CONTRACT NO. AG-3198-K-100043 (TASK 2)

ANNUAL MEASURE OF ERRONEOUS PAYMENTS TO WIC VENDORS: 2010

Final Methodology Report

July 8, 2011

Prepared by:
Richard E. Mantovani, Task Manager
ICF International Inc.
7315 Wisconsin Ave., Suite 400W
Bethesda, MD 20814-3202

Submitted by:
Mathematica Policy Research, Inc.
600 Maryland Ave., SW, Suite 550
Washington, DC 20024-2512

CONTENTS

EXECUTIVE SUMMARY

The generation of improper payment estimates based on Special Supplemental Nutrition Program for Women, Infants, and Children (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. In developing these updated estimates, overcharge 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 underlying idea was that investigations are by their very nature biased toward vendors disposed toward overcharging and other violation-prone behaviors and that some adjustment was necessary to align these vendors to the population. Undercharge estimates were developed from predictive models based on data collected as part of the 2005 WIC Vendor Management Study since there is no other data source available for undercharges.

This report has two objectives: 1) to explain the approaches that have been used in the update studies and that will be used in the 2010 update and 2) to perform an exploratory analysis of alternative methodologies that could be used to generate estimates.

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 the investigative activity on these vendors by the state and other investigative agencies. When choosing vendors to investigate, state and other investigative agencies use a variety of criteria, which are not necessarily selected according to an a priori statistical scheme[1] that would allow extrapolation of the results to the population. The utilization of a post-stratification method such as raking allows this translation by providing weights for each investigated vendor based on characteristics that

[1] In other words, vendors are not selected by a probabilistic sampling method that based on probabilities of selection. Because of this, sampling weights are meaningless and extrapolation to the population, without post-stratification techniques is meaningless.

describe the bias in the investigative data in terms of the outcomes of interest (overcharging). 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 redemption-based estimate.

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 violating. 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. The estimates obtained by each of these methods reflect, to an extent, on the variables selected for the raking and undercharge predictive equations.

For the second objective, we explored alternative specifications of the variables and of the methodologies used. In general, the findings indicate that with regard to overcharging, three of the variables used in the raking process (store type, urbanization of the neighborhood in which the vendor is located, and whether the store was privately or publicly owned) were still effective predictors. Redemption level and the poverty level of the neighborhood in which the vendor is located were not significant predictors. In addition, two other variables (whether the vendor was newly authorized that year and whether the vendor was identified as high risk) were identified as critical for estimating overcharges. In determining whether the substitution of these two variables resulted in different results from those derived by the raking methodology that had been used in the 2005–2009 updates, four versions with differing combinations of variables were used to generate estimates. The results indicate that the newly formulated versions did not perform better than the version used for the 2005–2009 estimates.

With regard to undercharges, we found that the best predictor for undercharges was whether the vendor also had a transaction that was an overcharge. This makes sense if it is suspected that undercharges are the result of unintentional mistakes, which could result in an over- or undercharge. However, the use of this variable in conjunction with TIP would be difficult because only overcharges are collected for investigated cases.

1. 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 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,[2] and commissaries. In Federal fiscal year (FY) 2009, there were 47,828 authorized vendors in the United States and its territories and possessions.[3]

One of the programmatic concerns of the Food and Nutrition Service (FNS) is vendor overcharging.[4] Overcharging occurs when vendors, intentionally or not, charge the WIC participant more than a non-WIC customer for items prescribed by the food instrument. Overcharging results in additional funds for the vendor and thus reduces the funds available to serve more 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.

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 2011. These studies use covert purchases in a 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.

---

[2] 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.
[3] Some States, such as Mississippi and Vermont, operate food delivery systems that do not use retail vendors within that State.
[4] 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.
[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.

On an annual Federal fiscal year basis, FNS receives information on the characteristics of redemption activity for all WIC vendors as part of The Integrity Profile (TIP) data system. TIP also provides information on investigations and other 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 providing overcharge estimates since the 2005 WIC Vendor Management Study (also called the 2005 bookend study). However, because vendors that have shown a high-risk profile are usually selected for investigation, the data from State investigations alone would be expected to overestimate overpayments, thus providing a biased estimate for the population as a whole. Regarding undercharges, States do not record incidents of underpayment; therefore, TIP provides no investigative data on this type of error. However, TIP does provide information on vendor population characteristics and redemptions that can be used to "age" the 2005 WIC Vendor Management Study undercharge estimates.

Since 2005, FNS has generated annual estimates of the amount of redemptions 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 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 also be used for generating estimates for the 2010 data. Although there are other viable methodologies for generating these estimates, the use of current methodologies will produce an estimate that is consistent, at least in terms of the underlying methods, with previous estimates. In the following sections, we will first describe the current approaches and then explore alternative methodologies, first for overcharges and then for undercharges.

2. APPROACH FOR ESTIMATING OVERCHARGES

For the purpose of this study, 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 will provide almost all the information required to make an overcharge estimate. However, the 2005 bookend study statistics were used for certain adjustments, as explained below.

2.1 ESTIMATION OF OVERCHARGES

The estimation approach for overcharges involves three steps:

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

These steps are described in the following sections.

2.1.1 Estimation of Weights

This section discusses the estimation of weights that allows investigative information from TIPS to be translated to the WIC vendor population. The approach used for developing overcharge estimates is a post-stratification adjustment known as raking. An illustration that describes this technique is provided in Appendix A. Raking begins with defining the vendor population that overcharged. Exhibit 1 shows the number of vendors in the FY 2009 TIP file that were sanctioned for overcharging by type of oversight (or compliance investigation).[6] We will refer to this as the investigative sample. Compliance investigations are covert activities in which an undercover purchaser seeks to uncover instances of fraud and abuse.[7] Of the 5,777 vendors undergoing State-based compliance investigations, 903 vendors, or 14.7 percent, were identified as having overcharged.[8] 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 State Agency and Overall (TIP 2009)

Type of Oversight	Total Investigated	Total Overcharging	
		Sanctioned	Percent
Compliance investigations by State agency	5,777	797	13.8
Compliance investigations by State agency or other entity*	6,373	903	14.7

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

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 strata. The five characteristics over which the data were raked (i.e., vendor type, ownership, urbanization, poverty level, and redemption dollar quartile) were chosen on the basis of previous research in the Supplemental Nutrition Assistance Program (SNAP) showing a relationship between food

---

[6] There are other reasons for sanctions indicated in the TIP file. An assumption could be that any vendor found to show a pattern of abuse, regardless of its specific nature, would also be a potential overcharger. We did not adopt this assumption in previous erroneous payment update studies.

[7] 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).

[8] This number represents all cases that were undergoing or had undergone investigations. In theory, only those that had been marked as completed should be counted; however, there were "uncompleted" cases in TIP that were marked as having an overcharge sanction. While formulating the methodology in 2005, a decision was made to include these cases, and since that time this decision has been maintained to provide a consistent perspective on the estimate. In the regression analyses, we use the number of completed cases, which adds up to 4,255 cases in total and 835 vendors sanctioned for overcharges.

stamp trafficking and vendor and neighborhood characteristics.[9] That research substantiated a basic set of indicators that, when modified, would be useful for characterizing WIC transactions and examining WIC over- and undercharges. Exhibit 2 provides details on the variables used during the raking process, which together define 540 strata. The raking process sought to adjust or fit the marginal distribution of the sample as described by the five identified dimensions to the marginal distributions of the variables in the population. 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. The result is a weight for each stratum. The weights can be viewed as similar to sampling weights, and they have the same purpose of inflating or deflating the estimate made within each stratum.

[9] 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.

Exhibit 2. Variables Used in the Overcharge Raking Process		
Variable	**Categories**	**Justification**
Vendor Type	Large Retailers (Retailers Defined as Having More Than $800,000 in Gross Sales)	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.
	Small Retailers (Retailers Defined as Having Less Than $800,000 in Gross Sales)	
	WIC Retailers With Missing Information on Gross Sales	
	WIC-Only Stores	
	WIC Above-50-Percent Stores	
Ownership Type	Publicly Owned Stores	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.
	Privately Owned Stores	
	Ownership Unknown	
Poverty Level of the Vendor's Neighborhood (defined by ZIP Code as the number of households under the poverty level)	20 Percent or Less	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.
	More Than 20 Percent but Less Than 30 Percent	
	30 Percent or More	
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	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.
	More Than 50 Percent but Less Than 90 percent	
	90 Percent or More	
WIC Redemptions	Quartiles Based on Current-Year Redemption Distribution	This variable was introduced to control for the wide range of WIC redemptions between stores.

2.1.2 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 redemptions 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 that each of the investigated stores represents. 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.

The amount of redemptions overcharged was calculated using a similar process. At this point, the unadjusted value of overcharges was the sum of the weighted redemption dollars from vendors that were found to be overcharging within the sample. In other words, the redemptions associated with each overcharging vendor was multiplied by the weight and then summed.

2.1.3 Application of an Adjustment Factor

In FY 2009, approximately $342 million in redemption dollars was associated with vendors that overcharged. This estimate represents all redemption dollars associated with these vendors, as if they overcharged on every WIC purchase. This also assumes that no one buys food without a food instrument in these stores. It is more likely that vendors that intentionally overcharge either charged more than the shelf price for foods bought or charged for items not purchased during a partial purchase of the items specified on the food instrument.[10] Therefore, given the assumption that at least some of the items on the food instrument are purchased, the amount of the actual amount overcharged is a proportion of the $342 million in redemptions.

The 2005 bookend study provided data that were useful in computing this adjustment factor. The study included three types of buys (safe, partial, and substitution) in which a purchase was made 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 this study, only safe buys were used to generate this estimate, although the values were larger when partial buys of items on the food instrument were transacted.

[10] 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.

Exhibit 3 shows that the average overcharge 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.

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

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 is reduced to $37 million, which constitutes 0.86 percent of total redemptions. The redemption-based overcharge rate was the amount of overcharges found in the population of overcharging vendors divided by the total amount of redemption 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: WIC Erroneous Payments to Vendors: Annual Estimates for 2009.

2.1.4 Development of Confidence Intervals

It is important to note that the above approach provides a point estimate that is dependent on the investigative sample drawn. To establish confidence limits and investigate variation in the estimate, we employed a bootstrap methodology. This approach was based on the selection of multiple random samples from the larger sample of investigated vendors.[11] Each sample was subjected to the raking process to develop an estimate for the population and then aggregated to establish distributional statistics.

2.2 ANALYSIS OF THE EXISTING FNS OVERCHARGE METHODOLOGY

The initial discussions involving the development of annual over- and undercharge estimates were guided by the primary data sources available and the approaches that were in common use in the statistical community. Although the decision was made to employ a raking approach and a regression approach for estimating over- and undercharges, respectively, there was, as part of the previous contract, an effort to evaluate pertinent alternative methodologies and data sources. There was also an effort to validate the estimates from the selected raking approach against bookend study estimates. Although other alternatives were suggested and evaluated, no thorough data-driven evaluation took place. Since that time, experience has led to a better understanding of the data sources and how they mesh, as well as greater accessibility to methodologies that can perhaps provide better estimates.

Any effort to examine alternative methodologies should first focus on the variables used in the current approaches. The selection of variables used in the raking and regression methodologies (see Exhibit 2) were based on early analyses of SNAP trafficking. In those analyses, store type, store ownership, and store location (in terms of the urbanization and poverty levels of the vendor's neighborhood) were found to be factors in predicting retailer violation prone behavior. The use of these variables differs between the overcharge and undercharge methodologies. In the former, it is critical to examine how the TIP investigative sample relates to the population; in the latter, it is important to identify an optimal set of variables to produce a statistically efficient estimate.

[11] Samples were drawn from the investigative files and subjected to the raking algorithm. Each sample provided a mean. A grand mean and a standard deviation were estimated for all these samples.

As outlined in the previous section, the purpose of the raking approach is to produce weights that will translate results from a sample of investigations to the population of vendors. Given that the sample of investigations is not a probability sample,[12] a raking approach can be used to generate, post hoc, values that would approximate the weights were the sample drawn with probabilities of selection. The effectiveness of the approach depends to a large extent on the distribution of the sample across the various cells or strata, and this, in turn, reflects the variables selected and how the variables are structured into categories. There are two questions related to whether these variables are the appropriate ones.

- Are there other variables that might be better for generating weights? The primary issue here is whether the sample is significantly different from the population on the raking variables. If there are no differences on a particular variable, the use of that variable in the raking algorithm will provide little benefit and may cause issues in overly dispersing the sample across the set of cells in the raking matrix.
- Are the variables categorized correctly? Raking involves categorically defined variables or continuous variables that can be transformed into categories. It may be the case that certain categories should be combined or that selected cut-points for a continuous variable may be better than others. The objective of the categorization effort is to reduce the number of categories without losing information about how to transform the sample values into population estimates.

Other than the variables cited in Exhibit 2, several other variables were analyzed during the course of the WIC erroneous payment update studies. The most promising variable was an indicator distinguishing newly authorized vendors (i.e., vendors authorized in the last year) from previously authorized vendors. This distinction may reflect familiarity with WIC procedures and rules, which would lead us to think that newly authorized vendors will violate more frequently. Of course, we would expect that previously authorized vendors will have been vetted to a greater extent than newer vendors, and therefore this group will consist of lower-risk vendors (at least

[12] There is no reason to think that investigative samples were drawn in a systematic way that would produce selection weights, because the selection process is State dependent and is probably partially determined by a violation-prone profiling system and partially determined by leads and other informal targeting mechanisms that are subjective in nature.

from the standpoint of intentionally overcharging) since the overchargers will have been "weeded out."

In addition, other variables from the TIP file that may have some effect on overcharging include:

- Training—The expectation is that training provides solid guidelines for vendor behavior and should be associated with fewer instances of overcharging.
- Number of monitoring visits—Vendors may be reluctant to overcharge or violate in other ways when faced with a greater level of scrutiny imposed by the State.
- Risk designation—As a criterion for selection of cases to be investigated, we expect that this variable would be important for distinguishing the differences between the investigative sample and population.

In addressing the two questions posed above, we used an exploratory logistic regression approach modeling the dependent variable as whether the vendor was in the investigative sample and regression analysis to model redemptions.

2.2.1 Store-Based Analysis of Investigations

By comparing the investigative sample against vendors not in the sample, we are in essence comparing the investigative sample to the population.[13] Significant differences between the populations will reflect the potential need to employ post-stratification. The parameters generated by the logistic regression provide information on the probability of being investigated for each variable and variable category. The parameter estimates are in reference to a baseline category (incorporated in the intercept term), and they can be translated as the incremental probability of being selected for an investigation. For example, in Exhibits 5 and 6, which present the results of the logistic regression, publicly owned vendors display a value of -0.07, while privately owned vendors display a value of 0.347. This means that publicly owned vendors have a smaller chance of being in the investigative sample than the third category (i.e., other

[13] The major difference in this approach and an approach that tests the sample against the population is that the latter includes the sampled vendors in the population, and the population therefore contains the information present in the sample. In the case that we are pursuing, the comparison groups are separate and independent.

vendors for which ownership could not be established), which is represented in the intercept term, and that privately owned vendors have a greater chance of being included in the sample. Since the privately owned vendors vary significantly from the third category, this would signify that privately owned vendors would have a higher chance of being in the investigative sample and are therefore more highly represented in the sample than the population, when compared with other vendors. In Exhibit 5, we present an analysis of effect for each variable,[14] and in Exhibit 6, we present the parameter estimates, along with their standard errors, the standard error providing the variation associated with each parametric value, and the p value. Because of the relatively large number of vendors being analyzed, we restrict the significance to $p = 0.01$.[15]

Regression results can be summarized as follows:

- The redemptions variable, one of the variables included as a raking variable in the 2005–2009 overcharge estimates, was not related to investigations. This may reflect that, in the regression, we used redemptions amounts and not redemption quartiles, as were used in the raking approach.
- Store type was also not significant, although it would have been had we used the $p = 0.05$ level. This is somewhat surprising but, again, may be due to the categories used to characterize store type. The baseline category that was included in the intercept term was WIC above-50-percent stores. It may be the case that if large retailers were used as the baseline category, the small retailer category might have reached significance. Judging by the significance of two interaction terms that included small retailer type as a factor, we would assume that the major contrast would be between larger and smaller retailers and that inclusion of other retailers led to ambiguous results.

[14] The analysis of effect is the measure of whether the variable taken as a whole is significantly related to the dependent variable. The actual estimates provide information on the effect of the particular coding categories within the variable.

[15] The analysis at this stage was exploratory and, in this instance, was done on the entire WIC population of vendors. From a sampling perspective, all findings are significant. For our purposes, we are treating the population as a sample of vendors in a superpopulation. However, the large sample still would result in a greater likelihood that a parameter will be significant, thereby reducing our ability to distinguish its effectiveness for producing credible overcharge estimates. This is essentially why we choose the more restrictive criterion than that which is normally used ($p = 0.05$) for judging significant. It should be noted that in our analysis of undercharges, we used the $p = 0.05$ value because of the limited sample size.

- The poverty level of the vendor's neighborhood was also found to be not significant. This is surprising because, from the perspective of previous data analyses, violation-prone behavior always increased with increasing levels of poverty. It should be noted that risk designation seemed to be confounded with poverty level, and when risk designation is deleted from the model, poverty level becomes significant.[16] This is not a surprising result since risk designation status determined by the states is probably derived from using algorithms that factor in neighborhood characteristics—which most likely correlate the kinds of stores located in neighborhoods, and the kinds of buyers that can be expected to utilize those stores.[17] Our approach for the 2010 estimate would be to replicate what was done previously (i.e., use only poverty level of the neighborhood) primarily to maintain a consistent accounting over the years. For future studies, which would make use of the next WIC Vendor Management study and which might use a non-raking approach, we would suggest attempting to understand the relationship between risk designation and poverty level more closely so it can be modeled correctly. This may involve understanding the algorithms the states use in assessing risk.
- The other two variables that form the basis for the raking approach, urbanization and store ownership, were significant. As shown in Exhibit 6, which provides parameter estimates, vendors that were more likely to be investigated included retailers in highly urbanized areas and privately owned stores.
- All four additional candidate variables mentioned above (i.e., risk designation, new vendor status, number of monitoring visits, and training) were significant. Of these four, Exhibit 6 shows that:

    - High-risk vendors were more likely to be investigated.
    - Recently authorized vendors were not as likely to be investigated.
    - Vendors without monitoring visits had a lesser chance of being investigated.
    - Vendors with interactive training were more likely to be investigated.

[16] In separate analyses not presented in this report, we found that risk and poverty level were highly correlated.
[17] This study did not have access to these algorithms, thus it could not be determined if they actually used poverty level information to identify high risk areas.

- Two variables representing interactions were significant. As Exhibit 6 shows:

  - Privately owned small retail vendors in highly urbanized areas were less likely to be investigated.
  - High-risk small retail vendors were less likely to be investigated.

What emerged from this analysis is that three of the variables used in the raking procedure proved to be useful, or potentially useful, in translating the investigative sample to the population; the redemptions and poverty-level variables did not. Other variables seemed to differentiate the sample from the population, in particular, the high-risk designation and the length of association with the program. Although some interactions were significant, others were not, raising questions about the use of the raking methodology, which generates weights based on a full specification of interactions.[18]

That leaves two issues to be explored. The first issue involves the extent to which cut-points describing one of the five raking variables make a difference in the magnitude and variation of the estimate. As was noted above, we did generate model based estimates using continuous measures of urbanization and poverty level—with urbanization being significant and poverty level being significant when risk-designation was omitted from the equation. For raking, we needed to develop discrete categories that both make sense in terms of populating the raking matrix and for creating some face validity to the concepts they represent. In other words, we want to avoid situations where we have few observations populating a particular urbanization or poverty level category, and we want to avoid creating classifications where vendors in a highly urbanized area, for example, are mixed with vendors within a suburban area. These two overall objectives could probably be satisfied by a number of solutions—all of which would have to be assessed against some objective criterion that would certifiably indicate that the solution results in better estimates. The criterion may be a "least variance" solution or perhaps a solution that is close to an estimate that is derived from a random sample of vendors (such as that obtained from

[18] The raking process, that generates cell-specific weights based on the relationship between the marginal distributions from the sample and the population, is limited by the extent to which the raking matrix, (i.e., all the cells defined by defined raking variables) contains cells with no sample observations. Our analysis has shown that the current raking approach may be overspecified, in that it may contain too many empty cells. The distribution of cases across the cells would lead to a wide variance in weights, and to potentially poor estimates.

a WIC Vendor Management Study). The current cut-points provide consistency with previous estimates and we believe that any adjustment should wait for both the current WIC Management Study to be completed and for up-to-date statistics on urbanization and poverty reflecting data collected from American Community Survey, which should be available for small areas in 2012.

The second issue is the use of alternative approaches, in particular Bayesian model-based approaches and micro-simulation. As demonstrated above, regression approaches can be used to model overcharges (and as we will see the next section undercharges as well). The specification of the models developed above assumes a limited number of post-stratification variables and their interactions. A full model would specify all the interactions specified in the raking matrix, which would be inordinate and place an excessive resource burden on the regression model estimation procedure. Hierarchical regression based on Bayesian approaches and implemented through Markov Chain Monte Carlo (MCMC) simulation techniques have the ability to generate cell estimates by considering the multi-level nature of the model (marginal and cell effects). This would provide more precise estimates given that the focus is on cells rather than marginals as in other approaches, and allows pooling. In addition, information from other sources can be used to modify the parameters, thereby producing estimates based on previous information (through the use of priors). One approach is to use the next WIC Vendor Management Study to produce information on these priors. As it stands now the current WIC Vendor Management Study cannot do this.

Micro-simulation is another approach to generating estimates. By micro-simulation, we mean a computer program that works through many iterations of the WIC participant purchase process to arrive at an overall estimate. The micro-simulation would involve using the WIC Vendor Management Study and state level data to provide overall basis for modeling purchases. For each purchase, a vendor would have a chance to over or undercharge depending on data from the WIC Vendor Management Study and TIP data. This “chance” would be determined by a probability distribution relating to whether an overcharge or undercharge occurs. The specific probability will vary in each iteration according the distribution. Over all the iterations, overcharge and undercharge rates will be estimated and the amounts overcharged and undercharged will be summed to determine the overall impact.

Exhibit 5. Analysis of Overall Regression Effects Pertaining to Modeling the Probability of an Investigation, Based on All Vendors (N = 41,611)			
Effect	**DF**	**Wald Chi-Square**	**Pr > ChiSq**
Store Ownership	2	53.9853*	<.0001
Store Type	4	12.6888	0.0129
Poverty Level of Store's Neighborhood	2	5.5816	0.0614
Urbanization Level of Store's Neighborhood	2	26.2710*	<.0001
FY 2009 Redemption Amount	1	2.4256	0.1194
Newly Authorized Vendor	1	567.2210*	<.0001
Monitoring Activity	2	182.7891*	<.0001
Risk Designation	1	4175.9555*	<.0001
Training Received	2	89.7635*	<.0001
Located in the Poorest Urban Areas	1	0.5052	0.4772
A Privately Owned Small Retailer	1	2.5422	0.1108
A Small Retailer Located in an Area With the Highest Poverty Rate	1	0.0460	0.8302
A Privately Owned Small Retailer in a Highly Urbanized Area	1	22.0161*	<.0001
A Privately Owned Small Retailer Located in an Area With the Highest Poverty Rate	1	1.0516	0.3052
A High-Risk Small Retailer	1	16.0920*	<.0001
R-Square 0.2275	**Max-rescaled R-Square**		0.4708

*Significant at p = 0.01

Exhibit 6. Maximum Likelihood Estimates Pertaining to Modeling the Probability of an Investigation, Based on All Vendors (N = 41,611)						
Parameter		**DF**	**Estimate**	**Standard Error**	**Wald Chi-Square**	**Pr > ChiSq**
Intercept		1	-2.8327*	0.1546	335.7862	<.0001
Privately Owned Vendors	0	1	0.3473*	0.0750	21.4455	<.0001
Publicly Owned Vendors	1	1	-0.0748	0.0789	0.8975	0.3435
Large Retail Vendors	1	1	-0.2030	0.1092	3.4590	0.0629
Small Retail Vendors	2	1	0.0453	0.2187	0.0429	0.8358
Retailer Size Unknown	3	1	0.2259	0.1247	3.2831	0.0700
WIC-Only Stores	4	1	0.0381	0.2688	0.0201	0.8874
Areas in Which Poverty Level Is Less Than 10 Percent	1	1	-0.1389	0.0598	5.3981	0.0202
Areas in Which Poverty Level Is Between 20 and 30 Percent	2	1	-0.1168	0.0650	3.2338	0.0721
Areas in Which the Urban Population Is Less Than 50 Percent	1	1	-0.0194	0.0407	0.2272	0.6336
Areas in Which the Urban Population Is Between 50 and 90 Percent	2	1	-0.1506*	0.0393	14.7115	0.0001
Redemption Amount		1	-1.98E-7	1.269E-7	2.4256	0.1194
Retailer Authorized Prior to 2009	No	1	1.0401*	0.0437	567.2210	<.0001
No Monitoring Visits	0	1	-0.3788*	0.0394	92.6674	<.0001
One Monitoring Visit	1	1	0.2416*	0.0397	36.9526	<.0001
Low Risk	-1	1	-1.7485*	0.0271	4175.9555	<.0001
Annual Training	Annual	1	-0.2245*	0.0413	29.4786	<.0001
Interactive Training	Interactive	1	0.2140*	0.0428	24.9807	<.0001
Located in the Poorest Urban Areas	-1	1	0.0629	0.0885	0.5052	0.4772
A Privately Owned Small Retailer	-1	1	0.2208	0.1385	2.5422	0.1108
A Small Retailer Located in an Area With the Highest Poverty Rate	-1	1	-0.0584	0.2722	0.0460	0.8302
A Privately Owned Small Retailer in a Highly Urbanized Areas	-1	1	-0.2609*	0.0556	22.0161	<.0001
A Privately Owned Small Retailer Located in an Area With the Highest Poverty Rate	-1	1	0.2775	0.2706	1.0516	0.3052
A High-Risk Small Retailer	-1	1	-0.1993*	0.0497	16.0920	<.0001

* Significant at p = 0.01

2.2.2 Redemption-Based Analysis of Investigations

The raking methodology generates a vendor-based estimate and a redemption-based estimate; therefore, a similar analysis was performed on redemptions. In this case, the dependent variable is the log of the vendor's redemption amount. A general linear model was generated, including four distinct terms:

- An intercept that represented the average log redemption amount for all the categories not explicitly stated in the model
- An investigation term whose coefficient indicated the incremental log redemption amount for vendors that were investigated
- A main effect for each variable specified (or explicitly stated) in the model. The coefficient is interpreted as the incremental log redemption amount for a vendor with particular characteristics. Variables used in the previous logistic regression were modified to better fit the linear regression framework. The modifications were as follows:

    - Vendors whose ownership could not be discerned were combined with privately owned vendors.
    - A variable representing large vendors was constructed with other vendors, providing a contrast group.
    - A variable representing small retail vendors was constructed with other vendors, providing a contrast group.
    - A variable representing vendors in the area with the highest poverty rate was constructed and contrasted with other vendors in other areas.
    - A variable representing vendors in highly urbanized areas was constructed and contrasted with other vendors in other areas.

- An interaction effect between each main effect variable and the investigation variable. This in essence provides evidence of the degree to which redemptions vary between vendors with similar traits that were investigated. If this term is significant, it would indicate that redemptions associated with investigated retailers with certain traits do not equal redemptions of vendors with similar traits in the population.

Although the regression results (see Exhibit 7) present interesting findings about differences in the main effects, the interaction terms provide evidence on differences in redemptions between the sample and population and thus on our ability to distinguish redemption differences between investigated and other vendors. These interaction terms indicate (using the p = 0.01 level of confidence) that:

- Urbanization is the only variable used in the raking algorithm that distinguishes between investigated and other vendors. The coefficient for investigated large stores is marginally outside the acceptable level of confidence.
- Vendor authorization status in FY 2009 is also significant when comparisons are made between investigated and other vendors.
- Risk designation is significant when crossed with investigations.
- Investigated small private retailers differ in their redemption amounts from other stores.

Exhibit 7.
Regression Results Involving Differences in Redemptions Between the Investigated Sample and Other Vendors, Based on All Vendors (N = 41,611)

General Model Statistics

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	29	31044.4512	1070.4983	336.58	<.0001
Error	41582	132254.0138	3.1806		
Corrected Total	41611	163298.4650			

R-Square	Coeff. Var.	Root MSE	Mean of the Log Redemption Amount
0.190109	16.91719	1.783412	10.54201

Exhibit 7.—cont.				
Parameter Estimates				
Parameter	**Estimate**	**Standard Error**	**t Value**	**Pr > |t|**
Intercept	9.929967883*	0.04484139	221.45	<.0001
Investigated Vendors	0.090925105	0.04484139	2.03	0.0426
Ownership (Publicly Owned Vendors = 1)	0.217108635*	0.02349481	9.24	<.0001
Interaction Between Investigated Vendors and Ownership	-0.014720122	0.02349481	-0.63	0.5310
Large Retail Vendors	0.389358345*	0.03476195	11.20	<.0001
Interaction Between Investigated Vendors and Large Retail Vendors	-0.082758716	0.03476195	-2.38	0.0173
Small Retail Vendors	-0.040296159	0.09814925	-0.41	0.6814
Interaction Between Investigated Vendors and Small Retail Vendors	0.067154053	0.09814925	0.68	0.4939
Vendors Located in Areas of Highest Poverty	0.272360541*	0.06086419	4.47	<.0001
Interaction Between Investigated Vendors and Vendors Located in Areas of Highest Poverty	0.049573329	0.06086419	0.81	0.4154
Vendors Located in Very Urbanized Areas	0.262088692*	0.02014277	13.01	<.0001
Interaction Between Investigated Vendors and Vendors Located in Very Urbanized Areas	0.082982864*	0.02014277	4.12	<.0001
Previously Authorized Retailer	-0.618673978*	0.03104612	-19.93	<.0001
Interaction Between Investigated Vendors and Previous Authorized Retailers	0.249820028*	0.03104612	8.05	<.0001
Routine Monitoring Conducted	0.066616294*	0.01505246	4.43	<.0001
Interaction Between Investigated Vendors and Routine Monitoring Conducted	-0.026706883	0.01505246	-1.77	0.0760
Vendor Designated as High Risk	0.182805237*	0.02353756	7.77	<.0001
Interaction Between Investigated Vendors and High-Risk Vendors	-0.073008279*	0.02353756	-3.10	0.0019
Vendors Located in Highly Urbanized Areas With High Levels of Poverty	-0.203151454*	0.06334082	-3.21	0.0013
Interaction Between Investigated Vendors and Vendors Located in Highly Urbanized Areas With High Levels of Poverty	-0.081334806	0.06334082	-1.28	0.1991
Small Privately Owned Vendors	-0.252978249*	0.09286267	-2.72	0.0064
Interaction Between Investigated Vendors and Small Privately Owned Vendors	-0.039169635	0.09286267	-0.42	0.6732
Small Retailers Located in Areas of High Poverty	-0.080260329	0.18961409	-0.42	0.6721
Interaction Between Investigated Vendors and Small Retailers Located in Areas of High Poverty	-0.028372282	0.18961409	-0.15	0.8811
Small Privately Owned Retailers in Highly Urbanized Areas	0.053341511	0.03947412	1.35	0.1766
Interaction Between Investigated Vendors and Small Privately Owned Retailers in Highly Urbanized Areas	-0.150034053*	0.03947412	-3.80	0.0001
Small Privately Owned Retailers in Areas of High Poverty	-0.053015377	0.18834703	-0.28	0.7783
Interaction Between Investigated Vendors and Small Privately Owned Retailers in Areas of High Poverty	-0.039640606	0.18834703	-0.21	0.8333
Small Retailers With a High-Risk Designation	0.024008644	0.04339361	0.55	0.5801
Interaction Between Investigated Vendors and Small Retailers With a High-Risk Designation	0.046967905	0.04339361	1.08	0.2791

* Significant at p = 0.01

2.2.3 Store-Based Analysis of Overcharge Violations

The above analysis focused on identifying variables important to adjusting the investigative sample to the population. Conceptually, each cell or stratum in the raking matrix has a value that represents the number of overcharging vendors, or the amount of redemptions associated with overcharges. This value is multiplied by the weight derived from raking. The value, either the number of vendors or the amount of redemptions, is currently treated as a constant for that cell or stratum, since the influence of the variables used in the raking process are treated as fixed for that cell or stratum (unless they are specified differently). However, these cell values could also be seen as variables that are dependent on other variables not in the current raking mechanism. In this section, we will repeat our analyses of the last section but with a focus on whether there are overcharge violations among those in the investigative sample.

There were 4,255 vendors for which investigations were completed in FY 2009, with 835 found to be overcharging.[19] Exhibits 8 and 9 present the results of a logistic regression of violations on the same variables specified in Exhibits 5 and 6. The results from Exhibits 8 and 9 indicate that:

- 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.
- Previously authorized vendors were most likely to overcharge.
- High-risk vendors were most likely to overcharge.

These results suggest that all the original variables used in the raking algorithm, except the poverty-level indicator, are useful in identifying overcharging. Their continued inclusion in the raking algorithm is justified by these findings. However, two additional variables, previous vendor authorization and risk designation, may assist in obtaining more precise estimates. These

[19] Again, this includes only vendors in which the investigation was marked as completed. The updates used all ongoing investigations and thus included a larger number.

variables are therefore candidates for replacing some of the current raking variables—in particular poverty level and redemption quartiles.

Exhibit 8. Analysis of Overall Regression Effects Pertaining to Modeling the Probability of a Violation, Based on Investigated Vendors (N = 4,255)			
Effect	**DF**	**Wald Chi-Square**	**Pr > ChiSq**
Store Ownership	2	10.8409*	0.0044
Store Type	4	58.1837*	<.0001
Poverty Level of Store's Neighborhood	2	0.8083	0.6675
Urbanization Level of Store's Neighborhood	2	41.2906*	<.0001
FY 2009 Redemption Amount	1	0.0216	0.8833
Newly Authorized Vendor	1	19.4559	<.0001
Monitoring Activity	2	0.9513	0.6215
Risk Designation	1	15.0138	0.0001
Training Received	2	5.5568	0.0621
Located in the Poorest Urban Areas	1	0.0738	0.7859
A Privately Owned Small Retailer	1	0.0208	0.8854
A Small Retailer Located in an Area With the Highest Poverty Rate	1	0.3016	0.5829
A Privately Owned Small Retailer in a Highly Urbanized Area	1	0.4838	0.4867
A Privately Owned Small Retailer Located in an Area With the Highest Poverty Rate	1	0.4410	0.5066
A High-Risk Small Retailer	1	0.6983	0.4034
R-Square	0.1356	Max-rescaled R-Square	0.2157

* Significant at p = 0.01

Exhibit 9. Maximum Likelihood Estimates Pertaining to Modeling the Probability of a Violation Based on All Vendors (N = 4,255)						
Parameter		**DF**	**Estimate**	**Standard Error**	**Wald Chi-Square**	**Pr > ChiSq**
Intercept		1	-3.0236*	0.4153	53.0095	<.0001
Privately Owned Vendors	0	1	0.2468	0.1745	1.9992	0.1574
Publicly Owned Vendors	1	1	-0.5568*	0.1991	7.8179	0.0052
Large Retail Vendor	1	1	-0.7177*	0.2679	7.1755	0.0074
Small Retail Vendor	2	1	0.7457	0.5033	2.1957	0.1384
Retailer Size Unknown	3	1	0.4445	0.2838	2.4526	0.1173
WIC-Only Stores	4	1	-1.4491	0.8440	2.9477	0.0860
Areas in Which Poverty Level Is Less Than 10 Percent	1	1	-0.0300	0.1651	0.0330	0.8559
Areas in Which Poverty Level Is Between 20 and 30 Percent	2	1	-0.1139	0.1723	0.4372	0.5085
Areas in Which the Urban Population Is Less Than 50 Percent	1	1	-0.6631*	0.1443	21.1090	<.0001
Areas in Which the Urban Population Is Between 50 and 90 Percent	2	1	-0.1061	0.1362	0.6069	0.4360
Redemption Amount		1	4.001E-8	2.725E-7	0.0216	0.8833
Retailer Authorized Prior to 2009	No	1	0.4667*	0.1058	19.4559	<.0001
No Monitoring Visits	0	1	0.0692	0.0859	0.6486	0.4206
One Monitoring Visit	1	1	-0.0144	0.0850	0.0289	0.8651
Low Risk	-1	1	-0.4273*	0.1103	15.0138	0.0001
Annual Training	Annual	1	-0.1470	0.0816	3.2451	0.0716
Interactive Training	Interactive	1	-0.1718	0.0831	4.2737	0.0387
Located in the Poorest Urban Areas	-1	1	0.0644	0.2372	0.0738	0.7859
A Privately Owned Small Retailer	-1	1	0.0428	0.2966	0.0208	0.8854
A Small Retailer Located in an Area With the Highest Poverty Rate	-1	1	-0.2346	0.4271	0.3016	0.5829
A Privately Owned Small Retailer in a Highly Urbanized Area	-1	1	-0.0963	0.1385	0.4838	0.4867
A Privately Owned Small Retailer Located in Areas With the Highest Poverty Rate	-1	1	0.2776	0.4180	0.4410	0.5066
A High-Risk Small Retailer	-1	1	0.1223	0.1464	0.6983	0.4034

* Significant at p = 0.01

2.2.4 Redemption-Based Estimates of Overcharge Violations

In this section, we focus on differences in the logs of redemptions among vendors that have overcharged and vendors that were investigated and found not to have overcharged. The model is the same as the model whose results are shown in Exhibit 7, with the exception that the "investigation" variable is replaced by an "overcharge violation" variable, and the interaction terms use the overcharge violation variable. The regression uses only those vendors that were investigated. The results, presented in Exhibit 10, with regard to distinguishing differences between violating vendors and nonviolating vendors in their redemption levels, indicate that:

- The public/private distinction is significant and indicates that public and private vendors differ in redemption levels depending on whether they overcharged. This is the only significant difference among the four substantive variables (redemption quartiles being the other) used in the raking algorithm.
- The indicator that differentiates newly authorized and previously authorized vendors as overcharging vendors is significant.
- The redemption levels for the vendors receiving monitoring visits and overcharging or not overcharging variable is significant.

The results for redemptions seem to indicate fewer differences between vendors found to be overcharging and other investigated vendors. It is interesting that risk, a variable of significance when we examined stores, was not a factor in this analysis. It is also of note that only one of the variables used in the raking algorithm was found to be significant.

Exhibit 10.
Regression Results Involving Differences in Redemptions Between the Overcharging Vendors and Other Investigated Vendors Based on All Vendors (N = 4,255)

General Model Statistics					
Source	**DF**	**Sum of Squares**	**Mean Square**	**F Value**	**Pr > F**
Model	29	2115.19965	72.93792	23.20	<.0001
Error	4225	13283.63815	3.14406		
Corrected Total	4254	15398.83780			
R-Square	**Coeff. Var.**		**Root MSE**	**Mean of the Log Redemptions**	
0.137361	16.77017		1.773149	10.57323	

Exhibit 10.—cont.				
Parameter Estimates				
Parameter	**Estimate**	**Standard Error**	**t Value**	**Pr > |t|**
Intercept	10.0587517*	0.09037321	111.30	<.0001
Overcharging Vendors	-0.40994342	0.30063526	-1.36	0.1728
Ownership (Publicly Owned Vendors = 1)	0.23635010*	0.04654245	5.08	<.0001
Interaction Between Investigated Vendors and Ownership	-0.8441884*	0.24622307	-3.43	0.0006
Large Retail Vendors	0.29450925*	0.07433975	3.96	<.0001
Interaction Between Overcharging Vendors and Large Retail Vendors	0.05414242	0.15223122	0.36	0.7221
Small Retail Vendors	0.06777041	0.21383600	0.32	0.7513
Interaction Between Overcharging Vendors and Small Retail Vendors	0.08553606	0.46345533	0.18	0.8536
Vendors Located in Areas of Highest Poverty	0.31717742*	0.12105721	2.62	0.0088
Interaction Between Overcharging Vendors and Vendors Located in Areas of Highest Poverty	0.17290382	0.40530959	0.43	0.6697
Vendors Located in Very Urbanized Areas	0.34970395*	0.04042980	8.65	<.0001
Interaction Between Overcharging Vendors and Vendors Located in Very Urbanized Areas	-0.0725401*	0.16480733	-0.44	0.6599
Previously Authorized Retailer	-0.4641373*	0.06417077	-7.23	<.0001
Interaction Between Overcharging Vendors and Previously Authorized Retailers	0.63006624*	0.17414439	3.62	0.0003
Routine Monitoring Conducted	-0.02226230	0.03140261	-0.71	0.4784
Interaction Between Overcharging Vendors and Routine Monitoring Conducted	0.34164265*	0.07183775	4.76	<.0001
Vendor Designated as High Risk	0.09436109	0.04402387	2.14	0.0321
Interaction Between Overcharging Vendors and High-Risk Vendors	0.07901292	0.21760290	0.36	0.7165
Vendors Located in Highly Urbanized Areas With High Levels of Poverty	-0.24130520	0.12541983	-1.92	0.0544
Interaction Between Overcharging Vendors and Vendors Located in Highly Urbanized Areas With High Levels of Poverty	-0.44031827	0.42937915	-1.03	0.3052
Small Privately Owned Vendors	-0.39550749	0.19932384	-1.98	0.0473
Interaction Between Overcharging Vendors and Small Privately Owned Vendors	0.23289061	0.45060530	0.52	0.6053
Small Retailers Located in Areas of High Poverty	-0.03576640	0.44321763	-0.08	0.9357
Interaction Between Overcharging Vendors and Small Retailers Located in Areas of High Poverty	0.39057378	0.73523435	0.53	0.5953
Small Privately Owned Retailers in Highly Urbanized Areas	-0.04513060	0.08034306	-0.56	0.5743
Interaction Between Overcharging Vendors and Small Privately Owned Retailers in Highly Urbanized Areas	-0.18489078	0.24269555	-0.76	0.4462
Small Privately Owned Retailers in Areas of High Poverty	-0.13258757	0.44186215	-0.30	0.7641
Interaction Between Overcharging Vendors and Small Privately Owned Retailers in Areas of High Poverty	-0.33127580	0.71755278	-0.46	0.6443
Small Retailers With a High-Risk Designation	0.11088036	0.08817418	1.26	0.2086
Interaction Between Overcharging Vendors and Small Retailers With a High-Risk Designation	-0.24566309	0.27682019	-0.89	0.3749

* Significant at p = 0.05

2.2.5 Assessment of Alternative Raking Models

After examining the exploratory regression analyses, four versions of the raking results were generated for vendors and redemptions. The estimates for each version were generated using bootstrap sampling in conjunction with the raking algorithm and resulted in 3,000 samples of estimates, which were then summarized to provide relevant overall estimates. The four versions can be summarized as follows:

- Version 1—The raking procedure used the following five variables and the levels defined in Exhibit 2:

    - Store type
    - Vendor ownership
    - Poverty level
    - Urbanization level
    - Redemption quartiles

- Version 2—The new vendor variable replaced the redemption quartiles variable. The new vendor variable was coded as “Yes” or “No.”
- Version 3—The risk designation variable replaced the redemption quartiles variable. The risk designation variable was coded as “High Risk” or “Non High Risk.”
- Version 4—The poverty level and redemption quartiles variables were eliminated. They were replaced by the new vendor variable and the risk designation variable.

The results are shown in Exhibit 11 (store-based estimates) and Exhibit 12 (redemption-based estimates). It should be noted that there is no valid approach for judging whether these 2009 estimates are true values, since the last attempt to obtain a true value was the 2005 WIC Management Study. These analyses, however, do provide information about the consistency of the estimates and about the range of deviation for the estimates.

Exhibit 11 presents the results of the raking alternatives for providing the vendor-based overcharge rates. The data show that versions 1, 3, and 4 have roughly similar profiles, with mean rates approximating 9.3–9.4 percent. Version 2, which used the new vendor status variable instead of a redemption quartiles measure, had a mean of 12.4 percent. Of note, however, are the differences in the range between the 5th and 95th percentiles. For version 1, that range was 2.8 percent; for version 2, it was 3.6 percent; for version 3, it was 4.5 percent; and for version 4, it was 4.3 percent. These range figures indicate that version 1 provided the most consistent results over a large number of subsamples drawn from the investigative sample.

Exhibit 11.
Overcharge Store-Based Estimates Obtained From Raking and Bootstrapping

Post-Stratification Schema		Confidence Intervals and Mean Value		
		5th Percentile	Mean	95th Percentile
Version 1—Estimated for 2009	No.	3,294	3,885	4,463
	Rate	7.9%	9.3%	10.7%
Version 2—New Vendor Variable Substituted for the Redemption Quartiles Variable	No.	4,548	5,152	5,767
	Rate	10.3%	12.4%	13.9%
Version 3—Risk Designation Variable Substituted for the Redemption Quartiles Variable	No.	3,007	3,913	4,860
	Rate	7.2%	9.4%	11.7%
Version 4—New Vendor and Risk Designation Variables Substituted for the Redemption Quartiles and Poverty-Level Variable (Other Variables Have Been Modified)	No.	3,027	3,881	4,769
	Rate	7.2%	9.3%	11.5%

The data on redemptions (Exhibit 12) reflects the findings on vendors. Again, versions 1, 3, and 4 have the same approximate means, with version 2 being somewhat different. In terms of consistency, version 1 clearly shows the least range between the 5th and 95th percentiles. However, although it has a higher mean, the range associated with version 2 is the same as the 0.5 percent exhibited by version 1. Version 3 has a range of 0.7 percent, and version 4 has a range of 0.6 percent.

Exhibit 12. Overcharge Redemption-Based Estimates Obtained From Raking and Bootstrapping				
Post-Stratification Schema		**Confidence Intervals and Mean Value**		
		5th Percentile	**Mean**	**95th Percentile**
Version 1—Estimated for 2009	No.	$27,385,948	$36,741,354	$48,188,591
	Rate	0.6%	0.9%	1.1%
Version 2—New Vendor Variable Substituted for the Redemption Quartiles Variable	No.	$36,506,644	$46,612,513	$57,410,891
	Rate	0.9%	1.1%	1.4%
Version 3—Risk Designation Variable Substituted for the Redemption Quartiles Variable	No.	$20,937,166	$33,759,440	$49,277,408
	Rate	0.5%	0.8%	1.2%
Version 4—New Vendor and Risk Designation Variables Substituted for the Redemption Quartiles and Poverty-Level Variable (Other Variables Have Been Modified)	No.	$21,912,184	$33,667,636	$47,046,271
	Rate	0.5%	0.8%	1.1%

3. DEFINITION AND DISCUSSION 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 are not recorded in TIP and have not been used to issue sanctions. Therefore, for this study, both 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) and applied to the TIP data. This means that, when applied to TIP data in subsequent years, the total expected value of undercharges will change strictly as a function of changes in redemption dollar amounts and the characteristics of the population of WIC vendors.

3.1 METHODOLOGY FOR ESTIMATING UNDERCHARGES

The 2005 bookend study allowed retailers to undercharge on any of three types of buys. As shown in Exhibit 13, 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 13.
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: WIC Erroneous Payments to Vendors: Annual Estimates for 2009.

The proportion of vendors undercharging by type of buy is presented in Exhibit 14. 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 14.
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 number of buys, not the number of vendors.

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

With regard to dollar amount, the average undercharge in a safe buy was $0.94 for vendors undercharging (see Exhibit 15). 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 15.
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: WIC Erroneous Payments to Vendors: Annual Estimates for 2009.

In Exhibit 16, 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 16.
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: WIC Erroneous Payments to Vendors: Annual Estimates for 2009.

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.[20] In developing a predictive equation, logistic regression was used to model the probability of a vendor undercharge, and ordinary least squares regression techniques were used to model the amount of an undercharge.

[20] 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.

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,[21] 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 redemption 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:

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

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 redemptions for that vendor

LR_v is whether the vendor is a large retailer

[21] 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.

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
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.

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 redemption dollars were substituted for the value of the food instrument that was used when generating the equation from the 2005 bookend study data. The prediction equation is specified as:

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

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 redemptions 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

The third 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 redemption undercharge rate was calculated by determining the total amount of undercharges as a percentage of all redemption dollars.

3.2 APPROACHES FOR ESTIMATING UNDERCHARGES

As described above, underpayments were estimated using a three-step process. The attempt here is to generate an equation for assigning a probability that a vendor will undercharge and an expected undercharge value. These two variables would be multiplied to yield an expected amount of undercharge for every vendor in the population. The weights in this case, which are equivalent to the vendor-based raking weights, are the original predicted values from the logistic regression equation. The original estimating equations, first developed for the 2005 estimates, were 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. 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 incentive to undercharge, since the vendor is, in essence, subsidizing WIC (an alternate explanation is that the vendor uses WIC transactions as a loss leader, although the exact nature of the loss leader is not clear). Undercharging would be expected to occur as a random event reflecting a lack of particular controls regarding charging a consistent price. Including overcharges as an additional variable could provide information on the extent to which this occurs.

Exhibit 17 provides results from a logistic regression fitting undercharge violations across all three buys on a variety of regressors. The dependent variable in this case is whether the vendor undercharged or not. The significant findings indicate that being a publicly owned vendor reduced the probability of undercharging, as did being a large store. Stores in highly urbanization areas were associated with an increase in the probability that the vendor will undercharge. Most dramatically, however, the effect of whether the vendor also overcharged is significant and positive. It increases the odds of undercharging to a 2 to 1 ratio. Risk is also significantly related to undercharging.

Exhibit 17.
Estimates Generated From a Logistic Regression of the Probability That a Vendor Will Undercharge
(N = 1,485)

Variable	Effect (Wald Chi-Square)	Estimate	Standard Error	Odds/Ratio
Intercept		-2.0564*	0.0888	
Public	71.7332*	-0.3711*	0.0438	0.690
Percent of Households Under Poverty Level	1.3579	0.2268	0.1947	1.255
Large Retailer	19.1758*	-0.2697	0.0616	0.764
Small Retailer	2.7631	0.1169	0.0703	1.124
Percent Urban	5.0106*	0.1166	0.0521	1.124
Redemptions	0.0408	-1.77E-8	8.749E-8	1.00
Previously Authorized Vendors	1.8930	-0.2781	0.2021	0.757
High Risk	1.7111	-0.0571*	0.0436	0.945
Found to Be Overcharging	226.1546*	0.6993*	0.0465	2.012

* Significant at p = 0.05

Source: 2005 WIC Vendor Management Study (All Buys).

From this analysis, it is clear that the overcharge variable is the strongest predictor of undercharging. However, the overcharge variable is available only for investigated vendors, not the entire WIC population. Therefore, this predictor cannot be used in conjunction with TIP information to produce an undercharge estimate. One approach to dealing with this issue is to predict undercharges for the TIP investigative sample, then weight through raking or a regression approach those sample values to the population. Using the equation above and the current raking approach, we will attempt to derive an alternative estimate.

The second stage of providing an estimate is to provide an expected value of an undercharge. A second equation that was generated for the WIC erroneous payment update studies was a regression with the amount of underpayment as the dependent variable. The regression was only run for those vendors with underpayments (see Exhibit 18). The initial result clearly shows that vendor characteristics have no relationship to the amount of underpayment. The only significant variable is the amount of the purchase, which indicates that underpayment is larger for smaller purchases. An optimal model would be the intercept, whether the store is publicly owned, and the amount of purchase. The limitation of using this model at this point is that this undercharge model assumes a relationship across the three types of purchase made during the WIC Vendor Management Study rather than the current model that is based an assumption that safe buy purchases constitute the totality of purchases.[22] For future efforts, however, this model could be further developed and used, if the next WIC Vendor Management Study is structured in a way that can provide the required variables.

Exhibit 18.
Results of Regressing the Amount of the Undercharge on Vendor Characteristics for Overcharging Vendors (N = 187)

Variable	Estimate	Standard Error
Intercept	-0.91769	0.79039
Public	-2.6007	0.47449
Percent of Households Under Poverty Level	1.20056	2.15480
Large Retailer	0.31559	0.61376
Small Retailer	0.24661	0.67587
Percent Urban	-0.17358	0.52199
Previously Authorized Vendors	2.44232	2.05786
Purchase Total	-0.04574*	0.00846

* Significant at p = 0.05

Source: 2005 WIC Vendor Management Study (All Buys).

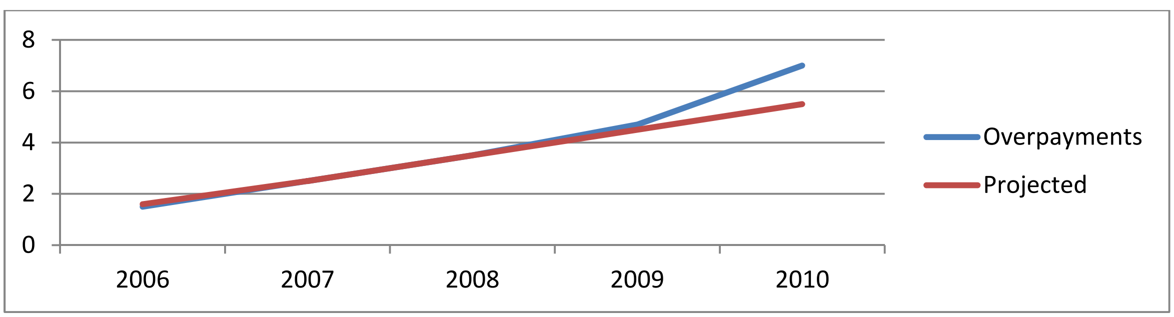
[22] This assumption was made because the natural occurrence of safe or full buys, partial buys and substitutions are not known.

4. FOOD PACKAGE CHANGES IMPACT ON ERRONEOUS PAYMENTS

FNS required that State Agencies implement a revised WIC food package by October 2009. In addition to adding foods that have more appeal to individuals with different ethnic backgrounds, it also offered the opportunity for WIC participants to purchase fruits and vegetables with a dollar denominated instrument, rather than one that specifies the product. With these changes, the potential for errors increases, since the new packages add new foods, and new processing complications. 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 sizes of package or product weight.

This analysis would ideally determine the degree to which the addition of the fruit and vegetable vouchers and the other changes has affected erroneous payments. Unfortunately, there are no direct and specific measures of erroneous payments (such as would be available from TIP) due to changes in the food packages (TIP does not collect information on food packages offered). Sources such as scanner information provide no direct information on whether the charge was in line with shelf prices and thus likewise cannot be used to directly measure this effect. In addition, scanner data would not be available for many smaller stores in which the confusion is likely to occur. Indirectly, however, we can assume that errors due to the food package changes would be expressed during regular covert purchases. This expression would yield a greater number of violations than normally experienced. In other words, if a vendor was investigated before and after the implementation of the new package, we could see an increase in overcharging from TIP data, if the changes were confusing the vendor. However, since vendors are not usually investigated at periodic intervals, we do not have the option of looking at violations over time for particular vendors. Instead, we suggest an approach that clusters vendors by their characteristics. Each cluster could then be followed over time using unweighted TIP data, thereby eliminating the need to follow particular vendors. Differences in violation rates within these clusters before and after the implementation of the program would yield information about overcharge patterns reflecting changes in the WIC food package. Other variables that have caused a temporal increase in overpayments will be controlled to some degree by use of other variables available on TIP that define the clusters. Since the changes are being implemented in FY 2009, but not uniformly by States during that year, differences may occur to some degree in FY2009, and perhaps to a larger degree in FY 2010. To address this staggered implementation, states conducting most of investigations will be examined to determine trends in overpayment from 2005-2008 in these states (See Exhibit 1). We will then examine 2009 and 2010 investigations within these states to determine if the values vary from the trend. These values will then be used to adjust the overpayment estimates for 2010 for the entire population.

Exhibit 1 Measuring the Effects of the Change in the WIC Food Package
(Effect =difference between actual overpayments and projected overpayments.)



This information will then be used within the raking approach to estimate the change in the overpayment component of erroneous payments due to the change in the food package. That is, we will add it as another assumption in assessing the effect of the food packages changes; much in the same way we use the WIC Vendor Management Study to establish some basic parameters in the raking process. It should be noted we hypothesize that vendors, once becoming familiar with the food package, will probably make fewer errors as time passes and regress to their previous erroneous charge rates. Thus, for estimates for FY 2011 and onward, this adjustment will need to be reconsidered.

The effect of the food package changes on underpayments is trickier, since there is no extant set of data (like TIP) that can accurately represent data points beyond the 2005 WIC Vendor Management Study. Consistent with the discussion of possible improvements in methods above, we will suggest a change in methodology to incorporate the values from the overpayment analysis in the last paragraph into the regressions used to generate the underpayment estimates.[23] This will involve examining the relationship between undercharges and overcharges from the WIC Vendor Management Study and using the relationship to extrapolate the potential underpayment estimate from the overpayment estimate. As indicated earlier, there are two equations generated for predicting undercharge estimates. The first involves estimating the probability of undercharging by any vendor in the population, and the second is focused on generating the expected amount of the overcharge given that vendor overcharged. The effect of the overcharge variable will be examined only to determine the probability of an undercharge, but not the expected amount of the undercharge. Although it is easy to conceive that the action of undercharging is linked to overcharging (See Appendix A), it is more difficult to conceive of how overcharging would be linked to the amount of the undercharge. The regression equation previously used for predicting the probability that any vendor undercharges will include this relationship. Predictive values for 2009-2010 will be compared to predictions using values from 2005-08 to determine the effect of the new WIC food packages.

[23] See Appendix A for further information on this relationship.

5. CONCLUSIONS

5.1 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. In developing these updated estimates, overcharge 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 two objectives: 1) to explain the approaches that have been used in the update studies and that will be used in the 2010 update, 2) to provide a preliminary design for assessing the effect of changes to the WIC food package on under and overcharges, and to perform an exploratory analysis of alternative methodologies that could be used to generate estimates.

5.2 SUMMARY OF APPROACHES USED FOR THE 2005–2009 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 redemption-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.

The estimates obtained by each of these methods reflect, to an extent, on the variables selected for the raking and underpayment predictive equations. For the second objective, we explored alternative specifications of the variables and of the methodologies used.

5.3 SUMMARY OF RESULTS

The analysis of overcharges consisted of five separate analyses:

- An analysis of the probability that a vendor will be investigated. This analysis provides information on whether the population of WIC vendors selected for investigation resembles those not selected for investigation. To the degree that these populations differ, the analysis would signify that post-stratification adjustment is necessary. The results show that three of the variables used in the raking process (store type, urbanization of the neighborhood in which the vendor is located, and whether the store was privately or publicly owned) were still effective predictors. Redemption level and the poverty level of the neighborhood in which the vendor is located were not significant predictors. In addition, two other variables (whether the vendor was authorized in the year of the estimate and whether the vendor was identified as high risk) were identified as critical for estimating overcharges.

- An analysis of whether there are redemption differences between investigated vendors and non-investigated vendors. This would determine whether there needs to be any adjustment made based on redemptions. The results indicate that urbanization is significant, as is whether the vendor was authorized in the year of the estimate. Risk and store type are significant when small retailers are the focus.
- An analysis of the probability of overcharging. This analysis provides additional evidence on factors that result in overcharging. Store ownership, vendor type, urbanization, whether the vendor was authorized in the year of the estimate, and high-risk status were identified as significant variables in this analysis.
- A sensitivity analysis comparing four different specifications of the raking algorithm based on the previous analyses. Four versions with different combinations of variables were used to generate estimates. The results indicate that the newly formulated versions did not perform better than the version used for the 2005–2009 estimates.

Regarding undercharges, we found that the best predictor for undercharges was whether the vendor also had a transaction that was an overcharge. This makes sense if it is suspected that undercharges are the result of unintentional mistakes, which could result in either an over- or undercharge. However, the use of this variable in conjunction with TIP would be difficult because only overcharges are collected on investigated cases.

5.4 ASSESSING THE EFFECT OF CHANGES IN THE WIC FOOD PACKAGE

One additional component described in this report is an examination of the effect of changes in the WIC food package on over and undercharging. The thought was that vendors would be more apt to mistakenly over or undercharge given the new food acceptable for purchase. The effect would be measured by examining the trends over the last number of years in under and overcharge behaviors. By using projections based on over and undercharges prior to the change in the WIC food package, and comparing these projections with estimated over and undercharges after this change, we can determine the extent to whether the change had an effect.

5.5 SUGGESTIONS AND RECOMMENDATIONS

The results indicate that the current raking procedure seems sufficient for producing consistent estimates that produce values with relatively acceptable confidence limits. Particularly intriguing is the role of risk and new vendor status as additional variables that could improve statistical efficiency of the estimates and, perhaps, provide better estimates. Because of the need to limit the raking approach to the number of variables that can be accommodated, there may be consideration of model-based approaches such as are discussed in Andrew Gelman's journal article "Struggles with Survey Weighting and Regression Modeling."[24] Such modeling would accommodate continuous variables and interactions without needing to address the need for values for each of the cells in the raking matrix. The value of the average overcharge could be more precisely modeled in terms of the latitude for overcharging provided by not-to-exceed limits on the food instrument.

The analysis of the undercharging estimates indicates that the models can be improved if the propensity to overcharge can be established for the population as a whole, not just for the investigative sample. Undercharges can be generated from models that envision a random process reflecting random transaction errors on the part of vendors.

We would, however, suggest that any changes to the methodology be explored in detail and be geared to the upcoming 2011 WIC Vendor Management Study.[25] That study should be closely synchronized to the update methodologies for any future improper payment update studies.

[24] *Statistical Science*, 2007, Vol. 22, No. 2, 153–164.
[25] 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. 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 Redemptions 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