urbanization—based on the percent of population in the ZCTA that lived in tracts designated by Census as being urban. This percentage was resolved into areas that were largely rural (50 percent or less lived in urban areas), moderately urban (50 to 90 percent of the population living in urban areas) or highly urban (90 percent or more of the population living in urban areas).
- Poverty Level—based on the percent of households in the ZCTA who were had household incomes less than the poverty level. Three poverty levels were established: 1) poverty level at or below 20 percent, 2) poverty level between 20 and 30 percent, and 3) poverty level over 30 percent.[7]

## 3.5 IDENTIFY LEVEL OF REDEMPTION ACTIVITIES.

As described in Section 3.1, total WIC redemptions claimed by retailers were resolved into food outlays. Vendors were identified with one of four levels depending on their food outlays quartile and this served as another dimension for comparing the investigation sample and the population. Unlike the other vendor descriptive variables, redemption activity varies from one year to the next. The variable represents the total WIC business, which may reflect factors other than store size and type. For instance, a Safeway may not provide service to many WIC participants

---

[6] Although there is some error is involved because the characteristics of the ZIP code designations may not exactly match the characteristics of the ZCTA code, this approach certainly provides better information than if county was used to designate urbanization and poverty level characteristics. A more costly but more accurate approach would focus on geocoding the stores and matching them to Census tracts. Neighborhood characteristics can then be drawn from tract information, or from aggregations of tracts, formed by nearest neighbor criteria. It should also be noted that Census ZCTA level characteristics reflect small area estimates themselves, and are subject to error.

[7] These limits were based on an analysis that examined how retailers were distributed by Census poverty and urbanization across Census ZCTAs. The aim was to attempt to identify ranges that would contain a sufficient number of retailers to be statistically viable yet be conceptually meaningful in denoting differing levels of poverty and urbanization. The original analysis was completed as part of SNAP trafficking studies in the mid-90's. For WIC, we reduced the number of categories used in SNAP from four to three to accommodate the smaller population of WIC food stores. This was accomplished by combining the two lowest categories, so that it maintained highly urbanized areas, and high poverty areas as separate entities.

because of its food choices or prices, while a much smaller store may predominately serve WIC customers in relative high volume.

## 3.6 DETERMINATION OF OVERCHARGES

TIP contains a variable denoted as “Sanction Reason” and which contains a series of letters identify the particular violations found in an investigated store. A code of “O” indicates an overcharge. This code was extracted wherever it appeared and recoded to 1 to indicate an overcharge violation and 0 to indicate no violation.

# 4. APPROACH FOR ESTIMATING OVERCHARGES

An overcharge occurs when the WIC Program makes a payment to a vendor for a food item that is greater than the price that a non-WIC customer would have paid. This definition guides activity related to establishing whether the vendor overcharged during covert investigations. As indicated above, TIP data present a general profile of the WIC vendor population and also of those vendors that are investigated. The working assumption is that TIP investigations serve as the primary source of information required to make an overcharge estimate, with the 2005 bookend study statistics used for certain adjustments, as explained below.

## 4.1 DISCUSSION OF THE CURRENT METHODOLOGY

The estimation approach for overcharges using the current methodology involves three steps:

- the estimation of weights that allow information on investigated vendors to be translated to the population of vendors;
- the application of those weights to vendor overcharge estimates; and
- the application of an adjustment factor for characterizing vendors' erroneous payment behaviors.

These steps are described in the following sections.

### 4.1.1 Estimation of Weights

The approach used for developing overcharge estimates is a post-stratification adjustment known as raking.[8] An illustration that describes this technique is provided in Appendix A. Raking begins with defining the vendor population that was investigated by a State or other investigative agency. Exhibit 1 shows the number of vendors in the FY2009 and FY2010 TIP files that were

---

[8] Raking is a technique that derives weights by adjusting the sample distribution on a single or multiple variables to the population distribution. It is usually done when joint distributions over the post-stratifying variables are not present. If joint distributions are present, which they are in this case, we could simply calculate the direct weights relating the sample to population within each cell. Raking was originally proposed and accepted by FNS as a way to smooth the weights to avoid extremes. Since the original discussions, the raking approach has been utilized as a way to maintain longitudinal consistency of the estimates.

investigated and who were sanctioned for overcharging by the investigative source.[9] Compliance investigations are covert activities in which an undercover purchaser seeks to uncover instances of error, fraud, and abuse.[10] In FY2009, there were 6,373 vendors undergoing compliance investigations, and in FY2010, there were 5,929 vendors.[11] The number and percentage of those overcharging for each year were: 903 vendors, or 14.7 percent in FY2009 and 781 or 13.2 percent in FY2010.[12] In identifying overcharging, only violations in which the State indicated that the reason for sanction was an overcharge were included. Other violations, such as substitutions or trafficking, were not counted as violations for this study.

**Exhibit 1.**

**Frequency of Overcharges by Year (TIP 2009 and 2010)**

| Investigative Source | Total Investigated | Total Overcharging | |
|---|---|---|---|
| | | Sanctioned | Percent |
| Compliance investigations by State agency or other entity in 2009* | 6,373 | 903 | 14.7 |
| Compliance investigations by State agency or other entity in 2010* | 5,929 | 781 | 13.2 |

* The TIP User Guide Data Dictionary defines investigations by other entities as "compliance investigations conducted by an outside agency, such as another State agency or the Food Stamp Program, or a Federal law enforcement agency."

Source: Annual Measure of WIC Erroneous Payments to Vendors: 2011 Methodology Report.

[9] There are other reasons for sanctions indicated in the TIP file. For these estimates, we did not attempt to estimate the extent of violations related to these reasons, nor did we attempt to investigate the association between various types of sanctions.

[10] A compliance buy is a covert onsite investigation in which a representative of the program poses as a participant, parent or caretaker of an infant or child participant, or proxy; transacts one or more food instruments; and does not reveal during the visit that he or she is a program representative (7 CFR 246, p. 314).

[11] The number does not include vendors located in ITOs or in North Dakota, Hawaii, Alaska, Vermont, Mississippi, and the territories.

[12] This number represents all cases that were undergoing or had undergone investigations. The TIP file makes a distinction between cases marked as completed and cases that are ongoing. The distinction leads us to consider two denominators in calculating a violation rate. Because there is no tracking of individual cases in TIP, it is sometimes difficult to detect when or if initiated cases are completed. In some instances, it appears that some vendors may have been subject to two investigations within the fiscal year, but only yielding one outcome. Because of this, a decision, while formulating the methodology in 2005, was made to include all cases in the estimate, not just completed cases. Since that time we have included all investigation in the denominator to provide a consistent perspective on the estimate.

As indicated above, the raking procedure attempts to translate investigative sample results to the population through a set of characteristics, which are then organized into a matrix. The five characteristics over which the data were raked (i.e., vendor type, ownership, urbanization, poverty level, and food outlay dollar quartile) were chosen on the basis of previous research in the Supplemental Nutrition Assistance Program (SNAP) showing a relationship between food stamp trafficking and vendor and neighborhood characteristics.[13] That research identified a basic set of indicators that, when modified to fit the WIC population, would be useful for characterizing WIC transactions and examining WIC over- and undercharges.[14] Exhibit 2 provides details on the variables used during the raking process, which together define 540 cells. The raking process, as discussed in Appendix A, repeatedly churns through the data until the marginal distributions for the sample equal the marginal distributions for the population across all dimensions. The result is a weight for each of the 540 cells in the matrix. The weights can be viewed as similar to sampling weights, and they have the same purpose of inflating the estimate made within each cell to the population.

### 4.1.2 **Application of Weights**

Raking weights were used in conjunction with overcharge information to form two estimates. The first was an estimate of the number of vendors overcharging, and the second was an estimate of the amount of food outlays overcharged. The population estimate of vendors that overcharged was the sum of the weighted number of vendors found to be overcharging within the sample. In other words, each vendor in the investigative sample was assigned a weight as a result of the raking process. The weight is interpreted as the number of stores in the population represented by each of the investigated stores. The sum of these weights for investigated stores that overcharged provides the number of stores that overcharged in the population. The vendor-based overcharge rate was the weighted number of overcharging vendors divided by the total vendor population.

---

[13] See U.S. Department of Agriculture, Food and Nutrition Service, Office of Analysis, Nutrition and Evaluation (2003). *The Extent of Trafficking in the Food Stamp Program: 1999–2002, FSP-03-TRAF, by Theodore F. Macaluso, Ph.D.,* Alexandria, VA, and U.S. Department of Agriculture, Food and Nutrition Service, Office of Analysis, Nutrition and Evaluation (2000). *The Extent of Trafficking in the Food Stamp Program: An Update, by Theodore F. Macaluso, Ph.D.,* Alexandria, VA.

[14] Other variables examined in the 2010 Methodology Report (risk and new vendor status) showed some potential in predicting overcharges in particular. However, when used in the context of the raking approach, the estimates were similar to those produced with the five primary variables.

| Exhibit 2.<br>Variables and Variable Categories Used in the Overcharge Raking Process | | |
|---|---|---|
| **Variable** | **Categories** | **Justification** |
| Vendor Type | Large Retailers (Retailers Defined as Having More Than $800,000 in Gross Sales)<br>Small Retailers (Retailers Defined as Having Less Than $800,000 in Gross Sales)<br>WIC Retailers With Missing Information on Gross Sales<br>WIC-Only Stores<br>WIC Above-50-Percent Stores | Vendor type was found to be significant in the WIC Vendor Management Study relative to differences in both over- and undercharging. SNAP trafficking studies have reinforced the idea that smaller retailers are more violation prone than larger retailers. The size of the retailer, derived from the Stores Tracking and Redemption System (STARS), differentiated among retailers identified in the TIP data. The value used to distinguish between large and small retailers was derived from previous studies of SNAP trafficking and from our need to limit the number of categories. |
| Ownership Type | Publicly Owned Stores<br>Privately Owned Stores<br>Ownership Unknown | Public and private ownership values were drawn from the STARS database. Values for stores for which the ownership type was unknown were largely retailers and other stores that could not be matched to STARS. Using TIP data as well as data on SNAP retailers, public ownership was found to be associated with fewer violations, probably due to the greater need for corporate controls. |
| Poverty Level of the Vendor's Neighborhood (defined by ZIP Code as the number of households under the poverty level) | 20 Percent or Less<br>More Than 20 Percent but Less Than 30 Percent<br>30 Percent or More | Vendors in poorer neighborhoods were found to be associated with higher levels of SNAP violations, and this variable was therefore carried over to the WIC erroneous payment update studies. |
| Urbanization Level of the Vendor's Neighborhood (defined by ZIP Code as the number of individuals who live in urbanized tracts within the area.) | 50 Percent or Less<br>More Than 50 Percent but Less Than 90 percent<br>90 Percent or More | Vendors in more urbanized neighborhoods were found to be associated with higher levels of SNAP violations. This variable, particularly in conjunction with the poverty-level variable, was a powerful predictor of places in which rules and regulations may be relaxed to permit certain illegal behaviors. |
| WIC Food Outlays | Quartiles Based on Current-Year Food Outlay Distribution | This variable was introduced to control for the wide range of WIC food outlays between stores. It should be noted that the focus is on food outlays rather than total WIC redemptions, which includes formula and food rebates. |

The maximum amount of food outlays that were overcharged was calculated using a similar process. Food outlays for overcharging vendors were multiplied by the food outlay-based weights to obtain overpayments for vendors with a particular set of characteristics and these were added over all vendors to obtain the total overpayment amount, which represents the unadjusted value of overcharges. The next step would be to adjust these values to arrive at an estimate of the actual amount of food outlays overcharged.

### 4.1.3 Application of an Adjustment Factor

Approximately $342 million and $281 million in food outlay dollars were associated with vendors that overcharged in FY2009 and FY2010, respectively. This overcharge estimate represents all food outlay dollars associated with these vendors, as if they overcharged the entire amount on the food instrument on every WIC purchase.

Overcharging can occur in several ways, all of which implies that overcharging represents a portion of the entire food outlay amount.[15] First, the vendor may, intentionally or not, charge more than the shelf price for foods bought. In this case, the total charge includes the amount of the item plus the overcharge. Second, they may charge for items not purchased during a partial purchase of the items specified on the food instrument.[16] In either case, at least some of the items on the food instrument are purchased, and the amount of the actual amount overcharged will be a proportion of actual food outlays associated with the retailer.[17]

The 2005 bookend study provided data that were useful in computing this adjustment factor. The study included three types of purchases (safe, partial, and substitution) with a food instrument from a particular sampled vendor. The study provided information on the overall charge for each type of buy and the amount that was supposed to be charged. Therefore an overcharge adjustment factor can be identified as a percentage of the total value of the food instrument that was redeemed. For the purposes of the estimate generated from the raking methodology, only safe buys were used, although the values were larger when partial buys of items on the food instrument were transacted (Exhibit 3).

Exhibit 3 shows that the average overcharge from the last bookend study was $1.82 for safe buy violations. It should be noted that this amount reflects the activities of only those vendors that overcharged, which were very few. The data also show that the amount of the overcharge was very small in many cases. For example, for safe buys the minimum overcharge was $0.02, with 25 percent of all safe buy overcharges valued at less than $0.20.

---

[15] If the total value of the food instrument was redeemed by the vendor, it would probably be an instance of trafficking.

[16] During a purchase in which all items on the food instrument are bought, additional charges could occur up to the not-to-exceed amount specified on the food instrument. During a partial buy, charges for items not purchased could occur. Because the documentation has no information on what is bought, there is no way for the State to assess these additional charges. Electronic Benefit Transfer States, by assessing what is bought in real time, avoid this situation.

[17] It is also the expectation that vendors would overcharge only on a proportion of WIC purchases. The extent to which is the case is unknown. We have basically assumed that an overcharge was made on every purchase. This assumption generates a conservative estimate—that is one that would overestimate the amount of overcharges. In counterbalance, since the same vendors are not being monitored over the entire year, the data represents a point estimate for each vendor. Thus, vendors who overcharge occasionally may not be observed to overcharge within a very limited span of time, and thus would be counted as non-violating. Our expectation is that a greater number of vendors would be associated with overcharging if they were monitored over the entire year.

| Exhibit 3. Weighted Distribution of Overcharges in the 2005 Bookend Study, by Buy Type | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Buy Type** | **No. of Buys** | **Average** | **Minimum** | **25th Percentile** | **Median** | **75th Percentile** | **Maximum** |
| Safe | 46 | $1.82 | $0.02 | $0.20 | $0.64 | $2.01 | $10.00 |
| Partial | 65 | $7.86 | $0.02 | $0.44 | $2.39 | $7.87 | $65.54 |
| Minor Substitution | 39 | $4.38 | $0.01 | $0.30 | $0.71 | $2.40 | $67.00 |
| Major Substitution | 24 | $1.57 | $0.02 | $0.20 | $0.60 | $2.16 | $9.30 |

Source: Annual Measure of WIC Erroneous Payments to Vendors: 2011 Methodology Report.

The 2005 bookend study provides a method for determining the average percentage overcharge. When overcharge data were aggregated and weighted across retailers, it was found that of those retailers that overcharged, the overcharge was 10.74 percent of the total purchase, which was used as the adjustment factor in the raking analysis. Exhibit 4 presents summary statistics on safe buy overcharges. Using the adjustment factor, the amount in FY2009 is reduced to $37 million, or 0.86 percent of total food outlays, and the amount in FY2010 is reduced to $30 million, or 0.74 of total food outlays. The food outlay-based overcharge rate was the amount of overcharges found in the population of overcharging vendors divided by the total amount of food outlay dollars reported in the population.

| Exhibit 4. Mean 2005 Bookend Study Overcharge as a Percentage of the Food Instrument for Safe Buys Only | | | | |
|---|---|---|---|---|
| **Number of Safe Buy Overcharges** | **Mean Overcharge Percent** | **Standard Deviation** | **Minimum** | **Maximum** |
| 46 | 10.74% | 77.87% | 0.07% | 73.64% |

Source: Annual Measure of WIC Erroneous Payments to Vendors: 2011 Methodology Report.

#### 4.1.4 **Derivation of Sensitivity Estimates**

It is important to note that the above approach provides a point estimate that is dependent on the investigations conducted in a particular year. The investigations are a sample of all possible draws from the WIC vendor population, and thus a different sample might produce a different outcome. The sample, as stated previously, is drawn with consideration of risk profiles determined by the States. The point estimate can be expected to reflect the particular sample chosen and its relationship to the entire WIC vendor population. In other words, if there is a large variance in the relationship of the investigative sample relative to the WIC population, the point estimates would be expected to vary accordingly. However, if the investigative sample was fairly homogenous, we would see smaller variations in the point estimates. To investigate this situation, the variation in the point estimate was examined by taking different random draws of the investigated vendors. Each draw was subjected to the raking procedure previously described. In all, the process involved at least over 1,000 draws, thus creating at least 1,000 estimates, from which we calculated mean values and 5$^{th}$ and 95$^{th}$ percentile intervals. It should be noted that the interpretation of these statistics reflects the consistency of the investigative sample in producing overpayment estimates. The percentile values reflect not only the consistency in the selection of cases for investigation, but also the size of the subsamples that are used in the iterative estimation approach. A wide range of vendors will result in large variances as their results are translated to the larger population, as will smaller samples. Thus, the percentile intervals should be examined more as a reflection of how the investigative sample is structured, and the sensitivity of the overall estimates to that structure.

## 4.2 ALTERNATIVE OVERCHARGE METHODOLOGY

The initial discussions prior to producing the first "update" study in 2005 were guided by the primary data sources available and the statistical approaches in use for balancing sample information (TIP investigations) to population totals. For the 2010 Methodology Report, several analyses were conducted to examine alternative methods of estimation. In this report, we will continue efforts to generate an alternative model for generating overcharge estimates.[18]

[18] We will discuss alternative approaches for estimating undercharges in Section 5.2.

### 4.2.1 **Framework for Estimating Undercharges**

An appropriate theoretical model for estimating over- and undercharges is one that is vendor-based rather than food outlay-based[19], and assumes that over- and undercharges are generated intentionally or unintentionally by vendors. A model for over- or undercharge behavior is based on the following:

$$WC = \sum_{k=1}^{n} I_k P_{k,wc} \qquad (1)$$

$$NWC = \sum_{k=1}^{n} I_k P_{k,nwc} \qquad (2)$$

Where:

***WC*** is the total cost of the foods purchased from those designated on the food instrument
***NWC*** is the total cost of the foods (less tax) purchased by a non-WIC customer from those equivalent to those specified a food instrument,
$\boldsymbol{I_k}$ represents the k$^{th}$ item on the food instrument,
$\boldsymbol{P_{k,wc}}$ represents the price charged to the WIC customer for the k$^{th}$ item, and
$\boldsymbol{P_{k,nwc}}$ represents the price charged to the non-WIC customer for the k$^{th}$ item

An over- or undercharge occurs if WC ≠ NWC and the amount of the over- or undercharge is defined as WC-NWC. For a retailer, the net total overcharge or undercharge for a year is defined as:

$$NMC = \sum_{t=1}^{n} (WC - NWC)_t \qquad (3)$$

[19] Methods using the raking approach separate out the processes for estimating vendor- and outlay-based rates—however in this more theory-based approach, the notion is that vendors initiate an error, and that erroneous outlays occur as a result of that error.

Where:

***NMC*** is the net total over- or undercharge,

***t*** represents a WIC transaction.

In this case, WC and NWC have to be measured in the same time frame to control for price changes within a store.

In theory, vendor behavior with regard to overcharging (and also to undercharging) reflects both a random component and a fixed component. The random component entails inadvertent (human) errors made by retailers during check-out. In general, we would expect this random component to yield overcharges equivalent to undercharges when viewed over all WIC transactions by that vendor. The use of scanning and other electronic devices will reduce these random errors by ensuring human error is minimized. It may be the case that certain items that change price relatively frequently may trigger errors, and therefore may introduce some systematic over- or undercharge errors.

The fixed component represents attempts by WIC vendors to intentionally overcharge. This may reflect 1) charging more than list price for items being purchased, 2) charging for items not purchased (in the case of a partial buy), or 3) other behaviors that discriminate between what retailers charge the WIC program and what they charge other customers. Such overcharges are expected to be reduced substantially when real-time scanning equipment is used in conjunction with EBT services. In some cases, the overcharge violation may be triggered by the amount of benefit that might be gained. In other words, overcharging by a few cents may not be worth the vendor’s effort, but the potential for overcharging may increase if the vendor realizes a dollar benefit of 5 to 10 dollars from the transaction. This assumed relationship may explain the finding from the 2005 WIC Vendor Management Study that partial buys, which provide a chance for large overcharges, actually do result in higher overcharges.

We are unaware of any fixed component to intentionally undercharge since undercharging results in no monetary benefit to the retailer.[20]

### 4.2.2 **Post-Stratification**

The approach described in 2.1 is a post-stratification strategy that allows translation from the TIP investigations file to the entire population. This strategy can be represented as:

$$Y = \sum_{k=0}^{n} \mathrm{w_k v_k} \quad (4)$$

where Y is the total number of vendors estimated to be overcharging; k is a cell designation representing the combinations of store type, ownership type, poverty level, urbanization level and food outlay level; $w_k$ is the weight that is derived for each k cell through raking; and $v_k$ is the number of overcharging vendors defined by the characteristics associated with cell k. Both v and w are affected by how the cells are defined. In the case of post-stratification procedures described in Section 4.1, the five variables and the variable categories chosen to define k were done based on previous analyses for estimating food stamp trafficking (see Exhibit 2). However, the selection of these variables and levels result in some of the cells as having no or very few representatives in the sample or the population, or for those dimensions that represent continuums (poverty and urbanization) demonstrating little sensitivity towards changes in how the categories are determined. Of the 540 cells that constitute the raking matrix, just less than half (268) were represented in the TIP population.[21] Exhibit 5 represents the distribution of investigated cases among those 268 cells. In general, just over a third of cells contained no information on the investigated cases, and almost two-fifths contain less than 10 observations. Together, these two categories represent over three-quarters of the cells. When no investigations are conducted within a cell, the violation rate for that cell is by definition zero, although the rate

[20] This Statement reflects an assumption that vendors do not want to subsidize WIC by not fully capturing the costs of the items on the food instrument. There may be other reasons, not fully understood, for intentional undercharges. The 2005 bookend study, in fact, found that undercharges exceeded overcharges, which might suggest a bias toward undercharging.

[21] The total number of cells equals 5 store type levels times 3 ownership type levels times 3 poverty levels times 3 urbanization levels times 4 food outlay quartiles for a total of 540 cells.

is really an unobservable.[22] When the investigations are less than 10, the raking weight--which extrapolates to the general population--will be relatively imprecise because of the small cell sample size.[23]

| Exhibit 5. Number of Cells by Total Number of Investigations Within Cells | | |
|---|---|---|
| **Number of Investigations Within Cells** | **Number of Cells** | **Percent of Cells** |
| Cells With Zero Investigations | 96 | 35.8% |
| Cells With 1 to 10 Investigations | 111 | 41.5% |
| Cells With 11 to 25 Investigations | 23 | 8.6% |
| Cells With 26 to 100 Investigations | 26 | 9.7% |
| Cells With 101 or More Investigations | 12 | 4.5% |

Source: Annual Measure of WIC Erroneous Payments to Vendors: 2011 Methodology Report.

Three limitations of the raking procedure include the inability to:

- cluster cells to obtain adequate sample sizes for estimating violation rates,
- deal with continuous variables that would ease restrictive requirements related to categorizing the dimensions, and
- include all factors important in translating results from investigation to the population.

The last is particularly important since with additional factors, the number of cells in the raking matrix increases thereby reducing the number of observations within the cells. The resulting sparsely populated cells further affects our ability to obtain reliable estimates of violation rates. At some point the weights derived represent very few cases within each cell, leading to the possibility of extreme variations in the raking weights and instability of the estimates.

[22] The absence of any investigations does not imply that overcharging is not occurring for the population represented by the cell—there is just no evidence for that.
[23] The imprecision or variance within each cell could be modeled by a Bernoulli process in which an overcharge with a cell represent a failure and an investigation with that same cell represents a trial.

### 4.2.3 A Regression Approach to Post-Stratification

A methodology that was initially discussed in the 2010 Methodology Report was a regression approach aimed at controlling for factors that account for differences between the sample and population. This topic was addressed by Gelman[24] as a way to tackle complicated post-stratification challenges in surveys. Using the definition in equation 4, the $v_k$ component is formulated as a probability function of a set of variables including main effects and all significant interactions, and the predicted probability of overcharging in $po_k$. The k in this case is defined as the set of predictors used in the linear function. Instead of raking to obtain the w's, the cell estimate "$po_k$", or the probability of an overcharge violation, can be applied to each vendor in the population. In this instance, the sample becomes self-weighting.[25] A regression approach has the advantages of allowing flexibility in the specification of the model in terms of the number and form of the predictor variables. In other words, we do not have to specify every combination of the variables used as predictors.

Using this approach, this model-based regression estimate consists of six steps.

- First, using the 2010 TIP data, a probability of an overcharge is estimated from a regression model using investigated vendors.
- Second, this equation is applied to each vendor in the population to obtain a probability of overcharging for that vendor.
- Third, probabilities over all vendors are summed to yield the total number of vendors overcharging.
- Fourth, the average overcharge for a violating vendor is estimated from the 2005 WIC vendor management study and applied to each vendor's food outlay figure in the population to obtain an adjusted food outlay figure.
- Fifth, the probability of an overcharge is multiplied by the adjusted food outlay figure to obtain the expected overcharge amount for that vendor.

---

[24] Gelman, A. Struggles with Survey Weighting and Regression Modeling, Statistical Science, 2007, Vol.22, No. 2, pp.153-164.

[25] The function for predicting y takes into account the post-stratification factors, and thus provides for each individual vendor the probability of violation given the post-stratification factors.

- Sixth, the expected overcharge amount estimated for each vendor is summed to derive the estimated overpayment value for the entire WIC population.

### 4.2.4 Estimating the Vendor Overcharge Rate

The first step in estimating an overcharge is to generate the probability that a vendor will overcharge. Our starting point is the following model:

$$po(overcharge) = \frac{e^z}{1 + e^z} \qquad (5)$$

Where po is the probability of a vendor overcharging and z = $\sum_{k=1}^{n} \beta_k X_k + \varepsilon$ is a logistic function where z is whether the vendor has overcharged or not, the X's are the predictors of overcharges, the β's are the parameters to be estimated, and ε is the error term with zero mean. The vector of variables (X) could include all the main effects and interactions specified by the raking matrix, but there are good reasons not to. As indicated before, there are 540 individuals cells being represented in the raking matrix, which reflects the X regressors were we were to use the raking matrix as a model for developing the model. If we include the intercept, main effects, and two, three, four and five way interactions, the number of terms expands to 1,401. This is a considerable number of terms, many of which have no information due to lack of representation from the investigation sample. Also many of the predictors would probably lead to convergence or estimation issues.

In the 2010 Methodology Report, we generated results relating to completed investigations in 2009 using regression analyses on a subset of variables. That analysis provides some evidence on the specification of the model. The findings, using a logistic regression focused on violations, indicated that:

- Store ownership was significantly related to overcharge violations, with public stores showing a lower tendency to overcharge.
- Vendor type was significantly related to overcharge violations, with large retailers showing a lower propensity to overcharge.

- Urbanization was significantly related to overcharge violations, with lower density areas showing the lowest propensity to overcharge.
- Poverty level was not significant.
- Previously authorized vendors were most likely to overcharge.
- High risk vendors were most likely to overcharge.

Although we did not extensively look at interactions terms, the ones that were examined were not significant. The purpose of that analysis was to determine the extent to which the dimensions and variables used in the raking process were useful. The analysis did not attempt to generate estimates based on the model.

There are two tenets that underlie the regression approach to dealing with adjusting sample values to the population and for the selection of independent variables. First, all factors included in the model should be important in identifying differences between the population and the investigation sample and be useful in predicting a violation. Second, the regression model should be parsimonious—that is it should be as compact as possible. The first tenet would specify a model with a large number of factors, while the second would limit these factors depending on their relationship to overcharging and on the effort needed to derive stable models.

One of the issues associated with including too many "like" variables in the model is misspecification, which could lead to collinearity and over-specification issues. There is a large literature, from the classical frequentist perspective, addressing collinearity in particular within regression problems both as a result of sample size and specification issues. However, we will be utilizing Bayesian methods to develop our predictive equations. In terms of protecting against collinearity issues, Bayesian estimation methods allows us to specify appropriately specified priors to describe the uncertainty associated with the specification of model and thus reduce the effects of collinearity. (See Gelman et. al.[26]). Under this Bayesian approach, it is also possible to specify a large number of variables, which under classical assumptions would lead to overfitting, but within a Bayesian framework suggests that the cost-benefits favors keeping even poorly performing variables. It should be noted that the modeling in this section and in Section

[26] Gelman, A., Carlin, J.B., Stern, H.S. and Rubin, D.B. Bayesian Data Analysis (Second Edition). Chapman and Hall/CRC. Boca Raton, FL 2004. Pp. 369-372

5.2 are to be considered initial forays into this type of approach, mostly focusing on generating a proof of concept model that could be used as a basis for refinement, but also to generate substantive information about the relationships that would influence overcharges.

Given this previous investigation, we first sought to reduce the number of terms in the model to increase tractability. As a first step, we eliminated cells or categories that had fewer than 10 investigations. The elements that were identified from this process were then used within a logistic modeling framework to allow us to iteratively eliminate interactions that presented estimation problems or did not meet significance criteria (p=0.10). This left us with a limited set of predictors, which were then modified further to be more effective in modeling overcharge behavior.

- Store Type—after inspecting this variable as originally formulated, this variable was collapsed into a binary variable contrasting large retailers and all other stores, which contained small retailers, WIC only stores and WIC 50 percent stores, as well as stores that we could not classify in the raking approach described in Section 4.1. The theory is that large stores are more likely to have sophisticated systems to track purchases and thus are less likely to err. In addition, fewer large stores are chosen for investigations relative to their presence in the population and thus are not represented proportionately in the investigations sample.
- Percent of the Population that is Urban—this variable was specified in continuous form rather than the categories specified in the raking approach. The percentage of the population in urban areas was used to distinguish vendors operating in dense urban areas from those operating in more sparsely populated areas. We expect that investigations would be more likely to occur in urban areas than in rural areas and this variable acts to adjust the investigation sample to population totals.
- Percent of Households in Poverty—like urbanization, this variable was also specified in continuous form. This variable not only is critical for adjusting the investigation sample to the population, but also may be an important variable in designating neighborhoods in which vendors demonstrate a higher propensity for violating program rules.
- Food Outlays—using the food outlay quartile concept, after careful examination, the top three quartiles were combined, creating a two category variable. In addition to being an

important stratification variable, this variable captured a dimension relating to the overall level of WIC business.

After examining all interactions and eliminating those with sparse within-cell information, we decided to focus on specific contrasts rather than all the elements that would be included in the interaction terms. Contrast terms that identify vendors by specific combinations relate to interactions using store type, ownership, urbanization, and poverty and food outlays. The contrasts explored included:

- Large privately owned vendors (including those we were unable to classify) versus others vendors.
- Large publicly owned vendors versus others vendors.
- Small vendors in highly urbanized locations versus other vendors.
- Small vendors in rural areas versus other vendors.
- Vendors in poor rural areas versus other vendors.
- Vendors in highly urban, high poverty areas versus other vendors.
- Large, privately owned vendors located outside highly urban areas versus other vendors.
- Publicly owned vendors with low food outlays in high poverty areas versus other vendors.

In addition to those terms defined by the raking variables, we introduced two other variables representing vendor specific variables.

- Risk—this binary variable distinguishes high risk vendors from other vendors. Risk was determined by the State according to the methodologies used to define vendors that should be investigated.[27]
- New Vendor Status—vendors are classified as new vendors if they are authorized in the fiscal year associated with the TIP file. The theory is that these vendors should have a higher overcharge rate due to their lack of familiarity with program rules. However, as we found in our previous analyses, they actually have a lower rate. The reason for this

[27] This risk variable is certainly associated with the overcharge outcome, but not perfectly aligned with outcomes. We would expect it to be an important predictor of outcomes, and would certainly increase the overall explanatory value of the model. And since not all high risk vendors are investigated, nor do they demonstrate overcharge violations, this factor can be used to predict the probability of an overcharge for these vendors.

lower overcharge rate reflects a lag between the initiation of the investigation and the sanction, in this case a finding of an overcharge. Thus, the initiation occurs when they are classified in TIP as new, and the outcome, in the majority of cases, occurs in the following year, when they are no longer classified as new. This issue, however, should not prevent us from using this variable since the lag helps explain the delay that seems to be important in deciding differences between the sample and the population.

Four State-level variables were also considered. These variables sought to characterize differences among States in their vendor management policies.[28] The variables included.

- Monitoring Rate—this predictor is the number of vendors as a percent of the vendor population in that State that received at least one monitoring visit.
- Percentage of High Risk Vendors—this predictor is a State-level variable that represents the number of high risk vendors over the total number of vendors authorized by the State.
- Investigation Rate—this predictor represents the number of total completed investigations conducted in the State over the number of vendors authorized by that State.
- Violation Rate—this predictor represents the number of vendors sanctioned over the number of investigations conducted. This predictor may reflect the degree to which States differ in the investigative techniques and thus in outcomes. Although considered, it was not used in the final model.

We modeled two conceptually different indicators of overcharge violations.

- All Investigations—this indicator reflects our interest in obtaining estimates consistent with the raking estimates that have been reported previously. In any TIP fiscal year file, there are instances of completed investigations, initiated investigations, and on-going investigations. Although we might expect that only completed investigations would be resolved, sanctions have been issued for cases marked as initiated or ongoing as well,

[28] Another variable that might be of value reflects whether the State transacts its WIC FIs through paper or EBT instruments. There are a handful of States with EBT implemented statewide, and a few where EBT is implemented in a few counties. These systems require electronic processing at the point-of-purchase and therefore purchases have to be reconciled with store and WIC electronic databases. This leaves little room for overcharging. Thus, we might expect that the occurrence of mischarges to be substantially reduced or eliminated in the presence of EBT. Although not investigated here, it would be useful to examine the predictive value of EBT on overcharges in future studies.

although at a lower rate than that observed for completed observations. Part of the issue here is that TIP does not provide a case tracking system, and thus the logical progression of case through an initiation, ongoing and completion sequence does not occur all the time. In addition, some vendors seem to be subject to two investigations within a year, thus leading to the possibility that one investigation could be completed with a result but marked as initiated if a second investigation was initiated within the same year.

- Completed Investigations—this indicator reflects our interest in only those cases that should have been resolved. The use of this indicator eliminates some of the ambiguity associated with how investigations are coded in TIP.

To obtain final estimates we used Bayesian estimation techniques based on Monte Carlo Markov Chain (MCMC) sampling. The advantage of this Bayesian approach was that it allowed us (1) to simulate multiple forms of the regression to be used in the prediction, which is useful in characterizing the variance characteristics of the estimates, (2) to pursue a model that could easily accommodate State-level affects in the form of a hierarchical or multi-level model, (3) to express our certainty/uncertainty with the parameter estimates through credibility intervals on the posterior distribution.

Exhibit 6 provides the results of using a Bayesian logistic analysis of the probability of an overcharge. The Bayesian procedure sampled the database (TIP investigations) repeatedly and created 10,000 sets of predictor results (the posterior distribution). Each predictor was allowed to vary under the assumption that they were normally distributed with zero mean and variance of 1000.[29] The 10,000 samples were then averaged to create mean values for the parameters, and credibility limits for the parameters at the 5th and 95th percentiles.[30] Negative parameter values generally show a decrease in the probability of a violation when compared to the intercept and positive values show a higher probability of a violation. The analysis does not evaluate significance (as would be the case in the non-Bayesian situation) but is examined by our

[29] We view this as a relatively large variance, but saw it as a way to express our uncertainty with the process, and also to allow the production of a wider range of models that can be used in the prediction. The specification of priors, however, should be looked at more closely if this approach is adopted over the raking process.
[30] Credibility limits are interpreted as a probability Statement that the violation rate will be within the values, a different interpretation than that given by confidence intervals.

confidence in the values as indicated by the credibility limits. If the two credibility values cross zero, there is a particular probability that the parameter could be equal to zero.

The results (using completed cases) indicate that the factors that decrease the probability of an overcharge relative to the mean include:

- Public ownership
- New Vendor Status
- Being Located in Poor Rural Areas
- Being Large and Private in Rural Areas

The factors that increase the probability of an overcharge relative to the mean include:

- Being a small store
- Being in a large urban area
- Being in an area with a greater level of poverty
- Having a low level of food outlays
- Being defined as a high risk vendor
- Being large and privately owned
- Being small and located in a rural area
- Being publicly owned with low levels of food outlays and in low poverty areas
- Being located in a store with a high rate of monitoring activity
- Being located in a State with a high percentage of high risk vendors
- Being located in a State with a high investigation rate.

It should be noted that the form and predictors for the models focusing on all cases differ are similar in direction from those that include completed cases, although the parameters vary.

The regression specified the State-level variables as regressors in the final model at the vendor level. We attempted to examine intercept differences among the States and slope differences in the regressors but failed to obtain an adequately performing model. Despite this, there seemed from our initial efforts to be some potential in developing a hierarchical model.

| Exhibit 6. Modeling Results on Overcharge Violations for Completed Investigations and All Investigations Using 2010 TIP Data | | | | | | |
|---|---|---|---|---|---|---|
| **Variable** | **Model Using Completed Cases** | | | **Model Using All Cases** | | |
| | **Parameter Value** | **5th Credibility Limit** | **95th Credibility Limit** | **Parameter Value** | **5th Credibility Limit** | **95th Credibility Limit** |
| **Intercept** | -9.1197 | -12.9412 | -5.3754 | -11.5810 | -12.6740 | -10.4681 |
| **Ownership (Public Ownership)** | -0.7858 | -2.3009 | 0.5028 | -0.9423 | -1.8435 | -0.0435 |
| **Urban Population as a Percentage of Total Population** | 1.0820 | 0.2338 | 1.9312 | 1.4062 | 0.9084 | 1.9126 |
| **Poverty Level Households as a Percentage of All Households** | 2.6121 | 1.1533 | 4.0928 | 0.9263 | 0.0736 | 1.7992 |
| **Store Type (Small Vendors)** | 1.5683 | -0.2917 | 3.4516 | 2.0782 | 1.5512 | 2.6051 |
| **Lowest Food Outlays** | 0.2341 | 0.0014 | 0.4620 | * | * | * |
| **High Risk** | 1.7276 | 1.2719 | 2.2230 | 3.4232 | 3.0105 | 3.8619 |
| **New Vendor** | -0.9729 | -1.5219 | -0.4521 | -3.0192 | -3.4997 | -2.5656 |
| **Large Privately Owned Vendors** | 1.4447 | -0.4133 | 3.2871 | * | * | * |
| **Small Vendors in Rural Areas** | 0.7303 | 0.1080 | 1.3999 | * | * | * |
| **Vendors in Poor Rural Areas** | -0.4623 | -0.8243 | -0.1088 | * | * | * |
| **Vendors with Low Levels of Food Outlays in Rural Areas** | * | * | * | 0.3111 | 0.0759 | 0.5442 |
| **Large Vendors With Large Food Outlay Levels** | * | * | * | 0.4104 | -0.0282 | 0.8700 |
| **Large Privately Owned Vendors in Rural Areas** | -1.4967 | -2.2013 | -0.8381 | 1.2772 | 0.7378 | 1.8342 |
| **Publicly Owned Vendors with Low Levels of Food Outlays the Lowest Poverty Areas** | 2.0425 | -0.2041 | 3.9846 | 2.1687 | 0.2227 | 3.9705 |
| **State Level Monitoring Rate** | 0.3385 | -0.1794 | 0.8622 | 2.1476 | 1.5808 | 2.7352 |
| **Percentage of High Risk Retailers Identified in State** | 1.4610 | 0.7656 | 2.1707 | 2.9980 | 2.4929 | 3.4889 |
| **Percentage of Vendors Investigated in State** | 2.4703 | 1.6823 | 3.2624 | 0.3111 | 0.0759 | 0.5442 |

*Not included in Model
Source: Annual Measure of WIC Erroneous Payments to Vendors: 2011 Methodology Report.

Next we sought to apply the models to the 2009 data in hope of determining how well overcharge violations in that year are modeled. The results are displayed in Exhibit 7. The data show that the difference between the predictions and actual number of violations is about 2.6 percentage points (or a difference of about 12 percent) for completed investigations and 1.0 percentage point for all investigations (or a difference of about 6 percent). Thus, the model will tend to overestimate the number of violations in the population, at least for 2009.[31] Our estimates were not terribly off, but the results suggest that the generation of separate models should be considered for each year, or at the very least, the year to year variation in the model ought to be explored through a hierarchical approach. As was noted, although we attempted this approach, the initial results were not satisfactory and an extended effort is needed to develop models that can be used for over time estimates.

**Exhibit 7**

**Comparison of Predictions Generated by the Violations Model And the Actual Number of Overcharge Violations in 2009.**

| | **Percentage of Vendors Identified As Overcharging** | |
|---|---|---|
| | **Actual Number of Violations in TIP** | **Predicted by Model** |
| Completed Investigations | 19.62% | 22.27% |
| All Investigations | 14.17% | 15.15% |

Source: Annual Measure of WIC Erroneous Payments to Vendors: 2011 Methodology Report.

To obtain the actual values for the population, the models were applied to each vendor in the population using the vendor's characteristics. Using MCMC sampling processes, we generated 25,000 models. From these 25,000 sampled models, we randomly selected, with replacement, 200 models for each vendor that generated 200 different overcharge probabilities for each vendor. These probabilities allowed us to compute dispersion characteristics for the sample as a whole and for any set of retailer characteristics.

These individual level probabilities were summed to obtain the total number of vendors overcharging, the overall percentage of vendors overcharging, and the 5th and 95th credibility limits on the overcharge amounts. These are presented in Exhibit 8, which compares the values reported in 2009 and 2010 using the raking approach to the values predicted through the models.

[31] The model performance against the 2010 data set could be accomplished by comparing the distribution of probabilities for each of the 200 samples drawn from using the posterior predictor models to determine how well they fit the actual values provided by TIP. Another approach for examining the viability of the model would have been to estimate the model on a subset of investigations, and then to predict the values from the resulting models on the "hold-out" sample.

| Exhibit 8.<br><br>Comparison of Overcharge Rates (Percentage of Vendors Overcharging) Using the Standard Definition, and the Alternative Models | | | | | | |
|---|---|---|---|---|---|---|
| Year | 2009 | | | 2010 | | |
| | Generated Through Raking Process | Generated Through Regression Approach Using All Investigations | Generated Through Regression Approach Using Completed Investigations | Generated Through Raking Process | Generated Through Regression Approach Using All Investigations | Generated Through Regression Approach Using Completed Investigations |
| Estimated Number of Vendors Overcharging | 3,885 | 1,281 | 2,271 | 3,524 | 1,296 | 2,245 |
| Percent of Vendors Overcharging | 9.34% | 3.08% | 5.46% | 8.26% | 3.04% | 5.26% |
| 5th Percentile Limit | 7.91% | .00% | .09% | 7.09% | 0.00% | 0.9% |
| 95th Percentile Limit | 10.73% | 17.3% | 28.96% | 9.54% | 18.4% | 25.2% |

Source: Annual Measure of WIC Erroneous Payments to Vendors: 2011 Methodology Report.

There are three results of note.

- First, the overcharge violation rate using all investigations is about one-third of the ones generated by the raking procedure, while the one using completed investigations is about 60 to 67 percent of the one derived from raking. This means that the regression approach adjusts the investigative sample to the population to a larger degree than the raking procedure. We speculate that this reflects the introduction of the State-level variables into the regression. In other words, we know that some States have a high investigation and overcharge violation rate (New York) and other States have very little activity. The introduction of the State-level predictors essentially creates a blended rate, shrinking New York's higher rate toward the very low rates for other States. The raking approach, however, ignores State-level variables and thus does not adjust the values experienced in States like New York to the general population. In comparing the rates, it should be noted that neither is correct or incorrect—they reflect two different assumptions about the propensity of vendors to overcharge. One assumption would suggest that vendors in New York can be used to represent overcharging nationally (per the raking approach). On the other hand a blended rate would suggest that vendor overcharging varies considerably among States (regardless of their investigation activity) and New York's experience should be averaged with these other States (per regression approach). The beauty of the regression approach is that it is more flexible in terms of adding State-level and other predictors.

- Second, the differences between FY2009 and FY2010 estimates are not only in the same direction for all three sets of estimates, but they are roughly the same magnitude of change. Thus, if our interest is in the change in overcharging rather than the actual magnitude, all three approaches seem to give roughly the same results.
- Third, the regression results provide a much wider range of credibility than the raking estimates. This difference belies a fundamental difference between the raking and regression approaches. Whereas the intervals for the sampling process in the raking approach reflect variation in the weights due to repeatedly sampling from investigations, the sampling process for the regression approach focuses on variation in the probability of a violation. Another reason for the difference in credibility statistics posed by the two approaches may reflect the relatively large number of New York cases. In other words, the more the investigative sample is dominated by a State, the more it is likely to exhibit a smaller range of confidence of credibility, and this is what the raking approach does. The elimination of a State such as New York would first of all shift the mean estimate and deviance statistic toward the results manifested in States other than New York. This raises a question of whether the overcharge behavior found in other States is typical of what happens when WIC participants (not investigators) purchase WIC goods, or whether what occurs in New York State actually represents overcharge behavior. Our approach allows for integrating this uncertainty into the modeling process. However, more research is needed to fully understand the differences.

### 4.2.5 **Generation of Food Outlay Overpayments**

The critical estimate produced by the WIC Erroneous Payment Study is the total food outlays due to overcharges. The next step then is to generate a value that represents the total food outlays and percentage of food outlays that were overcharged over all the vendors in the population. For all WIC purchases for a specific vendor the overpayment is equivalent to:

$$\sum_{i=1}^{n} po_i o_i af_i \qquad (5)$$

Where

$po_i$ is the probability that the vendor overcharged for purchase i,

$o_i$ is the purchase amount of the items on the food instrument for purchase i

$af_i$ is the adjustment factor or the percentage of the total purchase amount that is overcharged for purchase i.

As indicated in our discussion of raking, TIP provides no information on the amount of dollar overcharged and no way to determine an adjustment factor. This value has to be estimated from the 2005 WIC Vendor Management Study. This percentage of food outlays overcharged (af) is represented as:

$$Percentage\ of\ Outlays\ Overcharged = \frac{Error}{Total\ Cost\ of\ Food\ on\ the\ Food\ Instrument}$$

Where the "error" is the overcharge recorded for violating vendors surveyed in the 2005 WIC Vendor Management Study and "total cost of food on the food instrument" pertains to the instance in which the overcharge occurred. As in our discussion of raking, the average amount of an overcharge for a purchase estimated in the 2005 WIC Vendor Management Study was about 10 percent of the face value of the items on the food instrument. Instead of using a fixed percentage as was done under the raking scenario, it was decided to model the food outlay percentage.

We first attempted to model the percentage as a function of those predictors used to identify the probability of a violation, but found that none of the predictors were significant. This meant that the amount and percentage of the overcharge could be represented as a random variable that varied about the mean percentage. The only critical predictor was whether the buy made within the study was a safe or partial buy. It was decided to generate two models for depicting the percentage of food outlays overcharged, one for safe buys only, and one for safe and partial buys.[32] This modeling was accomplished through a Bayesian MCMC approach that used the

[32] The estimate using a combination of safe and partial buys assumes that the proportion of each of these buys is equivalent. The actual proportion in the population is unknown. Previous estimates have assumed that only safe buys are made.

beta distribution as a prior to depict the average percentage of overcharge.[33] The results, along with the corresponding information from the 2005 WIC Vendor Management Study, are presented in Exhibit 9. The distributional statistics in this exhibit show that the percentage estimate is very close to the mean derived from the WIC Vendor Management study, and although there is a difference in standard deviation, that may be attributable to the small sample size for the 2005 bookend study. The ranges as indicated in the last two rows indicated an acceptable correspondence between the highest values obtained from the MCMC results and the WIC Vendor Management Study results.

**Exhibit 9.**

**Results Comparison the Percent Error Statistics Derived From the WIC Vendor Management Study and Bayesian Estimates Using a Beta Prior**

| | Safe Buys | | Safe and Partial Buys | |
|---|---|---|---|---|
| | **2005 WIC Vendor Management Study** | **MCMC Results Using Beta (.65,6 ) As Prior** | **2005 WIC Vendor Management Study** | **MCMC Results Using Beta (1.1,6 ) As Prior** |
| Mean | .097 | .097 | .155 | .158 |
| Std. Deviation | .108 | .151 | .127 | .193 |
| $5^{th}$ Percentile | .002 | .020 | .010 | .002 |
| $95^{th}$ Percentile | .322 | .347 | .400 | .567 |

Source: Annual Measure of WIC Erroneous Payments to Vendors: 2011 Methodology Report

The MCMC produced 25,000 values, and each vendor was randomly assigned, with replacement, 200 of the values. The total food outlays associated with overcharges or overpayment is represented by the formula below:

$$total\ overpayment = \sum_{i=1}^{n} po_i o_i af_i$$

Where $po_i$ is the probability of an overcharge for retailer i, and is generated as we specified in Section 4.2.2. Food outlays $o_i$ is the total food outlays recorded for retailer i, and is provided by TIP. The product of the three values is summed over all retailers and across all of the 200 values generated for the retailers and are presented in Exhibit 10.

[33] The beta distribution is represented by two parameters (a location and a shape parameter), and can take a variety of forms producing values between 0 and 1. The forms that we were attempting to estimate were those that resembled a highly left skewed distribution with the mean centered at about 10.

There are several results to note in examining the dollar food outlay figures.

- The actual amounts generated by the regression equations are smaller than those generated by the raking approach. The magnitude of these differences mirror the results generated for vendor-based violation rates. That is, they are from forty to sixty percent of the raking values. Even when the adjustment factors using safe and partial buys are used, the amounts and rates are less or the same as those obtained by raking. We again reference the regression approach use of State-level variables to adjust for the proportionately high level of violations in several States.
- The ranges generated by the regression approach are also more dispersed than those generated by the raking approach. It should be noted that the values still show a relatively low level (less than 3 percent) rate of overcharging relative to total WIC food outlays.

**Exhibit 10.**

**Overcharged Food Outlays for the Vendor Population By Year, Method and Investigation Status**

| Overcharge Measure | Type of Covert Buy Used for Generating Estimates | 2009 | | | 2010 | | |
|---|---|---|---|---|---|---|---|
| | | Raking Process | Regression Approach Using All Investigations | Regression Approach Using Completed Investigations | Raking Process | Regression Approach Using All Investigations | Regression Approach Using Completed Investigations |
| Estimated Amount (in thousands) | Safe Buy Only | $24,706 | $8,729 | $15,312 | $30,128 | $11,855 | $18,815 |
| | Safe and Partial Buy | * | $14,175 | $24,925 | * | $19,242 | $30,589 |
| Percent of Food Outlays Overcharged | Safe Buy Only | 0.86% | 0.21% | 0.36% | 0.74% | 0.29% | 0.46% |
| | Safe and Partial Buy | * | 0.33% | 0.59% | * | 0.47% | 0.75% |
| $5^{th}$ Percentile Limit on Percent of Food Outlays Overcharged | Safe Buy Only | 0.64% | 0.00% | 0.00% | 0.53 | 0.00% | 0.00% |
| | Safe and Partial Buy | * | 0.00% | 0.00% | * | 0.00% | 0.00% |
| $95^{th}$ Percentile Limit on Percent of Food Outlays Overcharged | Safe Buy Only | 1.13% | 0.69% | 1.51% | 0.97 | 0.91% | 1.74% |
| | Safe and Partial Buy | * | 1.30% | 2.54% | * | 1.67% | 2.93% |

Source: Annual Measure of WIC Erroneous Payments to Vendors: 2011 Methodology Report

### 4.2.6 **Summary of Discussion on the Alternative Approach**

The alternative approach for estimating WIC overpayments provided a basis for a modifying the raking approach that has historically provided overpayment estimates. In general, at least from looking at overcharges in trend terms, estimates made from the alternate approach tracks with estimates made from the raking process. It has several advantages over the raking approach, which include:

- The incorporation of factors other than those that could reasonably be incorporated in the raking approach. Thus, in addition to the five factors used in raking, it adds five additional variables that express differences between investigated stores and the population both in terms of their own characteristics, and in terms of State practices.
- It eliminates estimates made on predictors that are not well represented in the population, thus decreasing the differential values of the weights obtained from the raking process.
- It provides a theory-based approach to identifying error, as opposed to the mechanical approach used by raking, and in doing so, provides an approach that can be validated and that will provide better predictive quality.

The approach, although providing a proof of concept, needs to be further examined to address the following concerns.

- Further work to specify a model would include the following. First, there is the question about whether to include factors that would appear under a frequentist perspective to be equal to zero. The Bayesian approach does not require elimination of these factors under the rationale that zero is possible, but that other values, positive and negative are also possible and should be included in the prediction equation (especially when multiple draws are used to extract a posterior distribution). The question here is when these additional factors lead to adverse consequences, such as an overfitted model.

  The second task would entail the introduction a random effects in a hierarchical fashion. We had to abandon this line of inquiry because of convergence issues, but we believe that a more concerted effort could solve this problem. This would lead to being able to more precisely measure the difference in rates among States and their effect on overall estimates. The introduction of State-level variables showed promise in this regard.

The third task is to develop separate models for years either through treating year as a hierarchical variable within a comprehensive model, or by generating separate models for each year. The results in this report show that a model based on 2010 data was an acceptable but not optimal fit for modeling the 2009 data.

- Further work on the assumptions. In developing the model using Bayesian perspectives, we made a number of assumptions that need to be subjected to sensitivity testing.

  First, the assumption concerning the use of the normal distribution in the logistic regression predicting overcharge violation needs to be examined. Another distribution such as the inverse Chaucey distribution may be more appropriate, and the variance specified with the normal distribution (1,000 units) may be too large.

  Second, the assumption concerning the use the beta distribution for predicting the amount of the typical overcharge should be looked at relative to other distributions that might provide better fit with the actual data.

- Finally, there is the question about the correctness of the overcharge estimates. The raking approach provided a higher rate than the alternative method. We speculated that was because the raking approach did not account for State differences and therefore allowed a set of overly represented States to dominate in determining the estimates. The question therefore is whether the estimate is a result of State vendor management policy or the result of actual overcharging. Thus, if we compare two States—one with an aggressive policy and one without, would the respective violation rate differences a result from inherent differences between the behavior of the vendors in the two States, or from the behavior of State in investigating and pursuing cases? Addressing this question may mean that overcharge rates should be qualified by the State-level vendor policies of States and also by the presence of EBT systems, which are expected to reduce overcharging.

There is also a matter of how to deal with the anomalous finding from the 2005 WIC Vendor Management Study that estimates undercharges to be greater than overcharges. The Bayesian approach allows two ways to deal with this finding. First, we can use the data but place a lower level of credibility on it. This would probably involve some modeling that looks at the relationships between under- and overcharging. Second, we can pool the 2005 data with previous efforts to provide an overall result that is more consistent with theory.

# 5. ESTIMATION OF UNDERCHARGES

The 2005 bookend study defined an undercharge as a negative difference between the redeemed value of a food instrument and the best retail price for the food bundle as recorded by field data collectors. The 2005–2009 update studies also used this definition. Unlike overcharges, undercharges seem not to be an area of investigative interest for the States and therefore no information exists on such activity in TIP. Therefore, to obtain updated undercharge estimates, the probability of a vendor transacting an undercharge and the dollar amount of the undercharge were estimated using the 2005 bookend study (the only source on undercharges) through regression models and applied to the TIP data. This means that for generating estimates historically consistent with previous estimates, the total expected value of undercharges will change strictly as a function of changes in food outlay dollar amounts and the characteristics of the population of WIC vendors. It should be noted that in the alternative approach that we propose in Section 5.2, we alter the specifications of the model to include overcharges as a predictor, which makes use of TIP investigative outcomes.

## 5.1 METHODOLOGY FOR ESTIMATING UNDERCHARGES

In this section, we describe the approach for generating an undercharge estimate consistent with estimates presented between FY2005-FY2010. As was indicated, the estimates are solely based on the 2005 WIC Vendor Management Study. That study allowed retailers to undercharge on any of three types of buys. As shown in Exhibit 11, the percentage of vendors undercharging on any one of the three buys is approximately 10 percent, which is equivalent to the result for overcharging when all three buy types are taken into consideration.

**Exhibit 11.**

**Weighted Frequency of Vendors With Undercharges, 2005 Bookend Study**

| Number of Undercharges | Number | Percent | Cumulative Number | Cumulative Percent |
|---|---|---|---|---|
| No undercharges | 33,318 | 89.71 | 33,318 | 89.71 |
| One undercharge | 3,384 | 9.11 | 36,702 | 98.83 |
| Two undercharges | 346 | 0.93 | 37,047 | 99.76 |
| Three undercharges | 90 | 0.24 | 37,138 | 100.00 |

Source: Annual Measure of WIC Erroneous Payments to Vendors: 2011 Methodology Report

The proportion of vendors undercharging by type of buy is presented in Exhibit 12. The data show that the percentage of vendors undercharging on partial buys was lower than that for other buys. Vendors were more likely to undercharge for major substitutions than they were for partial or safe buys.

**Exhibit 12.**

**Weighted Frequency of Undercharges in the 2005 Bookend Study, by Buy Type***

| Buy Type | Undercharge | | No Undercharge | | Total | |
|---|---|---|---|---|---|---|
| | Number | Percent | Number | Percent | Number | Percent |
| Safe | 1,554 | 4.6 | 32,289 | 95.4 | 33,843 | 100.0 |
| Partial | 971 | 2.9 | 32,681 | 97.1 | 33,651 | 100.0 |
| Minor substitution | 1,131 | 5.1 | 20,995 | 94.9 | 22,127 | 100.0 |
| Major substitution | 656 | 6.0 | 10,308 | 94.0 | 10,963 | 100.0 |
| Total | 4,312 | 4.3 | 96,273 | 95.7 | 100,585 | 100.0 |

* Numbers represent the weighted number of buys, not the number of vendors.
Source: Annual Measure of WIC Erroneous Payments to Vendors: 2011 Methodology Report

With regard to dollar amount, the average undercharge in a safe buy was $0.94 for vendors undercharging (see Exhibit 13). In a partial buy, it was $1.43; in a minor substitution, it was $2.41; and in a major substitution, it was $0.96. As opposed to overcharges, undercharges became larger when partial buys replaced safe buys.

**Exhibit 13.**

**Weighted Distribution of Undercharges in the 2005 Bookend Study, by Buy Type**

| Buy Type | No. of Buys | Average | Minimum | 25th Percentile | Median | 75th Percentile | Maximum |
|---|---|---|---|---|---|---|---|
| Safe | 74 | -$0.94 | -$5.43 | -$1.16 | -$0.49 | -$0.18 | -$0.01 |
| Partial | 40 | -$1.43 | -$9.00 | -$2.09 | -$0.60 | -$0.20 | -$0.01 |
| Minor Substitution | 51 | -$2.41 | -$14.67 | -$3.00 | -$1.20 | -$0.40 | -$0.01 |
| Major Substitution | 23 | -$0.96 | -$3.00 | -$1.42 | -$0.50 | -$0.23 | -$0.02 |

Source: Annual Measure of WIC Erroneous Payments to Vendors: 2011 Methodology Report.

In Exhibit 14, undercharges ranged from 5.5 percent (major substitutions) to almost 12 percent (partial buys and minor substitutions) of the total value of the food instrument, which supports the claim that undercharges vary with the type of interaction that WIC participants have with WIC vendors. However, because the relative frequency of the natural occurrence of buy types cannot be determined and because these estimates are meant to build on the 2005 bookend study results, only safe buys were used to generate estimates of undercharges.

**Exhibit 14.**

**Weighted Distribution of Undercharges as a Percentage of Food Instrument Value in the 2005 Bookend Study, by Buy Type**

| Buy Type | No. of Buys | Mean Percentage | Minimum Percentage | 25th Percentile | Median | 75th Percentile | Maximum Percentage |
|---|---|---|---|---|---|---|---|
| Safe | 74 | 7.211 | 0.098 | 1.147 | 3.511 | 7.567 | 46.530 |
| Partial | 40 | 11.786 | 0.072 | 1.715 | 6.834 | 13.599 | 91.667 |
| Minor substitution | 51 | 11.759 | 0.031 | 1.105 | 6.651 | 16.534 | 71.030 |
| Major substitution | 23 | 5.483 | 0.314 | 1.401 | 3.840 | 8.186 | 25.063 |

Source: Annual Measure of WIC Erroneous Payments to Vendors: 2011 Methodology Report

Because the TIP files do not contain any information about undercharges, any estimate must be based solely on the undercharge behavior of vendors sampled for the 2005 bookend study, the only source of information on undercharges, as applied to the current TIP population. Since the 2005 bookend study provided the official improper payment estimates based only on safe buys, our approach involved developing predictive equations based on behaviors revealed in safe buys only.[34] In developing a predictive equation, logistic regression was used to model the probability

[34] Although safe buys were used in developing the model primarily for establishing consistency with the bookend study, the initial analysis conducted for the 2005 estimates included investigations into the use of other types of purchase types (i.e., partial buys and substitution buys). That analysis indicated that had partial and substitution buys been used, the results would have been somewhat different. One of the essential issues associated with combining the various types of buys, however, is that the natural occurrence of these buys in the population is not known. Therefore although it is possible to model each type of buy, developing a model for all three types of buys involves some tenuous assumptions.

of a vendor undercharge, and ordinary least squares regression techniques were used to model the amount of an undercharge.

The first step was to predict the probability of an undercharge. A predictive equation using a logit model was generated from the weighted 2005 bookend study sample. Because it is the probability of undercharging that is modeled at this stage, logistic regression is appropriate because it is nonlinear, allowing the modeler to take into account the fact that probabilities are bounded by 0 and 1. The vendor characteristics used as predictors include the following:

- Vendor type, expressed as a series of nominal variables, one each for large retail vendors, small retail vendors, and WIC-only vendors and an indicator for all other types of vendors. It should be noted that the 2005 bookend study did not include pharmacies that only provided special formulas and medical foods,[35] commissaries, direct vendors, or home delivery vendors in its sample. As a result, the indicator for all other types of vendors was necessarily estimated based on WIC above-50-percent vendors only.
- Ownership type, either public or private.
- Percentage of families within the vendor's ZIP Code living in a U.S. Census Bureau designated urban setting.
- Percentage of households within the vendor's ZIP Code living at or below the poverty level.
- Vendor's total annual WIC food outlay dollars in 2005.

Next, the logistic regressions, as estimated, were applied to all vendors in the TIP file, and the resulting log odds ratios were converted to probabilities. The equation that was applied is specified as:

$$P_v=\mathbf{1}/(\mathbf{1} + \exp(-(-\mathbf{1.8174} + \mathbf{0.0598}*U_v + \mathbf{1.5633}*PO_v - \mathbf{3.54}*(\mathbf{1/10^7})*R_v - \mathbf{1.6523}*LR_v - \mathbf{1.2922}*SR_v - \mathbf{0.4434}*WO_v - \mathbf{0.0475}*PU_v + \mathbf{0.0835}*PR_v)))$$

[35] Because the focus was on food outlays, it was difficult on a store-by-store basis to isolate formula sales from food outlay sales. We made a decision to exclude pharmacies because most would sell formula, and although some would sell food, they would probably account for a small portion of overall food sales.

Where: $P_v$ is the probability that the vendor undercharged;

$U_v$ is the percentage of the population living in urban areas within the vendor's ZIP Code;

$PO_v$ is the percentage of households living in poverty within the vendor's ZIP Code;

$R_v$ is the annual amount of food outlays for that vendor;

$LR_v$ is whether the vendor is a large retailer;

$SR_v$ is whether the vendor is a small retailer;

$WO_v$ is whether the vendor is a WIC-only store;

$PU_v$ is whether the vendor is publicly owned; and

$PR_v$ is whether the vendor is privately owned.

The second step was to predict the expected dollar value of an undercharge. Linear regression was appropriate because the predicted (dependent) variable is continuous, and unlike probabilities there was no reason to expect a nonlinear relationship. The regression used only those cases of undercharging in the estimation procedure. Therefore, it provided the amount of the average undercharge, given certain vendor characteristics, if the vendor undercharged.

The prediction equation is specified as:

$$EU_v = \mathbf{0.07302} - \mathbf{0.01322}*U_v - \mathbf{0.20337}*PO_v + \mathbf{2.496827}*(\mathbf{1/10^8})*R_v + \mathbf{0.04108}*LR_v + \mathbf{0.06282}*SR_v + \mathbf{0.03089}*WO_v - \mathbf{0.00542}*PU_v$$

Where: $EU_v$ is the expected amount of underpayments given that the vendor undercharged;

$U_v$ is the percentage of the population living in urban areas within the vendor's ZIP Code;

$PO_v$ is the percentage of households living in poverty within the vendor's ZIP Code;

$R_v$ is the annual amount of food outlays for that vendor;

$LR_v$ is whether the vendor is a large retailer;

$SR_v$ is whether the vendor is a small retailer;

$WO_v$ is whether the vendor is a WIC-only store; and

$PU_v$ is whether the vendor is publicly owned.

These predictive equations were applied to all vendors in the TIP file. Again, all values were predicted for each vendor using the parameters estimated based on safe buys. When predicting from the TIP file, total food outlay dollars were substituted for the value of the food instrument that was used when generating the equation from the 2005 bookend study data.

The last step was to obtain the expected amount of an undercharge for each vendor in the TIP file. Multiplying the probability of undercharging (step 1) by the average amount undercharged (step 2) produced an expected value for undercharges for each vendor. This value represents the total dollar amount undercharged. This is represented as:

$$AU_v = R_v * P_v * EU_v$$

Where $AU_v$ is the final adjusted undercharge for vendor v, and the other factors are as defined above. The vendor undercharge rate was calculated by summing the probabilities of undercharging across all vendors in the TIP file, and the food outlay undercharge rate was calculated by determining the total amount of undercharges as a percentage of all food outlay dollars.

## 5.2 ALTERNATE APPROACHES FOR ESTIMATING UNDERCHARGES

### 5.2.1 Introduction

As described above, underpayments were estimated using a three-step process, using regression procedures. In the original estimating equations, first developed for the 2005 estimates, the effort was limited to variables common to both the WIC Vendor Management Study and TIP. In examining the equations that have been used since the 2005 estimate, we noted three possibilities for improvement.

- First, the model specifications could be reformulated to focus on variables that have more relevance than those used previously. In particular, variables such as risk designation or new vendor status could be used in the estimating equations.

- Second, all undercharges could be examined, not just those associated with safe buys. The rationale behind this would be that partial buys and substitution buys would provide a larger number of cases from which to estimate the probability of an undercharge and the overpayments would reflect a more realistic assessment of the type of buys WIC participants make.
- Third, a term could be added that would express whether, in any of three buys, the vendor overcharged. The rationale for the inclusion of this variable is based on the perceived lack of economic incentive, on the part of the vendor, to undercharge. Undercharging would therefore be expected to occur as a random event reflecting a lack of particular controls regarding charging a consistent price. If it is a random event, we would expect that overcharges would occur with about the same frequency as undercharges. Including overcharges as an additional variable could provide information on the extent to which this occurs.

The same Bayesian MCMC approach that was described for overcharges was used for estimating undercharges. This consisted of:

- Estimating the probability that a vendor will have undercharged based on a model derived from the 2005 WIC Vendor Management Study.
- Applying the model for predicting the probability of undercharging to each WIC vendor.
- Summing the probabilities to obtain an estimate of the number and percentage of vendors who are undercharging.
- Estimating the average proportion of purchases that are undercharged.
- Developing an estimate of the undercharging rate in dollars for each vendor.
- Summing the adjusted food outlays to obtain an overall estimate of the amount and percentage of food outlays that are undercharged.

### 5.2.2 Estimating the Incidence and Amount of Undercharging

We first recast the model used in previous studies (Section 5.1) by considering the variables in the alternative overcharge model. In addition, we added an "overcharge" variable.

The inclusion of an overcharge predictor, however, shifts the emphasis from predicting the probability of an undercharge on a safe buy to the probability of an undercharge across all the buys at each vendor. The 2005 WIC Management Study was based on two to three covert purchases at each of the vendors in the sample. The first purchase was a safe buy, the second was a partial buy and the third was a substitution buy. Over- and undercharges could occur during any purchase. Thus, under- and overcharge indicators can be defined as occurring if one of these events occurred during any of the three purchases. For example a vendor can be identified as overcharging and undercharging only if they record one instance of each. The major issue with this formulation is that each purchase encompasses a different type of purchase and therefore clouds the issue as to whether the undercharge and overcharge behavior reflects the purchase type or the general behavior of the retailer. With this caveat in mind, overcharges were included primarily because they showed considerable promise in the 2010 analyses.

After preliminary analysis and elimination of potential predictors, the following were selected for the model:

- Store Type—this predictor distinguishes between large and small stores, with the former being more likely to have scanning and other electronic equipment that would limit store checkout errors.
- Food Outlays—this predictor distinguishes between stores with a low level of WIC food outlays and those with moderate to high levels.
- Overcharge Indicator—this final predictor was an indicator representing whether the vendor had overcharged in any of three “buys”.

Exhibit 15 presents the results of a model regressing the incident of an undercharge on the above variables. The independent variable in this case was the occurrence of a safe buy. An MCMC sampling approach with a logistic link function was used for generating 10,000 samples using a normal prior distribution for the parameters with mean equals to zero and variance equal to 100.

**Exhibit 15.**

**Modeling Results on Undercharge Violations for Completed Investigations and All Investigations Using 2010 TIP Data**

| Variable | All Investigations | | | Completed Investigations | | |
|---|---|---|---|---|---|---|
| | Parameter Value | 5th Credibility Limit | 95th Credibility Limit | Parameter Value | 5th Credibility Limit | 95th Credibility Limit |
| **Intercept** | -2.9480 | -3.4610 | -2.4382 | -4.2530 | -5.0273 | -3.4799 |
| **Store Type (Small Vendors)** | 0.5784 | 0.2351 | 0.9236 | 1.0206 | 0.5303 | 1.5170 |
| **Lowest Food Outlays** | -0.6502 | -1.2404 | -0.0560 | -1.2114 | -2.3361 | -0.1768 |
| **Overcharged (Yes)** | 0.7249 | 0.2671 | 1.1670 | 0.1928 | -0.5481 | 0.9454 |

Source: Annual Measure of WIC Erroneous Payments to Vendors: 2011 Methodology Report

The results indicate that small vendors have a greater probability of undercharging, those vendors associated with the lowest amount of food outlays have a smaller probability of undercharging, and those which were observed to have overcharged have a higher probability of undercharging.

The second stage of providing an estimate is to generate an expected value of an undercharge for each vendor. The model parameters were applied to vendor characteristics to yield a probability of an undercharge. For each vendor, we sampled the posterior parameters 200 times to yield two hundred different probabilities. For the overcharge variable, we utilized the probability of an overcharge. The probability is the one that was estimated for each vendor using the overcharged equation described in exhibit 6. The probabilities were summed to yield the number of undercharging vendors and related statistics (Exhibit 16).

| Exhibit 16. Overall Undercharge Estimates by Year, and Method | | | | | | |
|---|---|---|---|---|---|---|
| | 2009 | | | 2010 | | |
| | Standard Approach | Alternative Model—All Investigations | Alternative Model—Completed Investigations | Standard Approach | Alternative Model—All Investigations | Alternative Model—Completed Investigations |
| Estimated Number of Vendors Overcharging | 2,034 | 1,840 | 1,850 | 2,071 | 1,965 | 1,979 |
| Percent of Vendors Overcharging | 4.89% | 4.42% | 4.45% | 4.85% | 4.62% | 4.64% |
| 5th Percentile Limit | 4.82% | 1.00% | 1.01% | 4.80% | 1.05% | 1.05% |
| 95th Percentile Limit | 5.00% | 10.80% | 10.9% | 4.91% | 11.07% | 11.15% |

Source: Annual Measure of WIC Erroneous Payments to Vendors: 2011 Methodology Report

The estimates provided in Exhibit 16 can be summarized as follows:

- The estimated number of vendors overcharging using the alternative approaches is relatively close to those generated through the standard approach. For 2009, the difference is 200 vendors, and for 2010, the difference is near 100 vendors.
- The percentages of vendors overcharging are relatively close as well, varying by less than a percentage point.
- The ranges for these percentages are different however, with the alternative designs showing a wider dispersion than the standard approach.
- In comparison with overcharges, the estimates are greater than those estimated from all cases, but less than those estimated from completed cases. This is a very different result from those results presented in the most recent estimates for erroneous payments. It suggests that overpayments and underpayments cancel each other out, and that vendors in general do not intentional reap financial payments from overcharging. However, the results do not go as far as the 2005 bookend study, which provide estimates in which underpayments exceed overpayments—a finding that is somewhat unexpected and different than that which was found in previous studies. Part of the issue here is the uncertainty related to overpayments relating to whether the estimates reflect different investigative practices on the part of States, or actual differences in how vendors behave across States. It also reflects the unexpected findings on the relationship of under- and overpayments in the bookend study.

### 5.2.3 **Undercharge Food Outlay Estimates**

In generating dollar estimates, our approach resembles the approach used for overcharges more than the approach described in Section 5.1, which used food outlays values as an explicit factor in a regression equation. Thus our alternative approach is to take the probability of undercharging and multiply it by food outlays, adjusted for the proportion of food outlays that are typically undercharged during a transaction.

This value is best modeled as a random variable; we estimated the mean as a random effect using a beta distribution. The modified approach for doing this was based on drawing about 10,000 samples using a beta distribution with determined central tendency and shape parameters. Two set of samples were drawn—one for safe buys only and one for safe buys and partial buys. The results show a relatively close approximate to the actual distribution of values derived from the 2005 WIC Vendor Management Study (Exhibit 17).

**Exhibit 17.**

**Results Comparison the Percent Error Statistics Derived From the WIC Vendor Management Study and Bayesian Estimates Using a Beta Prior**

| | Safe Buys | | Safe and Partial Buys | |
|---|---|---|---|---|
| | **2005 WIC Vendor Management Study** | **MCMC Results Using Beta (.5,5.5 ) As Prior** | **2005 WIC Vendor Management Study** | **MCMC Results Using Beta (0.6,6 ) As Prior** |
| Mean | .071 | .085 | .087 | .089 |
| Std. Deviation | .104 | .107 | .131 | .103 |
| 5$^{th}$ Percentile | .002 | .000 | .000 | .000 |
| 95$^{th}$ Percentile | .368 | .310 | .368 | .306 |

Source: Annual Measure of WIC Erroneous Payments to Vendors: 2011 Methodology Report

The final dollar based estimated were derived by multiplying the probability of undercharging times the adjusted amount of food outlays. The values are present in Exhibit 18.

| Exhibit 18. Estimates of Undercharged Food Outlays Among WIC Vendors by Year, Method, and Type of Buy | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | 2009 | | | 2010 | | |
| | | Generated Through Standard Approach | Generated By Alternative Model—All Investigations | Generated By Alternative Model—Completed Investigations | Generated Through Standard Approach | Generated By Alternative Model—All Investigations | Generated By Alternative Model—Completed Investigations |
| Total Amount of Undercharges (in thousands) | Safe Buy Only | $13,057 | $15,179 | $15,282 | $13,906 | $16,307 | $16,393 |
| | Safe and Partial Buy | * | $16,620 | $16,699 | * | $17,847 | 17,939 |
| Percent of Total Food Outlays Undercharged | Safe Buy Only | 0.31% | 0.36% | 0.36% | 0.34% | 0.40% | 0.40% |
| | Safe and Partial Buy | | 0.39% | 0.39% | | 0.44% | 0.44% |
| 5th Percentile Limit | Safe Buy Only | 0.30% | 0.00% | 0.00% | 0.32% | 0.00% | 0.00% |
| | Safe and Partial Buy | | 0.00% | 0.00% | | 0.00% | 0.00% |
| 95th Percentile Limit | Safe Buy Only | 0.32% | 1.59% | 1.60% | 0.35% | 1.70% | 1.71% |
| | Safe and Partial Buy | | 1.66% | 1.66% | | 1.78% | 1.78% |

Source: Annual Measure of WIC Erroneous Payments to Vendors: 2011 Methodology Report

The results indicate that dollar undercharges are about two to four million dollars more using the alternative approaches than for the standard approach currently in use. The direction and magnitude of the shift is comparable between the alternative and standard approaches. The percentage of vendors undercharging reflects this, although the total is relatively low (less than a half of a percentage point). The dispersion statistics show a wider range of potential error than the standard approach.

### 5.2.4 **Summary of Alternative Undercharge Approach**

The development of an alternative approach for estimating undercharges was based on the same approach as that developed for overcharges, except it used information from the 2005 WIC Vendor Management study as its main input. The final estimates were generally consistent with those generated by the previous approach, but showed greater dispersion. This greater dispersion reflected the introduction of more uncertainty through a specific of Bayesian priors rather than fixed parameter estimates used by the standard approach. It also introduced overcharges as a primary predictor of undercharges, under the assumption that if mischarging is a random situation, then undercharges should be matched by overcharges over the course of time.

The alternative model in this case serves as a successful proof of concept and has, for further development, all of the requirements specified in the summary for the overcharge estimates.

# 6. FOOD PACKAGE CHANGES IMPACT ON ERRONEOUS PAYMENTS

FNS required that State agencies implement a revised WIC food package by October 2009. This package included foods that have more appeal to individuals with different ethnic backgrounds, and it allowed for the purchase of fruits and vegetables. Although these additional foods still use the traditional food instrument that specifies the type and quantity of the product (and not the price, although a maximum allowable amount may be identified), the purchase of fruits and vegetables is facilitated through a separate dollar-denominated voucher. With these changes the potential for errors would be expected to increase, because the new packages may result in processing complications for the WIC retailer. For example, errors may result from the fact that the new fruit and vegetable benefit is dollar denominated and offered on a separate voucher, while traditional WIC vouchers are defined based on number and size of package or product weight. However, this change would probably dissipate with time as vendors became familiar with the new instrument and voucher. The 2010 WIC Erroneous Payment Study explored the relationship and in general found that the introduction of the new package had no effect. But, detection of an effect may have been obscured by the delay in resolving investigations. In other words, a case completed (i.e., resolved) after the introduction of the food package change actually reflected cases initiated before the change. The cases initiated after the introduction of the package would, for the most part, be completed in 2011. The analysis planned for the 2011 report should show effects, if there are any.

## 6.1 EFFECT ON OVERCHARGING

TIP has been the primary source of information on overcharging for the development of annual overcharge estimates. However, neither TIP nor any other currently available data source offers any specific information on how the investigations are conducted and thus no linkage to whether overcharges were related to changes in the WIC food packages. However, if we assume that investigations include the full range of purchase options available through food instruments and cash value vouchers, we might expect that any confusion due to the changes would show up during regular food purchase investigations. In other words, this expression would yield a greater percentage of violations than would have occurred had the new food packages not been implemented. To examine this question, an interrupted time series methodology is used to detect significant differences between pre-change and post-change outcomes, with outcomes being defined as an overcharge. The major comparison is based on constructing an intervention variable that represents the period in which the food package changes were in effect. That intervention variable would indicate the degree to which the outcome variable, as measured in the intervention period, changed relative to earlier periods.

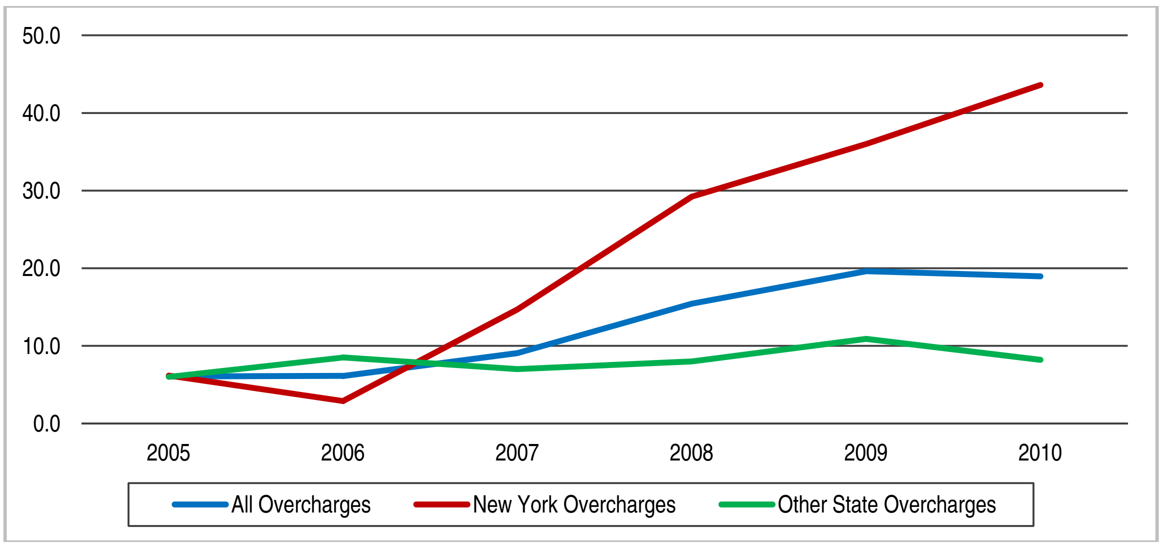
In estimating the overcharge impact related to the change in food packages in the 2010 study, TIP data from FY2005 to FY2010 were assembled, and the file was processed to include those records in which the investigation was completed. [36] We confined the study to the States and vendor types that were used for generating the FY2010 estimates that provided the overcharge estimates presented in the report. Two States, New York and Delaware, implemented the food package in early January 2009, about 3 months into the fiscal year. All others implemented the food package in the latter part of FY2009. Because TIP does not provide a monthly profile of investigative activity, we considered FY2009 and FY2010 as the intervention period for New York State and Delaware and FY2010 as the intervention period for all other States. For New York State and Delaware, we assumed that if the change in the food package had any effect, it would begin to show up in FY2009.[37]

In examining the changes in the overcharge rate over time, Exhibit 6 reveals that the violation rate (blue line) trends upward between FY2006 and FY2009. The rate declines slightly for FY2009 and FY2010, leading us to think that changes in the food package had no effect. It should be noted that New York State accounted for a very large proportion of investigations. Of the 3,872 investigations occurring in FY2010, 1,174, or 30.32 percent, were conducted in New York State. Thus, the analysis will look at New York separately from other States. The trend line for New York State showed dramatic increases in the overcharge rate, including the time period from FY2009 to FY2010. Although this might be consistent with the hypothesis that overcharging increased as a result of the introduction of the new food package, it does not

[36] As discussed previously, this sample contains a selection bias that tends toward including those vendors that are most prone to violate. However, we would expect that the sample would be consistently biased over the years with the no reason to assume that there are selection issues with the magnitude of the bias.
[37] One issue is the lag between the actual covert purchases made during an investigation and its completion. It could be argued that many investigations completed in FY2009 were conducted in FY2008, before the package was adopted.

account for the dramatic rise in the overcharge rate throughout the period. For other States, there was a slight decline in the rate. There is one seeming inconsistency related to the overcharge rates in FY2010. Because of the level of activity and the earlier implementation of the new food package in New York State, we will separately explore the effect in New York and in other States.[38] Because of issues with the FY2005 value (related to its similarity to the FY2006 value, which seemed to initiate a growth in the overcharge rate); we limited the regression analyses to FY2006 to FY2010.

**Exhibit 19. Trend in the Overcharge Violations Rate**




Source: Annual Measure of WIC Erroneous Payments to Vendors: 2011 Methodology Report.

To measure the effect independent of the trend and changes in the composition of the vendors, we developed the following model.

$$\Pr(Overcharge(i)) = e^{w}/(1+e^{w})$$

Where y is a linear combination defined by:

$$\mathrm{w} = \alpha + b1\,trend + b2\,intervention \sum_{i=3}^{15} \beta iXi + \varepsilon$$

[38] As in the other chapters, ITOs, U.S. Territories, Alaska, Hawaii, Vermont, North Dakota, and Mississippi were eliminated from this analysis.

- $\alpha$ is an intercept term that contains the estimate for cases that are not explicitly included in the equation.
- $b1$ represents a linear trend term, which controls for the linear increase in the proportion of violations.
- $b2$ represents the effect of the impact of the change in the food package (the intervention) controlling for the trend and other variables.
- $\beta_i$ represent terms that will help ensure that any effect is not due to changes in the vendor population.
- $\varepsilon$ is the error term.

The $X_i$'s include the following covariates or control variables:

- Store Type:
  - Large Stores
  - Small Stores
  - Retailers of Unknown Size
  - WIC-Only Retailers
  - WIC Above-50-Percent Stores

- Ownership:
  - Public
  - Private
  - Ownership Not Known

- Poverty Level:
  - Twenty Percent Under the Poverty Level
  - Twenty to Thirty Percent Under the Poverty Level
  - Thirty or More Percent Under the Poverty Level

- Urbanization Level:
    - Less Than 50 Percent Urbanized
    - Between 50 and 90 Percent Urbanized
    - More Than 90 Percent Urbanized
- Vendor Authorization Status:
    - Authorized in Last Fiscal Year
    - Authorized Prior to Last Fiscal Year

The 2011 report will duplicate this analysis with three exceptions. First, it will add 2011 data to the time series thus extending the series from 2006 to 2011. Second, it will explore different intervention periods—that is it will assume that the effect will occur in 2011, rather than in 2010 as in the previous analysis. Third, it will use the respecified alternative model generated in Section 4.2.

## 6.2 EFFECT ON UNDERCHARGING

The effect of the food package changes on underpayments is more difficult to estimate because TIP does not contain information on occurrences of undercharges, and the only available source of information is the 2005 WIC Vendor Management Study. For the 2010 study, we constructed a two-stage approach. In the first stage, we estimated the probability of an undercharge. A model was used to generate set of parameters that were then applied to the dataset that included all complete investigations to obtain the probability of undercharging for each store in the investigative file.

For the 2011 report, we will utilize the revised model generated in Section 5.2. As before, we estimated the impact of the new food package on erroneous payments separately in New York State and in other States.

# 7. CONCLUSIONS

## 7.1 SUMMARY OF STUDY OBJECTIVES

The generation of improper payment estimates based on WIC vendor over- and undercharges was last estimated through a nationally representative sample of WIC vendors in the 2005 WIC Vendor Management Study. Since that time, yearly updates to the estimates have been made through the WIC erroneous payment update studies. Overpayment estimates, or the amount paid out by WIC that exceeded the price a non-WIC customer would pay for the foods purchased with WIC funds, were developed through a statistical procedure (raking) that produced weights, allowing the translation of investigative findings to the population. The idea was that investigations are, by their very nature, biased toward vendors that are disposed toward overcharging and other violation-prone behaviors; therefore, some adjustment was necessary in order to align those vendors to the population. Undercharge estimates were developed from predictive models based on data collected in the 2005 WIC Vendor Management Study, since no other data source for undercharges is available.

This report has three objectives: 1) to explain the current methodology that has been used in previous update studies and that will be used in the 2011 update; 2) to develop and test alternative models for generating over- and undercharge estimates; and 3) to provide a preliminary design for assessing the effect of changes to the WIC food package on under- and overcharges.

## 7.2 SUMMARY OF APPROACHES USED FOR THE 2005–2010 UPDATES

Regarding the report's first objective, approaches for developing overpayments have been based on a procedure to adjust State-conducted investigative cases to the population of WIC vendors. Investigative case information derives from TIP, which is an annual database provided by the States on all authorized WIC vendors, and the investigative activity on these vendors Investigative cases are selected using different techniques, including automated identification systems that target high-risk vendors, leads and other informal approaches. There is therefore no a priori statistical scheme that would allow the calculation of probabilities of selection, and therefore no way to reasonably translate the results of the investigations directly to the population. Post-stratification, and raking, allows this translation by providing weights for each investigative vendor using vendor characteristics that are critical to differentiating investigated vendors from the population. The weights obtained from raking are combined with the probability that a vendor with certain characteristics overcharges to produce both a store-based estimate and a food outlay-based estimate.

The estimation of an undercharge results from a three-step process using predictive equations derived from data supplied by the WIC Vendor Management Study. In the first step, each vendor in TIP is assigned a probability of undercharging. When summed, this provides the overall number of vendors violating. Second, each vendor is assigned, based on a predictive equation, the amount that they would undercharge, if they undercharged. Third, the probability of an undercharge is multiplied by the amount estimated in the second step, to produce the undercharge amount.

### 7.3 ALTERNATE APPROACHES

The alternate approaches developed within this report rely on model based regression methods. This development is based on a six steps for providing vendor-based over- and undercharge rates and erroneous charges. The six steps are the same for deriving the over- and undercharge rates.

Step 1—this step involves the development of predictive models based on a logistic regression of overcharges and separately undercharges on a variety of vendor characteristics. The predictive model for overcharges also includes indicators of State-level management practices and for undercharges, the probability of an overcharge. The regressions were estimated using Bayesian MCMC techniques. Overcharges were estimated from TIP investigative data while undercharges were estimated largely using information from the 2005 WIC Vendor Management Study.

Step 2—the second step takes the predictive models and applies them to each vendor in the WIC vendor population. The result is a probability of over- and undercharging for each vendor based on their characteristics. To allow for estimation of credibility intervals, we generated 200 probabilities for each retailer using the posterior distributions of the model parameters.

Step 3—the third step is to sum the probabilities over the entire WIC Vendor Population to obtain the number of overcharging and undercharging vendors. Credibility limits for the 5th and 95th percentiles were also established.

Step 4—this step is the first in deriving an amount representing actual over- and undercharges. A random sample of proportion was drawn representing the amounts over- and undercharged as a proportion of the total amount on the food instrument in the presence of a specific over- or undercharge. The derivation of this proportion was guided by data from the 2005 WIC Vendor Management Study. The sampling allowed us to develop credibility estimates associated with the final estimates.

Step 5—The fifth step applies the random sample of dollar values to the WIC Vendor population and obtained the adjusted over- and undercharge food outlays per vendor by multiplying the probability of an over/undercharge, total food outlays for that vendor, and the proportional amount of food outlays that were estimated to be over/undercharged.

Step 6—the over- and undercharged food outlays per vendor are summed to obtain the total over- and underpayment amounts for the year.

## 7.4 SUMMARY OF RESULTS USING ALTERNATIVE APPROACHES

Overcharge Rates—The alternative approach produced notably smaller violation and dollar overcharge rates than the raking procedure. Although the reason for this has not been fully explored, the initial belief is that it reflects the lack of attention in the raking process to State differences in investigative approaches. For instance, New York accounts for a large proportion of all cases, and has a very high overcharge violation rate. If not considered within the estimation procedure, the final estimate would resemble New York's violation rate (i.e., higher) to a larger degree than other States (i.e., lower). The alternative approach, by including State-level variables, compensated for this and thus created lower rates than were obtained through raking.

It should be noted that the situation in New York explains overcharging not only in New York but also in the other States. In other words, maybe the other States, with low violation rates, have WIC vendor management practices that tend to be not as focused as New York's in identifying overcharge violators.

Finally, with regard to overcharges, the range of estimates considered to be credible is larger than those in the raking situation. Part of this may be due to some assumptions used in establishing the regression based approach, but it also may be due to the high consistency among vendors associated with the investigative sample.

Undercharges—Undercharges estimated through the alternative regression approach show a high level of consistency with the undercharges estimated through the standard approach. The one difference would be the larger expression of deviance expressed by the alternative approach. This probably reflects more realistic assumptions concerning the performance of the models.

Comparison of Trends—although we were not able to do a trend analysis using alternative methods, we did compare the results from predictions made for 2009 and 2010 to the raking and standard over- and undercharge estimates. In general, although the absolute values of the estimates were different (for overcharges), the direction and magnitude of the changes between 2009 and 2010 were similar among the two sets of methods.

## 7.5 SUGGESTIONS AND RECOMMENDATIONS

The results indicate that the alternative methods succeeded at proof of concept, and offers a more flexible and understandable system than the standard procedures used in previous studies. And it allowed us to probe into some of the reasons why the estimates could differ and to examine assumptions underlying the approach. However, we believe that there needs to be further work on the modeling and on consideration of the primary assumptions underlying the modeling in order to fully develop the alternative approaches. This further exploration should be done in conjunction with the results from the 2012 WIC Vendor Management Study.[39]

[39] In fact, an optimal approach would be to investigate an optimal update strategy and ensure that the WIC Vendor Management Study uses that strategy in its data collection efforts.

# APPENDIX A:

# DESCRIPTION OF RAKING

The following illustration provides an explanation of the raking process. This process starts with a two-dimensional matrix with three categories in each dimension and assumes that the population consisting of 10,000 vendors is scattered across the cells, as shown in Exhibit A1. This process also assumes that the corresponding sample of 1,000 investigated vendors is scattered across the same 9 cells, as shown in Exhibit A2.

**Exhibit A1.**
**Vendor Population Distributed Across Two Dimensions**

| Dimension 1 (e.g., urbanization) | Dimension 2 (e.g., poverty) | | | |
|---|---|---|---|---|
| | Low | Medium | High | Total |
| Low | 300 | 400 | 300 | 1,000 |
| Medium | 1,500 | 1,500 | 1,000 | 4,000 |
| High | 700 | 600 | 3,700 | 5,000 |
| Total | 2,500 | 2,500 | 5,000 | 10,000 |

**Source:** WIC Erroneous Payments to Vendors: Annual Estimates for 2009.

**Exhibit A2.**
**Vendor Sample Distributed Across Two Dimensions**

| Dimension 1 (e.g., urbanization) | Dimension 2 (e.g., poverty) | | | |
|---|---|---|---|---|
| | Low | Medium | High | Total |
| Low | 40 | 60 | 100 | 200 |
| Medium | 100 | 200 | 200 | 500 |
| High | 60 | 40 | 200 | 300 |
| Total | 200 | 300 | 500 | 1,000 |

**Source:** WIC Erroneous Payments to Vendors: Annual Estimates for 2009.

A comparison of Exhibits A1 and A2, shows that the sample is not consistent with the population, overstating representation in certain categories and understating it in others. The object of raking is to determine weights that will allow the translation of the sample to the population so that the sample is truly representative of the population.

Exhibit A3 provides an example of the initial raking matrix. The cell entries represent sample values, and the marginal totals represent population values. As discussed above, the idea is to identify values for the cells that will add up to the marginal population values. Each value is assigned a weight that allows this transformation to occur. Multiple iterations are needed to accomplish this when the transformation involves two or more dimensions.

**Exhibit A3.**
**Initial Raking Matrix**

| Dimension 1 (e.g., urbanization) | Dimension 2 (e.g., poverty) | | | |
|---|---|---|---|---|
| | Low | Medium | High | Total |
| Low | 40 | 60 | 100 | 1,000 |
| Medium | 100 | 200 | 200 | 4,000 |
| High | 60 | 40 | 200 | 5,000 |
| Total | 2,500 | 2,500 | 5,000 | 10,000 |

**Source:** WIC Erroneous Payments to Vendors: Annual Estimates for 2009.

For the first iteration, the weight is calculated by dividing the population total by the sum of the cell sample values (see Exhibit A4). Thus, 1,000 is divided by 200 for a weight of 5. The weights are calculated for the first iteration. Note that the weights for the second iteration are not calculated.

| Exhibit A4. Marginal Frequencies and Percentages for the Population and Sample | | | | | | |
|---|---|---|---|---|---|---|
| Dimension | | Population (Marginals) | | Sample (Marginals) | | Weight |
| | | Number | Percent | Number | Percent | |
| Dimension 1 | Low | 1,000 | 10 | 200 | 20 | 5 |
| | Medium | 4,000 | 40 | 500 | 50 | 8 |
| | High | 5,000 | 50 | 300 | 30 | 16.7 |
| | Total | 10,000 | 100 | 1,000 | 100 | |
| Dimension 2 | Level 1 | 2,500 | 25 | 200 | 20 | * |
| | Level 2 | 2,500 | 25 | 300 | 30 | * |
| | Level 3 | 5,000 | 50 | 500 | 50 | * |
| | Total | 10,000 | 100 | 1,000 | 100 | |

* = no weight assigned at this stage.

**Source:** WIC Erroneous Payments to Vendors: Annual Estimates for 2009.

A new sample cell frequency is calculated by applying the weights to the original sample cell frequency (see Exhibit A5). These new cell frequencies will add to the Dimension 1 marginals but not to the Dimension 2 marginals. Therefore, we have to adjust the cell values to the Dimension 2 marginals.

| Exhibit A5. Weights Resulting From Initial Rake | | | | |
|---|---|---|---|---|
| Dimension 1 | Dimension 2 | Original Sample Cell Frequency | Weights From Initial Rake (Exhibit 4) | New Cell Frequency |
| Low | Low | 40 | 5 | 200 |
| | Medium | 60 | 5 | 300 |
| | High | 100 | 5 | 500 |
| Medium | Low | 100 | 8 | 800 |
| | Medium | 200 | 8 | 1,600 |
| | High | 200 | 8 | 1,600 |
| High | Low | 60 | 16.7 | 1,000 |
| | Medium | 40 | 16.7 | 760 |
| | High | 200 | 16.7 | 3,340 |

**Source:** WIC Erroneous Payments to Vendors: Annual Estimates for 2009.

The second step is to divide the population marginals for Dimension 2 by the new cell frequencies summed over Dimension 2. This gives a new set of weights as shown in Exhibit A6. Note that Dimension 1 is ignored in this iteration.

**Exhibit A6.**
**Marginal Frequencies and Percentages for the Population and Sample**

| Dimension | | Population (Marginals) | | New Cell Frequencies (Marginals) | | Weight |
|---|---|---|---|---|---|---|
| | | Number | Percent | Number | Percent | |
| Dimension 1 | Low | 1,000 | 10 | 1,000 | 20 | * |
| | Medium | 4,000 | 40 | 4,000 | 50 | * |
| | High | 5,000 | 50 | 5,000 | 30 | * |
| | Total | 10,000 | 100 | 10,000 | 100 | |
| Dimension 2 | Level 1 | 2,500 | 25 | 2,000 | 20 | 1.25 |
| | Level 2 | 2,500 | 25 | 2,660 | 27 | 0.94 |
| | Level 3 | 5,000 | 50 | 5,340 | 53 | 0.94 |
| | Total | 10,000 | 100 | 10,000 | 100 | |

* = no weight assigned at this stage.

**Source:** WIC Erroneous Payments to Vendors: Annual Estimates for 2009.

When the Dimension 2 weights are applied to the cell frequencies, we get the results displayed in Exhibit A7. When added, the cell values sum to the Dimension 2 marginals; however, they lose their coherence with the Dimension 1 marginals. To ensure that the cell values maintain coherence with both the first and second dimensions, we repeat the raking, first across Dimension 1, then over Dimension 2. Each repetition will result in values that are closer to the population values. Raking will be completed when the marginals calculated from the cell values are equal, or close to equal, to the population marginals for all dimensions. The ultimate weight after these iterations will represent the number of vendors represented by each sample point.

| Exhibit A7.<br>Weights Resulting From Initial Rake | | | | |
|---|---|---|---|---|
| **Dimension 1** | **Dimension 2** | **New Cell Frequency** | **Weights From Initial Rake** | **New Cell Frequency After Dimension 2 Rake** |
| Low | Low | 200 | 1.25 | 250 |
| | Medium | 300 | 0.94 | 282 |
| | High | 500 | 0.94 | 470 |
| Medium | Low | 800 | 1.25 | 1,000 |
| | Medium | 1,600 | 0.94 | 1,504 |
| | High | 1,600 | 0.94 | 1,504 |
| High | Low | 1,000 | 1.25 | 1,250 |
| | Medium | 760 | 0.94 | 714 |
| | High | 3,340 | 0.94 | 3,140 |

**Source:** WIC Erroneous Payments to Vendors: Annual Estimates for 2009.

# APPENDIX B:

# FREQUENCY TABLES ON THE POPULATION, THE INVESTIGATIVE SAMPLE, AND OVERCHARGING VENDORS, BY SELECTED VARIABLES

**Exhibit B1.**
**Number and Percentage of Investigated and Total WIC Vendors, by Vendor Characteristics**

| | Investigated Vendors | | Total Vendors | |
|---|---|---|---|---|
| Total Number of Vendors | 4,255 | 100% | 41,612 | 100% |
| Retailer Type | | | | |
| Large Retailer | 2,102 | 49.4 | 30,510 | 73.3 |
| Small Retailer | 1,925 | 45.2 | 9,676 | 23.2 |
| Retailer Type Unknown | 159 | 3.7 | 1,061 | 2.6 |
| WIC Only | 17 | 0.4 | 152 | 0.4 |
| WIC Above-50-Percent Stores | 52 | 1.22 | 213 | 0.5 |
| Ownership | | | | |
| Public | 538 | 12.6 | 11,887 | 28.6 |
| Private | 3,661 | 86.0 | 28,985 | 69.7 |
| Ownership Not Known | 56 | 1.3 | 740 | 1.8 |
| Poverty Level of Area | | | | |
| 20 Percent or Less | 2,489 | 58.5 | 31,012 | 74.5 |
| More Than 20 Percent but Less Than 30 Percent | 969 | 22.8 | 6,680 | 16.1 |
| 30 Percent or More | 797 | 18.7 | 3,920 | 9.4 |
| Urbanization Level of Area | | | | |
| 50 Percent or Less | 643 | 15.1 | 8.051 | 19.4 |
| More Than 50 Percent but Less Than 90 Percent | 590 | 13.9 | 9,750 | 23.4 |
| 90 Percent or More | 3,022 | 71.0 | 23,811 | 57.2 |
| Vendor Authorized in the Last Year | | | | |
| Yes | 249 | 5.9 | 3,497 | 8.4 |
| No | 4,006 | 94.2 | 38,115 | 91.6 |
| Type of Training Provided | | | | |
| Annual | 1,871 | 44.0 | 19,423 | 46.7 |
| Interactive | 2,213 | 52.0 | 20,546 | 49.4 |
| No Training | 171 | 4.0 | 1,643 | 3.9 |
| Monitoring Activity | | | | |
| No Visits | 2,391 | 56.2 | 28,639 | 68.8 |
| One Visit | 1,647 | 38.7 | 11,545 | 27.7 |
| More Than One Visit | 217 | 5.1 | 1,428 | 3.4 |
| Risk Profile | | | | |
| High Risk | 3,430 | 80.6 | 7,364 | 17.7 |
| Non High Risk | 825 | 19.4 | 34,248 | 82.3 |

**Exhibit B2.**
**Statistics on Food Outlays for Investigated Vendors and the Total Vendor Population**

| Statistic | Investigated Vendors | Vendor Population |
|---|---|---|
| N | 4,255 | 41,612 |
| Mean | 100,015 | 102,166 |
| Standard Deviation | 188,244 | 188,681 |
| First Quartile | 19,746 | 17,835 |
| Median | 47,696 | 51,306 |
| Third Quartile | 115,815 | 126,640 |

| Exhibit B3. Number and Percentage of Overcharging and Investigated WIC Vendors by Vendor Characteristics | | | | |
|---|---|---|---|---|
| | Investigated Vendors | | Total Vendors | |
| Total Number of Vendors | Overcharging Vendors | Investigated Vendors | Overcharging Vendors | Investigated Vendors |
| **Totals** | **835** | **100%** | **4,255** | **100%** |
| Retailer Type | | | | |
| Large Retailer | 155 | 18.6 | 2,102 | 49.4 |
| Small Retailer | 613 | 73.4 | 1,925 | 45.2 |
| Retailer Type Unknown | 44 | 5.3 | 159 | 3.7 |
| WIC Only | 1 | 0.1 | 17 | 0.4 |
| WIC Above-50-Percent Stores | 22 | 2.6 | 52 | 1.22 |
| Ownership | | | | |
| Public | 22 | 2.6 | 538 | 12.6 |
| Private | 797 | 95.5 | 3,661 | 86.0 |
| Ownership Not Known | 16 | 1.9 | 56 | 1.3 |
| Poverty Level of Area | | | | |
| 20 Percent or Less | 396 | 47.4 | 2,489 | 58.5 |
| More Than 20 Percent but Less Than 30 percent | 215 | 25.8 | 969 | 22.8 |
| 30 Percent or More | 224 | 26.8 | 797 | 18.7 |
| Urbanization Level of Area | | | | |
| 50 Percent or Less | 30 | 3.6 | 643 | 15.1 |
| More Than 50 Percent but Less Than 90 Percent | 43 | 5.2 | 590 | 13.9 |
| 90 Percent or More | 762 | 91.3 | 3,022 | 71.0 |
| Vendor Authorized in the Last Year | | | | |
| Yes | 32 | 3.8 | 249 | 5.9 |
| No | 803 | 96.2 | 4,006 | 94.2 |
| Type of Training Provided | | | | |
| Annual | 355 | 42.5 | 1,871 | 44.0 |
| Interactive | 435 | 52.1 | 2,213 | 52.0 |
| No Training | 45 | 5.4 | 171 | 4.0 |
| Monitoring Activity | | | | |
| No Visits | 477 | 57.1 | 2,391 | 56.2 |
| One Visit | 325 | 38.9 | 1,647 | 38.7 |
| More Than One Visit | 33 | 4.0 | 217 | 5.1 |
| Risk Profile | | | | |
| High Risk | 769 | 92.1 | 3,430 | 80.6 |
| Non High Risk | 66 | 7.9 | 825 | 19.4 |