WP. 28
ENGLISH ONLY

UNITED NATIONS STATISTICAL COMMISSION and
ECONOMIC COMMISSION FOR EUROPE

CONFERENCE OF EUROPEAN STATISTICIANS

Work Session on Statistical Data Editing
(Ottawa, Canada, 16-18 May 2005)

Topic (iii): Electronic data reporting – editing nearer source and multimode collections

EDR AND THE IMPACT ON EDITING—A SUMMARY AND A CASE STUDY

Supporting Paper

Submitted by the Energy Information Administration, DOE, United States[1]

I. INTRODUCTION

1. It has been well documented that data editing is one of the most resource intensive aspects of the survey process, if not *the* most. Much of survey literature and research is dedicated to methodologies, approaches and algorithms focused on defining edit rules and building editing systems with the goal of identify and correcting actual/potential response error. Unfortunately, much less has been written documenting the actual survey performance of the varied editing approaches, and the net effect on data quality. Concerns have frequently been raised that surveys often suffer from over-editing which results in increased respondent burden and frustration from edit failure resolution, introduction of error and bias as edits are resolved, and increased survey processing time and cost. While recognition has been given to the prevention of errors through better survey forms design, respondent instructions, interviewer training, etc., these efforts have been limited in their effect. Data editing as a traditional post-collection process has been challenged to some extent by CATI and CAPI surveys, but now, electronic reporting and internet data collection/web surveys have expanded that challenge through self-administered surveys by providing the potential for editing at data capture by the respondent. The presence and the extent of edits in electronic data reporting through computer self-administered questionnaires (CSAQ) via web surveys, downloadable software, and e-mail attachments that are implemented at the initial data entry and capture, versus those implemented in the traditional data editing stage depend on: the amount of development resources dedicated; the sophistication of the electronic option selected; the security of the transmission that is required; the quality of the data that is required; the amount of respondent burden that is acceptable, and the related concern for increased non-response.

2. This paper examines the change in electronic data reporting options and usage from 2003 to 2004 for one statistical agency. One fully web-based survey that recently implemented an editing module is then examined in more detail to better understand the respondents' views and use of the edit feature. In particular, for this survey, respondents were asked how they used the edit function, as well as, about the clarity and usefulness of the information provided for edit failures. The responses regarding the edit function for this survey is further compared to a study of the edit log which records information each time the edit function is invoked.

[1]Prepared by Paula Weir (Paula.Weir@eia.doe.gov).

II. RECENT PROGRESS IN ELECTRONIC DATA COLLECTION

3. A review of surveys conducted by the U.S. Energy Information Administration (EIA) revealed that electronic reporting on 65 surveys had dramatically increased from 2003 to 2004. The U.S. Government Paperwork Elimination Act of 1998 was an encouragement for Federal statistical surveys to move into more electronic modes, but little progress had actually been made in the first years following the Act. However, the discovery of anthrax, which shut down the main post office for government mail, provided the impetus for change for historically mail surveys. The short-run solution of posting surveys in PDF format on EIA's web site, along with facsimile return, kept mail surveys operating, but this crisis-based approach did not represent the most efficient electronic collection method. As a result of the perceived threat to respondents, respondents were ready also to accept more electronic modes of data collection, especially methods for which they had already developed a certain comfort level. After the immediate surge in survey responses via unformatted emails and facsimiles, an alternative method was implemented fairly quickly making use of formatted Word files, or Excel files in the survey form image. To encourage secured reporting, a link was placed directly on the electronic survey form that directed the respondent to secured transmission.

4. The implementation of the formatted files on EIA's website was successful because respondents felt comfortable with this option. From the respondents' viewpoint, this method was convenient, simple and safe. From EIA's viewpoint, data were received more quickly and forms were more readable. Total or net values were calculated as the respondent entered data on the spreadsheet, so some potential errors were avoided, but very little editing at collection was attempted beyond integrity checks to insure loading of the data to a database. These mostly included automatic totals, checks on field lengths, numeric vs. character, or valid codes (state, month, etc). Despite these limitations, the ease of implementing this option resulted in the number of surveys offering this option increasing 17% from 2003 to 2004. Surprisingly though, of the surveys offering both secured and unsecured transfer options, 86% of the surveys had more respondents choosing unsecured transfer in 2003. But, the number of respondents choosing secured has grown as the number of surveys offering secured transfer has increased approximately 52% in 2004 (from 27 to 41 surveys), as shown in Table 1. Yet, for those surveys that had previously offered secured transfer, only a few of the surveys experienced a large increase in usage of secured transfers, ranging from 18 to 54%, while roughly half of the surveys had more modest increases, and two surveys experienced decreases (ranging from 10 to 20%) in secured transmission usage. This finding is interesting in view of the frequent reference that security concerns are respondents' primary concern about reporting via the Web.

5. While this electronic option of formatted files on the web has been appealing to the respondents, the benefits have also been limited because of the complexity of data capture. Although some of the surveys utilize a Visual Basic conversion and SQLLLOADER to an Oracle database, many surveys continued to print the electronic responses, and re-key the data into the respective survey processing systems/databases, potentially introducing new errors. This along with limited editing capability has restricted the benefits to the agency of these electronic forms.

6. Internet data collection (IDC), using a browser-based approach, has become the alternative, requiring more resources to develop, test, and implement. The usage of this reporting option for surveys that offered the option of IDC has been steadily growing as more surveys have provided this option. Most of this growth has occurred in the last year with 20 surveys offering IDC in 2004, compared to 12 in 2003. More importantly, the percent of IDC respondents choosing the IDC option for those surveys has significantly increased. One survey offering this option for the first time realized a 51% usage by respondents, while the other surveys with the IDC option showed an average increase in usage of approximately 40%. One particular series of surveys made a concerted effort to increase electronic reporting, resulting in the achievement of usage rates greater than 50% across their IDC option surveys. The IDC surveys have successfully incorporated editing by respondents using server-side information, such as, respondent's previous period's data. Fatal edits are clearly the most commonly implemented, driven by database requirements. Edit rules that depend on fixed values or within form reported values are the next most commonly implemented. Edit rules that depend on external values, such as, previous

period's report require that data be accessible to the respondent at data capture, or quickly returned to the respondent in an interactive mode. Therefore, these edits are more resource intensive to implement and require that security concerns be addressed for confidential data. The surveys with an IDC option vary in approach to the respondent's requirement for edit failure resolution. Some require data correction or edit override with comment (hard edits), while others require no response or action from the respondent (soft edits).

Table 1. Electronic Usage 2003 and 2004

Electronic Method	Number of surveys using method and range of % respondents using method (2003)	Number of surveys using method and range of % respondents using method (2004)	Change in surveys using method from 2003 to 2004	Data Capture?	Editing within electronic collection?
Unformatted e-mail	5	11	120.00%	no	no
	10-90%	.35-100%			
Unsecured transfer Word or Excel file	39	36	-7.69%	no	Only totals
	1-100%	.17-80.8%			
Secured transfer Word or Excel file	27	41	51.85%	some	simple
	1-70%	.16-55%			
Diskette/CD software (e-mail, fax or mail back)	4	9	125.00%	only if diskette is mailed back	yes
	3-57%	1-100%			
PEDRO (mail CD, install and electronic submission) and download software	23	23	0.00%	yes	simple
	1-27%	2.5-51.4%			
Internet	12	20	66.67%	yes	yes
	2-99%	.1-100%			

7. Despite the increased usage of editing at the data reporting phase, editing is still being performed in the traditional data processing stage for not only non-IDC respondents, but also across respondents from all reporting modes for edits requiring integration of responses, as well as for IDC respondents by-passing the (soft) edit failures. It is important in the survey process that the edits performed are consistent across collection modes, and, that data from all collection modes are integrated and higher level edits performed across respondents, or across surveys as appropriate. This is necessary to optimize the editing process, in an attempt to not only prevent error in the most effective manner by exploiting the respondents' knowledge at data capture, but also continue to draw on the more comprehensive information available for validation at post-collection. Some balance of the two phases of editing is viewed as optimal for improving efficiency and data accuracy without negative side effects on response rates, particularly for mixed mode data surveys.

III. CASE STUDY

8. Editing in internet surveys has brought about a new set of issues and concerns. In addition to the traditional problem of determining the most effective and efficient edit rules, the internet survey edit process has to address how and when to invoke the edits to maximize data quality and minimize respondent break-off. Should the edits be performed after each data item is entered or after all items have

been entered? Should hard edits be used requiring the respondent to take action, or soft edits to alert the respondent but require no action? How should the edit failures be presented—in a separate window or directly on the data entry/survey instrument screen? Should the edit failures be presented one at a time or all together? Edit messages take on a different role than in traditional editing, communicating directly with the respondent and taking on the role or "social presence" of the interviewer in resolving an edit failure. These messages need to be written in simple, non-confrontational language, and convey meaning to which the respondent can relate and take the appropriate action. How much information should be conveyed with the edit failures?

9. One fully web-based survey that recently had implemented an editing module was examined in terms of the how the respondents used the edit feature in reporting. In this survey, State officials are the respondents who reported prices charged by the sampled businesses operating in their State. The overall reaction to the new edit, which identified businesses whose price change since the previous report was out of range, was very positive. The edit module was intended to be invoked by the respondent after all the data for that period had been entered. The system required the edit to be run prior to submission, but did not require the respondent to make any changes to the flagged data prior to submitting the data. Respondents were encouraged to provide comments on the edit failures but were not required to do so. The system, however, actually allowed the respondents to "Review Prices Before Submit", thereby running the edit, at any point in their data entry. This was an area of concern because the edit rule was based on the mean change of all the prices that were entered, and different expectations for editing would result if run on partial data. After respondents had used the edit feature in the IDC for four reporting periods, they were sent a questionnaire asking five basic questions regarding the new function exploring: 1) when they invoke the price review/edit; 2) the process they use once the review screen is displayed (ignore, recall companies, etc); 3) their understanding of the information provided on the review screen; 4) the navigation between the review screen and the main survey screen for error correction; 5) other comments or suggestions regarding the edit function and the review screen.

Figure 1. Main Screen

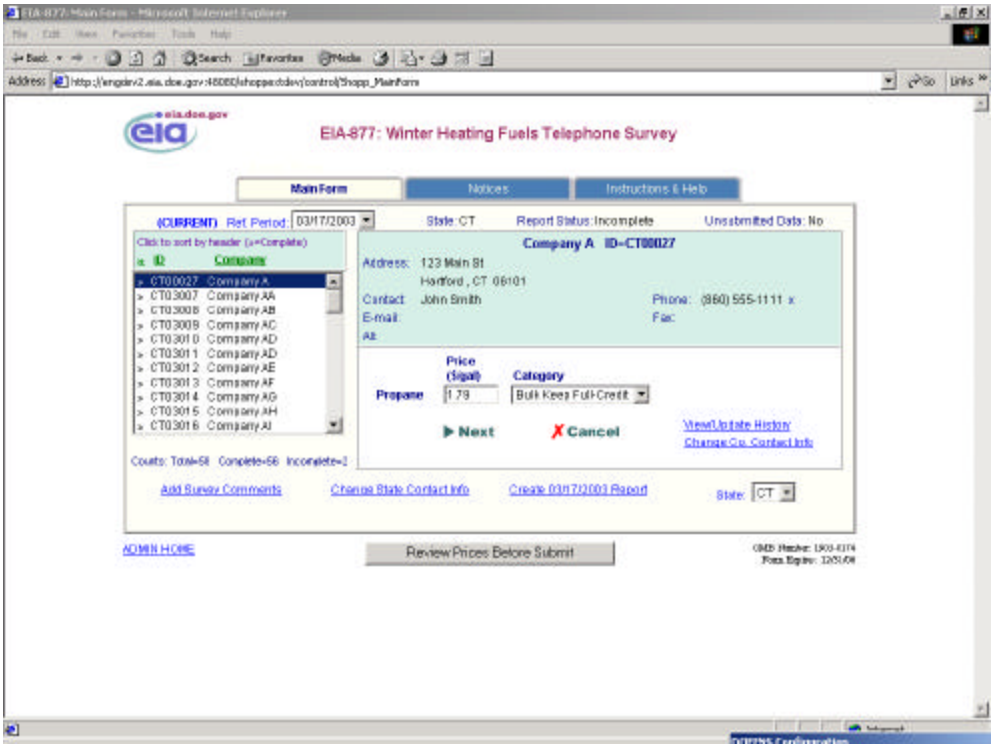
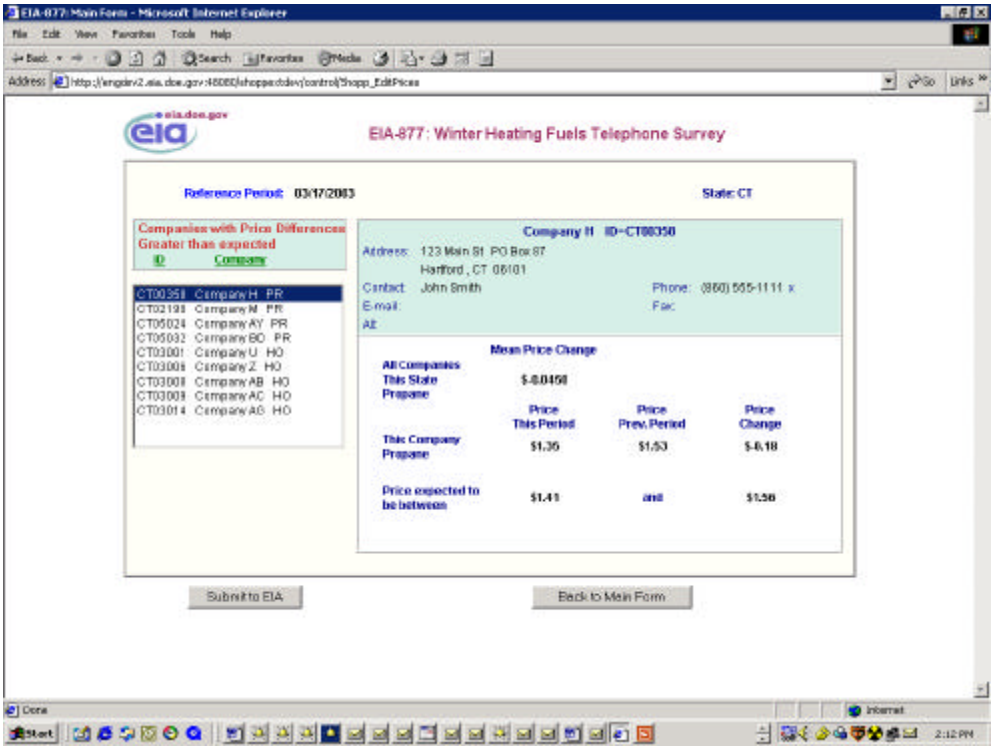


Figure 2. Review Screen



10. In order to better understand the process flow, two screen prints are provided. The first screen (figure 1) shows the data entry screen, the main screen. The respondent selects a company from the dialogue box on the left, and the company's information appears on the right side of the screen along with the boxes to enter the company's price and category of sale. After entering a price, the respondent selects the next company from the left dialogue box, enters their price, etc. Once all the companies' prices have been entered, the respondent clicks on the "Review Prices Before Submit" button located on the center bottom of the screen.

11. When the respondent clicks on this review button, the edit failures are displayed in the second screen (figure 2) shown. On the left side of the screen, the companies whose prices failed the edit are displayed. As the respondent clicks on a company, information about that company is displayed to the right. This information includes the mean price change for all companies in that State, the previous and current period's price, and the price change since the last period for the company selected, and the price interval the company was expected to fall within for the current period. To make a change to the data, the respondent clicks on the "Back to Main Form" button shown in the figure at the bottom right to return to the data entry screen. If no data changes are needed, the respondent clicks on the "Submit to EIA" button at the bottom left to send the data to EIA.

12. Of the respondents who returned the questionnaire, most respondents (86.7%) indicated that they ran the edit after all prices had been entered, just prior to submission of the data, as shown in Table 2. However, a few respondents (12.5%) indicated that they invoked the edit when an individual company's data seemed anomalous. Similarly, a few respondents indicated they used their own method outside the system to review the data, frequently making use of a longer historical series than just the previous period, or compared price changes to a fixed amount they had set. The respondents also varied as to their process for reviewing the prices flagged by the IDC edit.

Table 2. Edit Function Questionnaire Results

Q1. When invoked (can check more than one):	
After each entry	0%
After some but not all entries	12.5%
After all entries	86.7%
Other	20.0%
Q2. Process for review flagged entries (can check more than one):	
Ignore	25.0%
Review one-by-one, calling and correcting	25.0%
Review by noting outside system, call all, correct all, review again, then submit	25.0%
Review by noting outside system, call all, correct all, then submit (no final review)	6.7%
Other	25.0%
Q3. Understand:	
Why failed	80.0%
Mean change	80.0%
Mean includes only entered data	73.3%
Flagged company information	86.7%
Expect company value	86.7%
Q4. Is navigation to/from review screen a problem?	
No	100%

13. **Finding One**: Approximately 25% of the respondents reported that they ignored the information provided regarding the failed prices. Another (25%) reviewed the information for each failed company one at a time, verifying the information and correcting as necessary by returning to the main screen, before proceeding to the next flagged company. On the other hand, 31% of the respondents make note (outside of the system) of the edit failure information provided and then follow-up. Of particular concern is the finding that 25% of the respondents ignore the edit, and 7% perform no second review of edit failures after corrections are made in the main file, prior to submitting the data, despite the fact that that the re-edit information is presented to them on the submit (review) screen.

14. **Finding Two:** Also of interest is the finding that 87% of the respondents understand the information regarding the particular company's price that failed, and, in particular, the information regarding the company's reported price this and last period, and also understand the expected interval for the company's price for this period. Yet, only 80% understand why the price failed the edit, and understand the information on mean price change of all companies they report for. Furthermore, only 80% understand that the mean price change represents only the prices that were entered at the point at which the review button was hit. These last two findings can be used to further interpret the 87% --- 87% understand that the price is expected to fall within the specified interval, but a few of those respondents do not understand *why* the price should be within the interval, and therefore, do not understand *why* the price failed.

15. **Finding Three**: All the respondents reported that the navigation between the review screen and the main form (in order to correct a price) was not a problem for them. This finding was curious in view of the first finding that 20% of the respondents indicated that they flip back and forth between the main screen and the review screen for each company (after having entered all or most of the data), and that 33% of the respondents record information from the review screen to another location outside the system to further evaluate the data.

16. Each time the respondent selects the "Review Prices Before Submit" button, the edit failures that are displayed are also written to an error log. This log records for each edit failure the individual company's data, the reference period, the time the edit was invoked, and whether the price failed as too

high or too low. This log was analyzed for the first twelve reference periods. A summary of the edit failures written to the log is shown in Table 3.

17. As mentioned previously, each time the respondent clicks the "Review Prices Before Submit" button, each of the edit failures is written to the error log, regardless of whether the edit failure had been written before, as long as the price still failed the edit rule. Therefore, the number of failed records shown in Table 3 reflects the number of times the button was hit times the number of failures at that time. Even though this created a large number of virtually duplicate records written, that characteristic of the log made it useful for tracking and measuring the respondents' process flow. The number of first time failures shown in Table 3 was derived from the log to measure the set of unique edit failures. If the same set of failures occur each time the review button is clicked, then the ratio of the number of failed records to unique records indicates the number of times respondents went to the main screen from the review screen, and returned to the review screen. Across both products, respondents clicked on the review button an average of 5.7 times per reference period, but the rate by respondent and product varied from a low of 1.5 times to a high of 11.8 times review clicks, averaged across reference periods. These respondent rates for changing screens were further compared to the respondent average rate for edit failures to shed light on the respondent's process for invoking the edit and resolving the edit.

Table 3. Summary of Edit Failure Log

PRODUCT	PRICE HI or LO	Data	Total	Avg/wk.	Avg # Failures	Avg. # Screen Changes	Screen Changes/ Failure
Heating Oil	H	# Failed records	3478	289.8	13.2		
		# Changed and fail again	1		0.0		
		# First time failures	584	48.7	2.2	6.0	2.7
	L	# Failed records	4410	367.5	16.7		
		# Changed and fail again	5		0.0		
		# First time failures	728	60.7	2.8	6.1	2.2
Heating Oil # Failed records			7888	657.3	29.9		
Heating Oil # Changed and fail again			6	0.5	0.0		
Heating Oil # First time failures			1312	109.3	5.0	6.0	1.2
Propane	H	# Failed records	3841	320.1	13.3		
		# Changed and fail again	19	1.6	0.1		
		# First time failures	705	58.8	2.4	5.4	2.2
	L	# Failed records	2722	226.8	9.5		
		# Changed and fail again	13	1.1	0.0		
		# First time failures	504	42.0	1.8	5.4	3.1
Propane # Failed records			6563	546.9	22.8		
Propane # Changed and fail again			32	2.7	0.1		
Propane # First time failures			1209	100.8	4.2	5.4	1.3
Total # Failed records			14451	1204.3	50.2		
Total # Changed and fail again			38	3.2	0.1		
Total # First time failures			2521	210.1	4.4	5.7	1.3

The average respondent failure rate across both products was 4.4, compared to the 5.7 screen change rate, but again, the failure rate varied by product and respondent from a low of 1.4 failures to a high of 10 failures averaged across reference periods. The average number of screen changes per edit failure is displayed in the last column of Table 3. This shows a rate of 1.3 screen changes per failure overall, indicating that on average, the respondents return to the main screen one or more times per failure. This finding at first would lead one to think the respondents are viewing the edit failures one at a time and returning to the main screen to apply a correction. However, in general, this is not the case, because when the respondent returns to the review screen, the same edit failures with no data changes appear for

the most part. What the log of failures can not show, however, is the possibility that respondents return to the main screen and enter new data that do not fail nor impact the previously failed data sufficiently to change their edit failure status.

18. The screen change rates per edit failure were also compared to the findings from the questionnaire. Comparing the screen change to failure rates to the responses to question 2, we find for the respondents who answered that they ignored the information provided on the review screen, their screen changes to failure rates, lower than most, as determined from the log ranged from .2 to 1.3 failures per response period, across the products. Clearly, these respondents do not go back and forth from review to main for each edit failure, but they are using the review button more than would be expected, given that they said they ignored the information.

19. Similarly, we can compare information for the other extreme of respondents who reported on the questionnaire that they reviewed the information for each failed company one at a time, verifying the information and correcting as necessary by returning to the main screen, before proceeding to the next flagged company. This would imply screen changes to failure rates of 1.0 or greater. These respondents would be expected to have the largest screen changes to failure rates. The log showed that one of these respondents had a rate less than expected of only .6, but the remaining respondents ranged from 1.2 to 3.0 screen changes per edit failure across the two products.

20. Those respondents that reported making a note (outside of the system) and then following-up on the edit failures had, in general, the highest rates, ranging from 1.1 to 5.1 screen changes per failure across the products.

21. The concern regarding the finding that only 80% of the respondents to the questionnaire understood that the mean price change represents only the prices that were entered at the point at which the review button was hit was investigated. Review of the individual respondent logs for the 20% that did not understand showed that most of these respondents (75%) had very high screen changes per edit failure and showed evidence of editing data based on partial information, thereby effecting the edit rule and resulting failures. The logs for these respondents did not show the same set of edit failures each time the review screen was clicked. As expected, this was particularly true of the subset of respondents who had also answered on the questionnaire that they invoked the edit after each/some entries were made, but not all entries.

21. Further study of the edit failure log by reference period by create date revealed that a substantial number of the records were created after preliminary estimation by the survey, and some even after final estimation (one week after the preliminary estimate). The separation of failures by preliminary and final estimates' dates is now highlighted as an area of future study. Future research will also include an examination of the logs to determine the efficacy of the edit rule, and examination by company to highlight repeat failure companies to discover whether the data are repeatedly reported erroneously, or the edit rule is not appropriate for them.

IV. CONCLUSIONS

22. While substantial progress has been made in EDR, both in the number of surveys providing the option and the increased usage of the option by respondents, significant work remains in developing and EDR editing strategy. The strategy must recognize the balance of possibly conflicting quality goals of maximizing the use of the EDR option by respondents, and minimizing the errors on the submitted data. The EDR strategy must take into account when/if to use hard or soft edits, when to invoke the edit in the data entry process, how to present messages regarding the edit failures, and how to navigate efficiently to correct errors and/or submit the data. The use of cognitive testing and respondent interviews/questionnaires are useful in designing an EDR editing strategy or improving an EDR editing approach already in use.

23. EDR expands the "self-administered" role of the respondents to include their interaction with the edit process. As a result, new indicators on the performance of the edit process must be constructed and analyzed. As demonstrated in this case study, logs generated by the respondent actions are useful not only in measuring the performance of the edit rules, or to validate information obtained by cognitive testing or questionnaires, but just as importantly, as hard evidence on how the respondents actually use the edit process. This case study demonstrated that understanding the respondent process of how and when the edit is invoked, and how and when the failures are resolved may impact the resulting edit failures, which in turn effect the quality of the data.

References

[1] Best, Sam (2004): "Implications of Interactive Editing", FCSM-GSS Workshop on Web Based Data Collection, Washington, D.C.

[2] Nicholls, W.L. II, Baker, R.P., & Martin, J. (1997): "The Effect of New Data Collection Technologies on Survey Data Quality" in L. Lyberg, P. Biemer, M. Collins, C. Dippo, N. Schwarz, & D. Trewin (editors) *Survey Measurement and Process Quality*. New York: Wiley.

[3] Weir, Paula (2003): *Electronic Data Reporting—Moving Editing Closer to the Respondent*, UN/ECE Work Session on Statistical Data Editing, Madrid.

[4] http://www.eia.doe.gov/oil_gas/petroleum/survey_forms/pet_survey_forms.html

[5] http://www.eia.doe.gov/cneaf/electricity/page/forms.html

[6] http://www.eia.doe.gov/cneaf/electricity/edc/contents.html

---