Evaluation of Selective Editing for the US Census Bureau Foreign Trade Data[1]

Maria M Garcia[2], Andreana Able[2], Christopher Grieves[2]
[2]US Census Bureau, 4600 Silver Hill Rd., Washington D.C., 20233

Abstract
The Foreign Trade Division at the US Census Bureau is responsible for the production of monthly import and export statistics. Data processing begins with extensive micro-editing using an automated editing and imputation system; records that do not pass the edits are automatically imputed. However, imputation may not be successful for a small portion of the edit-failing records. Records for which imputation is not successful are sent to subject matter experts for manual review. Commodity analysts use their expertise to manually adjust failing records and must review a large number of records under tight time constraints before the publication of monthly statistics deadline. Due to the time and resource constraints, we have an ongoing effort to improve the current procedures while preserving (or improving) data quality. We investigate selective editing strategies for these data. Selective editing requires a score function to assign a priority ranking to erroneous records. We propose a score function that includes a measure of how suspicious a record is and a measure of the anticipated error in the suspicious record. We present methodology and results of an evaluation study of the application of selective editing to foreign trade export data.

Keywords: editing, score functions, selective editing

1. Overview of Foreign Trade Data

The Foreign Trade Division (FTD) at the Census Bureau processes monthly import and export transactions for the shipment of merchandise between the United States and its international trading partners and is responsible for publishing the official international trade statistics for the country. Foreign trade transactions are primarily filed via an online internet data collection system, mostly through the US Customs and Border Protection. The collection of these data is unusual at the Census Bureau because they are filed upon arrival or departure of merchandise goods and are not based on surveys or censuses that are sent to respondents soliciting responses.

Data items collected include commodity, country of origin or destination, port of arrival or dispatch, value, quantity, and shipping weight. Data processing begins with extensive micro-editing using the division's automated editing and imputation system that uses a parameter file called the Edit Master. The Edit Master verifies that numeric data fall

---

[1] This report is released to inform parties of ongoing research and to encourage discussion. The views expressed are those of the authors and not necessarily of the US Census Bureau.
[2] U.S. Census Bureau, 4600 Silver Hill Rd., Washington D.C., 20233

within the prescribed ranges and that the ratios of highly correlated items fall within prescribed commodity bounds. Records that do not pass the edits are automatically imputed. However, imputation may not be successful for a small portion of the edit failing records. Records for which imputation is not successful are marked as “rejects” and sent to subject matter experts for manual review. The analysts use their commodity expertise to manually adjust rejected records. They may also call back filers in an attempt to correct erroneous data. The commodity experts review a large number of records under tight time-constraints before the publication of monthly statistics deadline. Due to the time and resource constraints, the division has an ongoing effort to improve the current procedures while preserving (or improving) data quality. We investigate the application of selective editing to these data. Selective editing requires a score function. In this paper we present a score function that includes a measure of how suspicious a record is and a measure of the anticipated error in the suspicious record.

In this section, we provided background on the Census Bureau’s foreign trade data editing procedures. In Section 2, we present background on selective editing. In Section 3, we present the score functions and in Section 4 we present results of the evaluation study along with our recommendations. We close with a short summary in Section 5.

2. Overview of Selective Editing

Traditionally, the goal of data editing is to identify all errors in the data with the expectation of producing good, high quality statistics. Resolution of edit failures requires imputing erroneous data with plausible values; records for which imputation fail may require manual resolution. Manually following-up edit failures is costly, consuming a large amount of data editing resources. The purpose of selective editing is to reduce costs by concentrating the review effort on erroneous units with a large potential effect on publication estimates.

Research has shown selective editing methods’ potential for reducing editing costs without affecting the quality of the final publication estimates. Lindell (1997) reports on a study in which the highest ranked 20 percent of the erroneous records contribute to 90 percent of the total adjustment. Granquist and Kovar (1997) report on producing savings of 50 percent or more of the total editing cost while having a small impact on the final publication. Most statistical agencies include selective editing methods as part of data editing processing since the 1990s. Latouche and Berthelot (1992) developed score functions for an annual retail trade survey and Lawrence and McDavitt (1994) presented a score function for a quarterly average weekly earnings survey. Jäder and Norberg (2005) developed a score function for the Swedish foreign trade survey. Garcia and Bartha (2012) presented score functions along for the Census Bureau foreign trade data. Although there are a large number of studies describing applications of selective editing within statistical agencies, a theory to justify selecting only a small subset of the erroneous observations for further scrutiny is a more recent development. Arbués et al. (2011) set up selective editing as an optimization problem where the objective is to minimize the expected workload subject to minimal expected error on the aggregates. Their study provides a theoretical framework for selective editing.

Selective editing requires a score function to rank records; records with a score higher than a preset cut-off value are prioritized for manual review according to their score. All other records are accepted as reported (not edited) or edited using an automated system. The overall objective is to spend manual review resources on suspicious records that have a significant impact on the estimates without affecting overall data quality. In the next section we present score functions for selective editing of these data.

3. Selective Editing for Foreign Trade Data

In this section we present score functions for selective editing of these data. Our aim is not to re-engineer the current editing procedures; our goal is to include selective editing strategies as part of the data review process. Data processing using the Edit Master parameter file begins with ratio tests involving items Value, Quantity, and Shipping Weight denoted by variables V, Q, and SW, respectively. For any given shipment, the ratio of value to quantity $p = {}^{V}\!/_{Q}$ denotes the unit price of the shipped merchandise. For what follows, we focus on the unit price ratios only.

3.1 Hidiroglou-Berthelot method

The Hidiroglou-Berthelot edit (Hidiriglou and Berthelot, 1986) uses historical ratios to detect outlying observations in periodic data. For the trade data, it is not possible to use the method as described by Hidiroglou and Berthelot because current month data is not comparable to previous month data. The Hidiroglou-Berthelot method must be adapted for application to current ratios. For every record i we test ratios of two items in the current month data file.

We begin with the current month unit price ratio $p_i = {}^{V_i}\!/_{Q_i}$ and the median of unit prices p_{q_2}. Since the unit prices p_i are ratios, we transform them to ensure identification of outliers at both tails of the distribution of ratios. We use the transformation suggested by Hidiroglou and Berthelot,

$$S_i = \begin{cases} \dfrac{p_i}{p_{q_2}} - 1 \;\; if\ p_i > \ p_{q_2} \\ \\ 1 - \dfrac{p_{q_2}}{p_i} \;\; if\ 0 < p_i \leq \ p_{q_2}. \end{cases}$$

Note that this transformation centers the distribution of ratios about zero; it does not provide a symmetric distribution of the unit price ratios.

With economic data, we wish to ensure we are tracking errors associated with large units that affect many statistics the most. Hidiroglou and Berthelot (1986) suggest applying another transformation to exercise control over the influence of the magnitude of the data. They suggest comparing current month data to previous month data to ensure we place more importance on small deviations within large units, as opposed to large deviations within small units. However, their method applies to periodic surveys using previous and current data to identify suspicious units. This is not the case with the trade data. In these data, for most commodities previous month data may not be available or comparable to current month data; companies may have m number of shipments the current month and

$n \neq m$ (or no shipments) the previous month. The magnitude transformation must be adapted to using only current month data.

When using only current month unit prices, we use the median of unit prices and reported data to estimate an anticipated value for current month data (see Garcia and Bartha, (2012)). Since we are dealing with current month unit price ratios ($p_i = {}^{V_i}/_{Q_i}$), we use $p_{q_2} \times Q_i$ as the best possible anticipated value for V_i, with p_{q_2} denoting the median of current month unit prices as defined above. The size transformation adapted to current month unit price ratios is,

$$E_i = S_i \times \{max\left(V_i, p_{q_2} \times Q_i\right)\}^u, where\ 0 \leq u \leq 1.$$

In this application, we use the square root in the maximization part of the size transformation (i.e. $u = 0.5$).

Given the transformed unit price ratios E_i, we compute the first quartile, the median and the third quartile, denoted by E_{q_1}, E_{q_2} and E_{q_3}respectively. Using these statistics, we calculate a measure of the deviation of the first and third quartile of the transformed unit price ratios from the median as

$$d_{q_1} = \max(E_{q_2} - E_{q_1}, \left|a \times E_{q_2}\right|)$$

$$d_{q_3} = \max(E_{q_3} - E_{q_2}, \left|a \times E_{q_2}\right|).$$

We assign to every observation a score that is the ratio of the displacement of the transformed unit prices from the median and the appropriate distance from the median as measured by d_{q_1} and d_{q_3}:

$$Ratio_i = \begin{cases} \dfrac{E_{q_2} - E_i}{d_{q_1}} & if\ E_i < E_{q_2} \\ \\ \dfrac{E_i - E_{q_2}}{d_{q_3}} & if\ E_i > E_{q_2}\ . \end{cases}$$

The term $\left|a \times E_{q_2}\right|$ in the calculation of the distances ensures that d_{q_1} and d_{q_3} are not too small for observations clustered about the median. After testing with different values for a, we used the value suggested by Hidiroglou and Berthelot ($a = 0.05$) as it has worked well in our application.

3.2 Combine suspicion and potential error

We could think of the score function *Ratio* as a measure of how suspicious a record is (i.e., a measure of the probability that the record is erroneous). Garcia and Bartha (2012) developed score functions that also include a measure of the relative effect errors in the erroneous record have on the final totals, adapted from the *Diff* function described by Latouche and Berthelot (1992). Their study used a small subset of the foreign trade exports data file. They compute the expected error in the variable V as the absolute difference between the current month's reported value of V and an anticipated value for the current month's value of V. Using the product of median unit price ratios and quantity ($p_{q_2} \times Q_i$) as an anticipated value for V , the potential error is

$$Error_i = |V_i - p_{q_2} \times Q_i|$$

and the *Diff* score is

$$Diff_i = \frac{Error_i}{T_V},$$

where T_V represents the estimated total value of V calculated using current month data.

We found earlier in this project that the measure of potential effect, *Diff*, does not work as well when using the full data set. In this case, *Diff* becomes very small as the estimated total T_V becomes very large. Thus, we decided to use the measure of the potential error rather than the effect errors have on final totals. In addition, we assign to every observation an importance weight according to the weighting scheme described in Garcia et al. (2008.) For each record *i*, we compute a composite global score function based on how suspicious a record is ($Ratio$), the size of the expected error (*Error*), and the importance weight for the observation.

$$Score_i = Ratio_i \times Error_i \times Weight_i$$

Jäder and Norberg (2005) proposed a score function that is somewhat similar to *Score*. However, their measure of suspicion uses the interquartile range rather than d_{q_1} and d_{q_3}.

3.3 Calculate *Ratio* using Quartile method

The first transformation in the Hidiroglou-Berthelot method is an attempt to account for outliers at both ends of the distribution of unit price ratios. We can also use the log transformation to make the distribution of ratios more symmetric as in the quartile method. We begin by applying this transformation to the data, calculating the quartiles of unit price ratios using the log transformed unit price ratios instead of the E_i described above before computing d_{q_1}, d_{q_3} and *Ratio*. The distances $d_{q_1}$and d_{q_3} are defined as

$$d_{q_1} = \log(p_{q_2}) - \log(p_{q_1})$$

$$d_{q_3} = \log(p_{q_3}) - \log(p_{q_2}) ,$$

and the new variation for *Ratio* is computed as

$$Ratio1_i = \begin{cases} \dfrac{\log(p_{q_2}) - \log(p_i)}{d_{q_1}} & if \ \log(p_i) < \log(p_{q_2}) \\[2ex] \dfrac{\log(p_i) - \log(p_{q_2})}{d_{q_3}} & if \log(p_i) > \log(p_{q_2}) \end{cases} .$$

As in the Hidiroglou-Berthelot method, there may be commodities for which the median of unit prices p_{q_2} is too close to either the upper and/or lower quartile. In this case the denominator in the distances d_{q_1} and/or d_{q_3} is close to zero and is replaced by a small constant or a fraction of the median.

4. Evaluation Study

Garcia and Bartha's (2012) feasibility study was limited to a small subset of the 2004 exports trade data files. This evaluation study uses a larger, more recent data file. In addition, we investigate the possibility of identifying highly suspicious records earlier in the editing process, without the use of parameters. In this case, the strategy is to start editing processing by assigning a score to incoming records. Records with a score higher than a pre-set cut-off value will be marked for manual review; all others records are handled using the existing automated editing system. In this section, we describe the data file, the SAS program, and present results of the evaluation study.

4.1 Evaluation Data and Computer Program

We used four consecutive months of reported (raw) and final (edited) export trade data (October 2010 through February 2011). We also have available the Edit Master parameter file; the Edit Master contains editing parameters for each commodity with the necessary information for how to compute the ratios.

There are approximately 8,000 commodity classifications used to collect and compile export statistics. We note that the number of records within a commodity is important, as there must be a sufficient number of records within each commodity to calculate the simple statistics needed for assigning scores. If the number of records is too low, outliers in the distribution of ratios may be included in the computation of quartiles needed to assign scores. We decided to include only commodities having at least 30 records in the data file. Our final evaluation data file has over 3.25 million records for four consecutive months, with commodities having at least 30 records within the month.

We wrote a series of SAS macros for implementing selective editing. Starting with the Edit Master parameter file and trade records file as input, the program first extracts the commodities that are usable for selective editing according to the guidelines described above. It then merges the usable data to the Edit Master and then uses the Edit Master's instruction to compute the unit price ratios. Once the ratios are available the program calls separate macros to compute measures of suspicion for each record using either variation of *Ratio*. The user can decide which measure of suspicion to use. The program computes the measure of anticipated error and assigns importance weights within the main SAS macro. We combine the measures of suspicion and anticipated error into a global score for the record. The final step assigns to each record a priority ranking according to its score. We present results using *Ratio* to calculate suspicion; see Garcia and Bartha (2012) for results using the quartile method to calculate *Ratio* in *Score*.

4.2 Evaluation Results

Latouche and Berthelot (1992) suggest a simulation study to determine the effect of different percentage levels of review on the bias in the target parameters. They defined the absolute pseudo-bias as a measure to estimate the bias due to errors remaining in the target parameters after processing $p\%$ of the records through selective editing. In this evaluation, the target parameter is the total for a given variable. The absolute pseudo-bias estimates the relative discrepancy between the final publication total T_F, and T_{SE}, the selective editing estimated total, as $|T_F - T_{SE}|/T_F$.

To calculate the absolute pseudo-bias we proceed as follows:

1. Run the selective editing program to rank raw data records according to their scores.
2. Simulate that only the top ranked $p\%$ of records are flagged for analyst's review by replacing the raw data values of the top $p\%$ of the records with the available edited values while keeping raw values for records with a score lower than the chosen $p\%$ review level cut-off value.
3. For each $p\%$ review level, calculate selective editing item totals using the $p\%$ edited data and final item totals using the final edited data.

We graphed the absolute pseudo-bias versus the percentage number of records ($p\%$) flagged for review for several commodities at different levels of aggregation. The horizontal axis displays the percentage number of records corrected after ranking the records from the most to the least influential according to their score. Figure 1 displays the absolute pseudo-bias for the variable Q (quantity) for "*Automotive, diesel, or marine engine lubricating oils*" exported in January 2011. The graph illustrates how the absolute pseudo-bias rapidly decreases as the percentage number of records flagged for review increases and reviewing more than 15 percent of the highest ranked records does not affect the final estimated total. We could stop reviewing records at the 15 percent level of review when the effect of changes on the absolute pseudo-bias approaches zero and the estimated total approaches the final publication total.

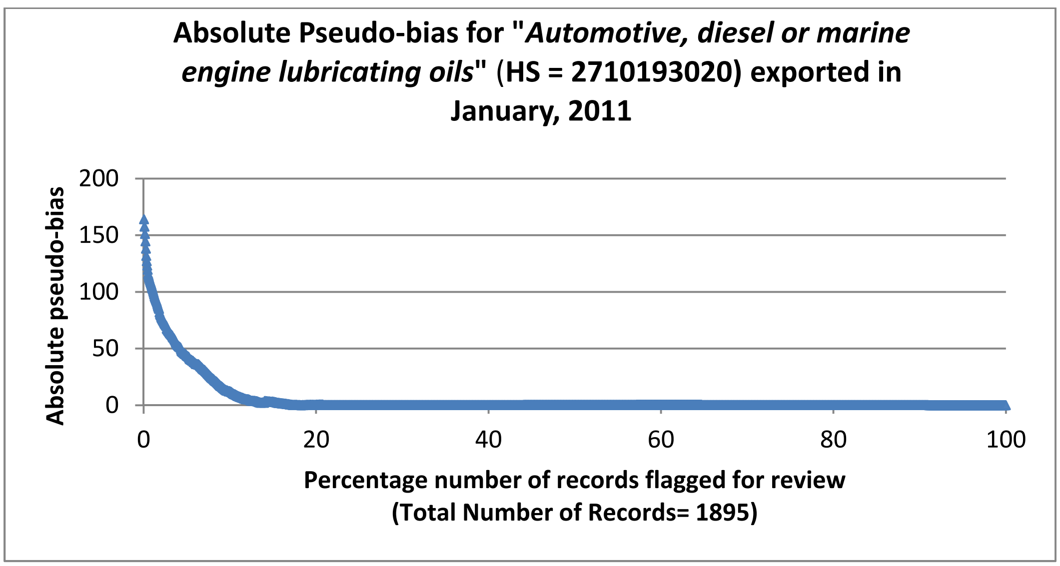

Figure 1: Absolute Pseudo-bias for "*Automotive, diesel, or marine engine lubricating oils*"

We observe a similar pattern for the commodity "*Meat of swine, fresh or chilled, hams and cuts thereof, with bone-in, processed*". Figure 2 shows the absolute pseudo-bias approaching zero as the percentage number of records flagged for review increases, and there are no significant changes in the amount of pseudo-bias remaining in the estimated total once the percentage follow-up level hits the 10 percent mark.

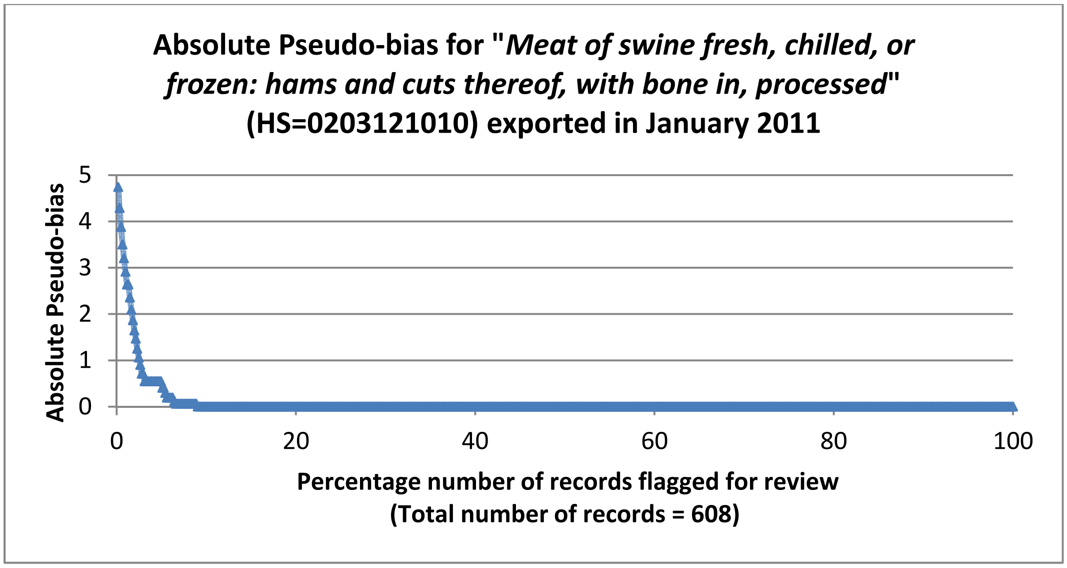

Figure 2: Absolute Pseudo-bias for "*Meat of swine, fresh or chilled, hams and cuts thereof, with bone in, processed*"

Our next example illustrates the case in which the decrease in absolute pseudo-bias is not as fast as that displayed in Figures 1 and 2. Figure 3 displays the absolute-pseudo-bias versus the percentage number of records review for the commodity "*Meat of bovine, fresh, chilled or frozen*" exported in January 2011.

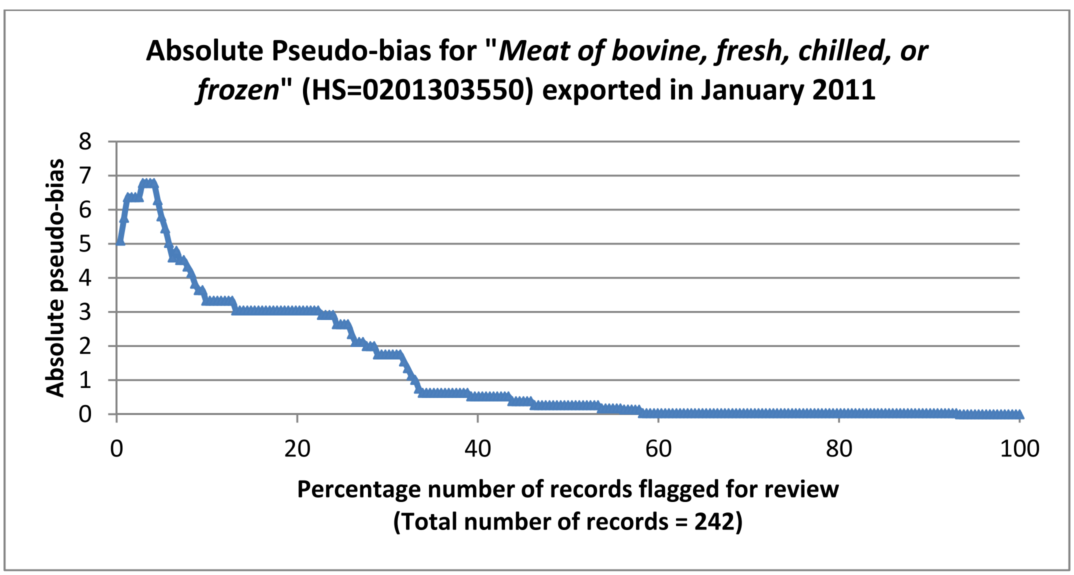

Figure 3: Absolute Pseudo-bias for "*Meat of bovine, fresh, chilled or frozen*"

In this case, the absolute pseudo-bias approaches zero when the percentage number of records flagged for review approaches 55 percent of the highest ranked records. This is not as encouraging as the pseudo-bias plots displayed in Figures 1 and 2; however, it verifies the manual correction of **all** flagged errors does not have a large effect in the final estimates

There is a caveat in this analysis: there are too many commodities. Thus, we cannot effectively determine review level percentage cut-off values by assessing the behavior of the absolute pseudo-bias by commodity, although we could determine cut-off values by measuring the absolute pseudo-bias at higher levels of aggregation.

Figure 4 displays the absolute pseudo-bias for the domain "*Meat of swine, fresh, chilled, or frozen - hams, shoulders, and cuts thereof*" (six-digit aggregation level) and Figure 4 displays the absolute pseudo-bias for the domain "*Meat of swine, fresh, chilled, or frozen*" (four-digit aggregation level). These domains include the commodity "*Meat of swine, fresh or chilled, hams and cuts thereof, with bone-in, processed*" (displayed in Figure 2). At these higher levels of aggregations, we observe a similar pattern: The slope indicates a fastest decrease of the absolute pseudo-bias as the percentage number of records corrected increases. The decrease is so fast that we graphed the pseudo-bias for the top ranked records only.

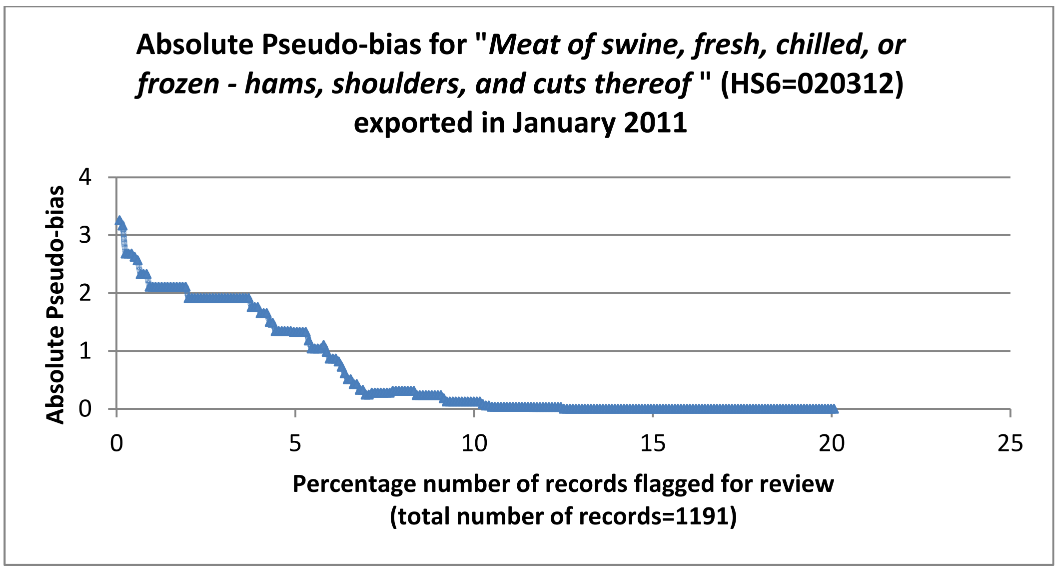

Figure 4: Absolute Pseudo-bias for "*Meat of swine, fresh, chilled, or frozen, hams, shoulders and cuts thereof*"

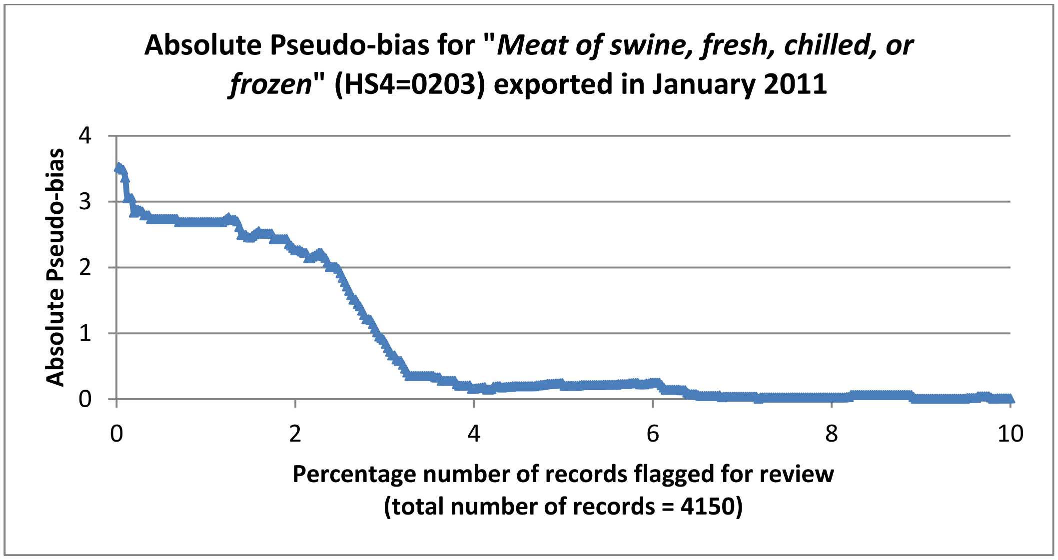

Figure 5: Absolute Pseudo-bias for "*Meat of swine, fresh, chilled, or frozen*"

4.2.1 Balancing low pseudo-bias and low analysts' referrals

Although we could determine review levels cut-off values using the absolute pseudo-bias, there are some concerns. Ideally, we would like records ranked as highly suspicious according to their score to correspond to records marked as rejects by the automated system. Moreover, we want to balance a fast drop in pseudo-bias with the overall objective of keeping the total number of records flagged for manual review (i.e. ranked as highly suspicious) low. Thus, we look at the following two statistics:

1. Proportion of records ranked as highly suspicious that are rejects (true rejects)
2. Proportion of records ranked as highly suspicious that are "passed" records (false rejects)

The proportion of records flagged as highly suspicious that are true rejects is a measure of how well the score function is able to track erroneous records. In this analysis, the proportion of records that passed the edits includes erroneous records for which imputation is successful; we called these records "passed" records. The reason we include automatically imputed records in the universe of passed records is that automatic imputation does not require additional manual review thus we treat these records as passed records when calculating these proportions.

Figure 6 displays the percentage number of records flagged for manual review that are true rejects representing an aggregate of commodities in the "Chemicals" sector (Section 2) for January 2011. The horizontal axis displays the percentage review levels and the vertical axis represents the percentage rejects flagged under each percentage review level. Note that the percentage review levels we display in the horizontal axis are small. This is because we look at the number of records flagged for analyst referral as a measure of analysts' workload. Given the available resources, we set the ideal follow-up level at 0.3 percent, which is the approximate percentage number of rejects when using the existing production system. We start by calculating these rates at very low review levels, and slowly increased this review rate up to and including a 10 percent review rate. The graph shows selective editing correctly flagged 28 percent of January Section 2 rejects for analysts' review at the 0.3 percent review level cut-off and 33 percent at the 0.5 percent review level. The proportion of rejects tracked correctly increases as the percentage of records flagged for review increases, with more than 75 percent of rejects flagged at the 10 percent review level.

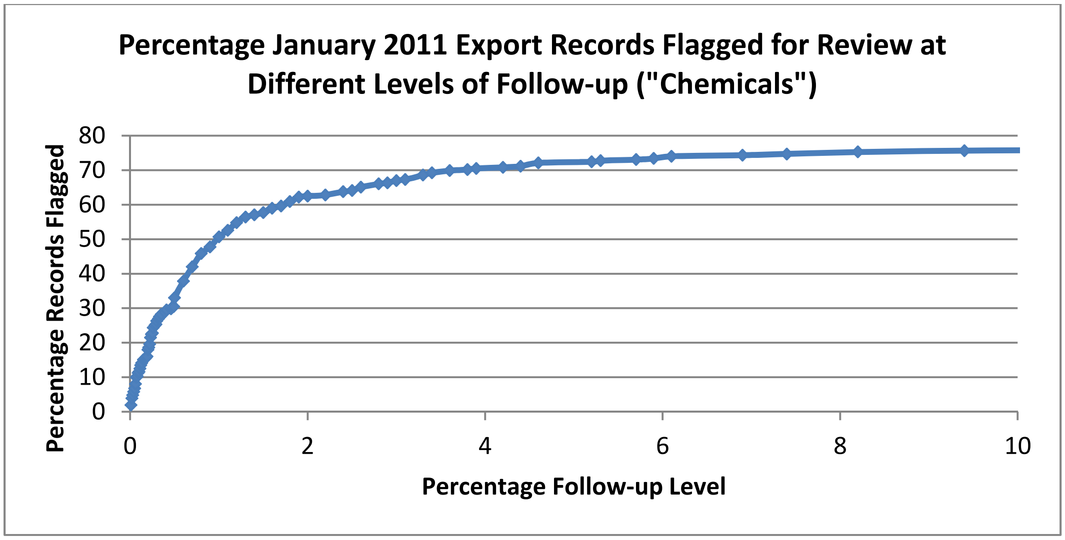

Figure 6: January 2011 Export Records Flagged at Different levels of Review by Selective Editing

Figure 7 displays the corresponding graph for November 2010 export of "Metals" (Section 5). In this case, selective editing performs better. It correctly identifies about 33% of rejects for analysts' review at the 0.3 percent review level and 50 percent of rejects at the 0.5 percent review level and more than 80 percent of rejects are flagged correctly at the 10 percent review level.

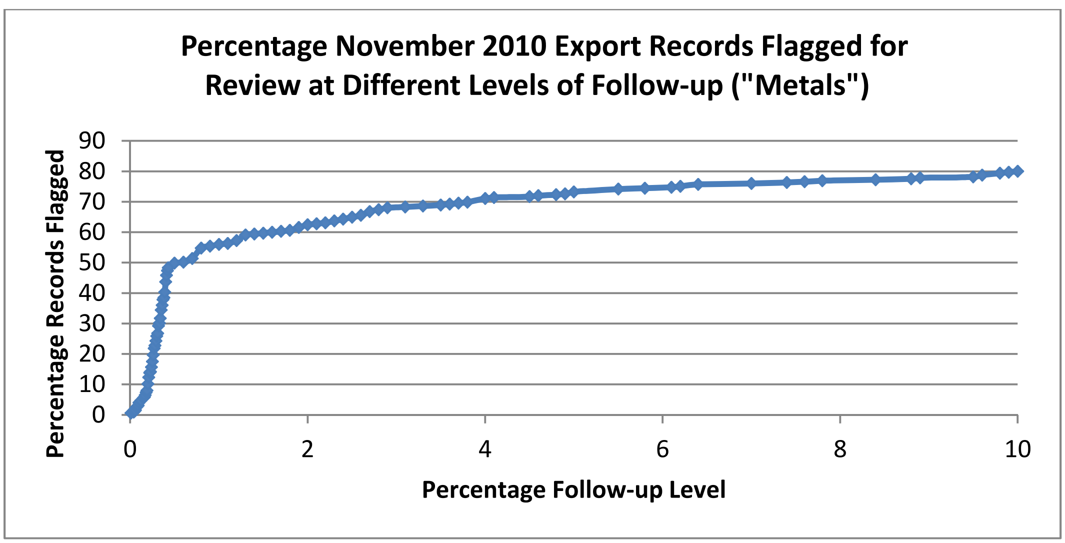

Figure 7: November 2010 Export Records Flagged at Different levels of Review by Selective Editing

Unfortunately, for these data, these measurements (percentage number of records hit) can be misleading. To understand why, let us look at the total number of passed records erroneously identified as highly suspicious at different cut-off levels of manual review.

Table 1 displays the percentage number of true rejects and the number of "passed" records flagged for review at different percentages review rates for commodities in the "Chemicals" sector (classified as Section 2) exported in January 2011. The first column represents cut-off levels of review under a selective editing application. Note that the percentage review levels are small; we start by calculating hit rates assuming 0.1 percent of top ranked records are to be reviewed and slowly increased this review rate up to and including an overtly pessimistic five percent review rate. In column two, we display the proportion of records ranked as highly suspicious according to their score that are true rejects for each review rate. In column three, we display the total number of records that fall under each percentage review rate. Table 2 displays similar information for the total number of records in the "Metals" sector (Section 5) exported in November 2010.

Table 1: Total number of January 2011 export records ranked as highly suspicious by rejects and passed records at different follow-up levels for "Chemicals."

Percentage review rate	Percentage of true rejects ranked as highly suspicious at this review rate	Number of "passed" records flagged for review at this review rate
0.1%	14.7%	210
0.2%	22.4%	348
0.3%	27.9%	515
0.4%	29.5%	628
0.5%	33.0%	845
1.0%	50.6%	1662
1.5%	57.7%	2470
2.0%	62.5%	3268
3.0%	67.0%	4979
5.0%	72.4%	8742

We would like the number of passed records marked as highly suspicious at any percentage cut-off level of review to be small as the number of records flagged for manual review is an indicator of the expected analysts' workload. Tables 1 and 2 illustrate that selective editing flags 72.4 percent of Section 2 rejects and 73.2 percent of Section 5 rejects at the 5 percent review level. This is an indication that the score function is working as expected in terms of correctly identifying rejected records. Unfortunately, the trade-off in the number of passed records flagged for review is not acceptable. We recall that erroneous records imputed by the automated system are included in the universe of passed records as the existent automated system (Edit Master) correctly identifies plausible imputes for these records. Selective editing processing prior to Edit Master processing identifies these records as highly suspicious and includes them in the pool of records flag for manual review. Although the percentage number of passed records flagged for review is small, the actual number of passed records flagged is not acceptable during production when resources and time constraints are crucial. Selective editing is feasible prior to automatic editing, although this places greater demands on the ability of the score function to discriminate between highly suspicious records and erroneous responses not needing manual review (automatic imputation is possible). Thus, our recommendation is to follow the guidelines in our previous study (Garcia, Gajcowski, and Jennings, 2008) and use selective editing to assign a priority ranking for guiding manual review of edit failing records identified as rejects (imputation was not successful) by the Edit Master.

Table 2: Total number of November 2010 export records ranked as highly suspicious by rejects and passed records at different follow-up levels for "Metals."

Percentage review rate	Percentage of true rejects ranked as highly suspicious at this review rate	Number of "passed" records flagged for review at this review rate
0.1%	4.9%	136
0.2%	16.3	203
0.3%	32.6%	252
0.4%	48.3%	289
0.5%	49.8%	404
1.0%	56.0%	864
1.5%	59.7%	1386
2.0%	62.5%	1902
3.0%	68.3%	2946
5.0%	73.2%	4890

5. Summary

In this paper, we presented a selective editing application and associated score functions for the Census Bureau foreign trade data. In traditional selective editing, we use previous period data for developing score functions; this is not possible for these data. We adapted available score functions to using only current period data by calculating anticipated values for the variables. We evaluated the application of selective editing to the full data set prior to automatic editing with our results showing that assigning a score to records prior to Edit Master processing leads to a large number of false rejects. The score function is not as powerful as the set of edit constraints. Thus, it identifies more records as highly suspicious, including truly erroneous records for which automatic imputation is possible. It is possible to compute score values for all records prior to automatic editing, but this places greater demands on the ability of the score function to discriminate between correct and erroneous responses. Our recommendation for a selective editing strategy for these data is a two-flow process in which records first pass through the Edit Master automated system, then records for which automatic imputation is not successful are marked as rejects and assigned a score to provide a ranked list of records for prioritizing manual review. The objective is to spend the bulk of manual review resources on these records ranked as highly influential according to their scores.

Acknowledgements

The authors thank Ryan Fescina and Yves Thibaudeau for their helpful comments.

References

Arbués, I., Revilla, P. and Saldaña, S. (2011). “Selective Editing as a Stochastic Optimization Problem”. UNECE Work Session on Statistical Data Editing, Ljubljana, Slovenia.

Garcia, M., Gajcowski, A., and Jennings, A. (2008). “Selective Editing Strategies for the US Census Bureau Foreign Trade data.” CSRM Research Report RRS2007/20, US Census Bureau, Washington D.C.

Garcia, M. and Bartha, E. (2012). Score Functions for Selective Editing of the US Census Bureau Trade Data. Data. CSRM Research Report RRS2012/10, US Census Bureau, Washington D.C.

Granquist, L. and Kovar, J. (1997). “Editing of Survey Data: How much is enough?” *Survey Measurement and Process Quality*, L. Lyberg, P. Biemer, M. Collins, E. De Leeuw, C. Dippo, N. Schwarz, and D. Trewin (Editors). Wiley, NY.

Hidiroglou, M., and Berthelot, J. (1986). “Statistical Editing and Imputation for Periodic Business Surveys”. *Survey Methodology*, V. 12, No. 1, 1986.

Jäder, A. and Norberg, A., (2005). “A Selective Editing Method Considering Both Suspicion and Potential Impact Developed and Applied to the Swedish Foreign Trade Statistics,” UNECE Work Session on Statistical Data Editing, Ottawa, Canada.

Latouche, M. and Berthelot, J., (1992). “Use of a Score Function to Prioritize and Limit Recontacts in Editing Business Surveys.” *Journal of Official Statistics*, V.8, No. 3.

Lawrence, D. and McDavitt, C.,(1994). “Significance Editing in the Australian Survey of Average Weekly Earnings. *Journal of Official Statistics*, V.10, No.4.

Lindell, K. (1997). “Impact of Editing on the Salary Statistics for Employees in County Council.” In UN Economic Commission for Europe, Statistical Data Editing, vol. 2, Geneva, 2-7.