Examining Blocks With Lister Error in Area Listing

Aliza Kwiat[1]
US Census Bureau
4600 Silver Hill Road, Washington, DC 20323

Abstract
Area listing is usually considered to have the best quality among available methods of developing survey frames. However, there has been very little research done to investigate the errors of area listing. This paper will look into the results when two different field representatives canvass the same sample of blocks. How consistent are their results? Are some blocks or types of housing units harder to list than others? What characteristics are more likely to produce inconsistent listing results?

Key Words: Area listing, error

1. Introduction

The U.S. Census Bureau currently uses area listing methods in several programs. Area listing is used to develop the address list for decennial censuses and to develop the rural portion of the demographic household survey frames. In area listing, the Field Representatives (FRs) start with a list of addresses from the most recent Master Address File (MAF) and make enhancements to the list by adding units, deleting units, and making corrections to existing address information. Information that the lister can see and update on the listing instrument includes address, unit status, and location description.

This type of area listing is usually considered to have the best quality among available methods of developing survey frames. However, there has been very little research done to investigate the errors of area listing. This paper will look into the results when two different FRs canvass the same sample of blocks. How consistent are their results? Are some blocks or types of housing units (HUs) less likely to produce consistent listing results than others? What characteristics are more likely to produce inconsistent listing results?

2. Methodology

For this study, we looked at the 301 tabulation blocks sent out for listing as part of the National Evaluation Sample (Loudermilk and Li, 2009). These blocks were listed from May through August 2007. In order to obtain the National Evaluation Sample, an initial universe was created of unit frame combined blocks which were stratified by region, block size, and percent of United States Post Service's Delivery Sequence File (DSF) adds and of area frame combined blocks which were stratified by permit status, block size, and percent E-911 addresses. The subsample for this study was selected from the block clusters in NES with non-zero DSF growth using systematic probability proportional to size, controlling for census region and mobile home percent category[2].

[1]Aliza Kwiat is a mathematical statistician in the Demographic Statistical Methods Division of the U.S. Census Bureau. This report is released to inform interested parties of ongoing research and to encourage discussion of work in progress. Any views expressed on statistical and methodological issues are those of the author and not necessarily those of the U.S. Census Bureau.

[2] In order to appropriately calculate variances, smaller strata were logically combined.

Two separate FRs were sent out to do a dependent listing[3] of each of these blocks. The FRs did not know that these listing assignments were different from regular listing assignments and we do not have any demographic information about the FRs, only information about the results of their listing. Our study looked at these results. Blocks where both FRs did not successfully complete the listing and those that were sent out with zero units were not included in this study, yielding a final sample of 201 blocks[4].

Following listing, the results were matched by identification number back to the January 2007 MAF extract to get some unit level geographic variables. In addition, certain block level variables were created from a variety of Census Bureau sources.

In order to produce the measures to answer the question, *What characteristics of HUs and blocks make a block (or HU) more likely to yield inconsistent listing results?,* we then created a "consistency rate", defined as the rate that both FRs agree on unit validity – this will be done weighted/unweighted and at the block level (or other levels). The unit validity was defined as having both a valid unit status code[5] and having a valid action[6] taken upon the unit by the FR. This validity definition is consistent with the definition used by current surveys to develop their sampling frame. Thus, it is possible that the FRs were inconsistent about the unit status of a certain HU, but as long as they agreed that the unit status was valid and took validating actions on that unit, that unit would be considered consistent by our definition.

We looked at consistency rates and certain associated variables. We wanted to determine which, if any, characteristics had a meaningful relationship with inconsistency. It is likely that inconsistencies between the two FRs are indicative of units (and blocks) that are harder to list and/or more prone to errors. By finding variables that are closely linked to low rates of consistency, we are hoping to determine what characteristics make a unit (or block) harder to list.

3. Limitations

There was a time lag of up to four months between the two listings done in the block. Some inconsistency can be a result of this time lag.

We did not account for inconsistencies due to differences in FRs' experience, ability, training, etc.

When the FRs disagreed about the validity of the unit, we do not know which FR was accurate, we only know if they were consistent with each other. This is why the statistic of interest is a measure of consistency and not of accuracy.

Because the HUs in a block are somewhat linked, the HU level results are not truly independent. For example, if the two FRs disagreed on the validity of one unit in a four unit building, it is likely that they disagreed about all four units.

[3] Since the MAF is a cumulative inventory of all addresses, past and present, it contains numerous records that cannot be found on the ground and sometimes multiple records for the same housing unit if different sources to the MAF provided different forms of the address. A filter is applied to the MAF to determine which records on the MAF are likely to be found on the ground and which are not. Each user of the MAF can determine their own filter based on their needs and requirements. For dependent listing, both units which do and do not meet the filter rules are sent out.

[4] Measures are weighted up, but weights were not adjusted to account for nonresponse.

[5] A valid unit status code is anything other than demolished, nonexistent, under construction, duplicate, unable to locate, physical merge and other uninhabitable.

[6] A valid action includes adding, verifying, changing or moving a housing unit.

In a dependent listing, the FR can see the existing address information from the listing instrument. While this may bias our results in that the FRs could have been influenced by seeing what was already present on the MAF, both FRs would have seen the same default value.

4. Results

4.1 Overall Results

The sampling methodology produced a sample of 201 sample blocks with a total of 56,420 HU records sent out for listing by the FRs. Figure 1, below, shows that the unweighted consistency rate of the FRs was 87 percent overall. This means that when listing the HUs in our sample, the two FRs agreed about 87 percent of the time as to whether or not a HU is valid, based upon the unit status they give to the HU and their actions overall upon the HU.

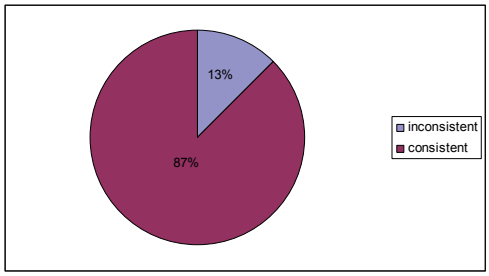

Figure 1: Unweighted Housing Units Consistency

Our sample is a sub-sample of the NES that excludes the sub-universe of no growth blocks and Table 1, below, shows the weighted results from the listing. Combining the cells where the two FRs agreed (in bold) yields a consistency rate of 84 percent (4 percent standard error). This means that when listing growth blocks, the two FRs agreed about 84 percent of the time as to whether or not a HU was valid.

Table 1: Listing Results from Two Field Representatives (s.e.)

		First Field Representative		
		Invalid	**Valid**	**Total**
Second Field Representative	**Invalid**	**44,504,011** (4,572,902)	3,342,896 (663,681)	47,846,907 (4,666,933)
	Valid	6,328,261 (2,605,470)	**6,327,643** (994,930)	12,655,904 (2,695,396)
	Total	50,832,272 (4,256,757)	9,670,539 (1,297,138)	60,502,811 (4,851,559)

There are many variables associated with each HU on the MAF. It is logical to think that there might be a link between some of the variables and a predisposition towards inconsistencies in listing. Some of these variables are block-level variables and some are HU level variables. The consistency breakdowns for a number of these variables are listed in the tables below. Table 2 shows the consistency rates by the four Census regions and they are all between 78 and 92 percent.

Table 2: Housing Unit Consistency Rates by Region

		Consistency Rate (s.e.)
Region	**Northeast**	87% (4%)
	Midwest	90% (4%)
	South	78% (9%)
	West	92% (3%)
	Overall	84% (4%)

4.1.1 Housing Unit Consistency Rates by Housing Unit Variables of Interest

Table 3, below, shows the consistency rates by certain HU level variables of interest. Our data[7] shows that there is a difference in HU level consistency, based upon chi-square tests at the 0.15 significance level, across each of the rows in the table.

The first row of the Table 3 shows that units that were previously categorized as eligible for sampling for American Community Survey (ACS)[8] on the MAF had consistency rates of 85 percent and those that were previously categorized as ineligible had consistency rates of 72 percent. Invalid units are probably harder to list, resulting in inconsistencies. Because we do not know which set of the two listing results were used to update records on the MAF (including unit status and action code) and why, we have no information about whether or not a listed unit was truly valid at the time of listing.

The listing instrument does show the unit status prior to listing and the FRs update that unit status with what they find. Over 99 percent of the units sent out for listing had unit status on the MAF of "valid living quarters". Less than one percent was "other uninhabitable". If all of the categories are grouped into valid/invalid unit status, we see on Table 3 below, that those units that were previously invalid had 63 percent consistent listing results from the FRs and those units that were previously valid had 84 percent consistency.

About one-third of the units sent out for listing were not in Census 2000. Units that were not in Census 2000 include a combination of units built since 2000 and units that were missed during Census 2000. Both of these categories of HUs can make it harder to identify which HUs are valid. Table 3, below, shows HUs that were not in Census 2000 had a consistency rate of 78 percent, while units that were in Census 2000 had a consistency rate of 86 percent.

Although mobile homes only made up about 3 percent of the units sent out for listing, they have traditionally been a source of problematic coverage, which can obviously lead to inconsistencies when listing. It may be that FRs are uncertain how to ascertain whether or not a mobile home has valid unit status. Table 3, below, shows that mobile homes had a consistency rate of 71 percent and conventional HUs had a consistency rate of 85 percent.

Slightly more than half of the units sent out for listing were units in multi-unit housing structures. This includes units that are in apartment buildings as well as private homes that were expanded or subdivided into

[7] When 2x2 frequency tables were produced for each of these variables, they each had high chi-squares with associated p-values < 0.0001.

[8] As mentioned earlier, there are various filters used for the MAF. We used the ACS filter to determine HU validity prior to listing. This filter does reflect how current surveys would select units for sample. A unit that passes the filter rules is considered valid for sample.

more HUs. As seen on Table 3, below, units in a multiunit structure had a consistency rate of 70 percent and single units had a consistency rate of 90 percent.

Table 3: HU Consistency Rates by HU Level Variables

	Consistency Rate (s.e.)		Consistency Rate (s.e.)
Previously Invalid for ACS Sample	72% (6%)	**Previously Valid for ACS Sample**	85% (5%)
Previously Invalid Unit Status	63% (10%)	**Previously Valid Unit Status**	84% (4%)
Not In Census 2000	78% (5%)	**In Census 2000**	86% (5%)
Mobile Home	71% (21%)	**Conventional Housing Unit**	85% (4%)
Unit in a multi-unit Structure	70% (13%)	**Single Unit**	90% (2%)

Table 4, below, shows the HU consistency rates by structure size[9], based upon the number of units at the basic street address (BSA). This breakdown is an indication not just of whether or not the unit was a single unit, but also the size of the multi-unit structure associated with that unit's BSA. The category of two to ten units includes smaller multi-unit structures as well as a number of unusual arrangements such as a single unit with a separate apartment added on in the back, a single unit that was split into two units, etc. As expected, it is this middle category that is most confusing to list, with only 62 percent consistency among the FRs in the listing, as compared to 77 percent for larger multi-units and 90 percent for single units.

Table 4: Housing Unit Consistency Rates by Structure Size

	Consistency Rate (s.e.)
1 unit	90% (2%)
2 – 10 units	62% (24%)
11+ units	77% (9%)

4.1.2 Housing Unit Consistency Rates by Block Variables of Interest

How challenging it is to list a HU may also depend upon the characteristics of the block in which it is located. Looking at certain variables of interest about the blocks, based upon our data[10], there is a difference in HU level consistency within each of the tables, according to a chi-square test at the 0.15 significance level.

One important characteristic about the block is whether or not there are any multi-units in the block. This has an effect upon the consistency results of the listing. Table 5, below, shows that units in blocks with some multi-units had a 79 percent consistency rate and those in blocks with no multi-units had a 91 percent consistency rate. It should be pointed out that the consistency rate was 68 percent (with 20 percent standard error) for those HUs in the nine blocks that were all multiunits.

[9] Our data shows that there is a difference in HU level consistency by structure size, based upon chi-square tests at the 0.15 significance level.

[10] When both weighted and unweighted 2x2 frequency tables were produced for each of these variables, they each had high chi-squares, likelihood ration chi-squares and mantel-haenszel chi-squares with associated p-valued < 0.15.

Table 5: Housing Unit Consistency Rates by Presence of Multiunits in Block

	Consistency Rate (s.e.)
No multiunits in block	91% (3%)
Some or all multiunits in block	79% (7%)

A block is considered to be permit-issuing area if the block is covered by a governmental entity that requires building permits to be taken out for all new residential construction. About 92 percent of the U.S. population at the time of Census 2000 lived in blocks classified as permit-issuing. Table 6, below, shows that an HU in a permit-issuing block has consistent listing results 84 percent of the time.

Table 6: Housing Unit Consistency Rates by Permit Status

	Consistency Rate (s.e.)
In Permit-issuing block	84% (5%)
Not In Permit-issuing block	86% (2%)

A block is considered to be either in an urban area or in a rural area, a classification based primarily upon population density. Table 7, below, shows that an HU in a rural block has consistent listing results 91 percent of the time.

Table 7: Housing Unit Consistency Rates by Urban vs. Rural

	Consistency Rate (s.e.)
Urban	81% (6%)
Rural	91% (3%)

Currently the universe from which the demographic household surveys select sample consists of four frames: unit frame, area frame, group quarters (GQ) frame and permit frame. The GQ frame consists of group quarters and the permit frame consists of new construction units in permit issuing areas. The listed blocks are all in either the area or unit frame. Whether a block is in the area frame or the unit frame depends primarily upon permit coverage and proportion of city-style addresses with the area frame blocks containing units in non-permit areas and/or areas with non-city style addresses (meaning addresses that do not include both a house number and a street name). Table 8, below, shows that HUs in area frame blocks had consistency rates of 83 percent and those HUs in unit frame blocks had consistency rates of 84 percent.

Table 8: Housing Unit Consistency Rates by Frame

	Consistency Rate (s.e.)
Area Frame	83% (2%)
Unit Frame	84% (5%)

It is also interesting to look at block size, as characterized by the number of HUs per block. Of the blocks sent out for listing, about half were classified as medium size blocks, one-third was classified as large, and the remainder was classified as small blocks. Table 9, below, shows that HUs in small blocks had a consistency rate of 52 percent, HUs in medium blocks had a consistency rate of 82 percent and HUs in large blocks had a consistency rate of 89 percent. This is counter-intuitive because people imagine that a large block is harder to list correctly. However the HUs in small blocks could be harder to find and to list

properly. Additionally, it is important to note that less than 1 percent of HUs in our sample were in small blocks.

Table 9: Housing Unit Consistency Rates by Number of Housing Units in Block

	Consistency Rate (s.e.)
Small (0-6 housing units in block)	52% (13%)
Medium (7-159 housing units in block)	82% (6%)
Large (160+ housing units in block)	89% (2%)

Block density is defined as number of HUs per square mile and is another measure of block size. The definitions used to classify block density as low, medium, or high is a Census Bureau convention. The table below shows the lowest consistency rates are for units in blocks in the highest range of block density, with a 91 percent consistency rate for those HUs in blocks with the lowest density. This is consistent with the findings for urban and rural blocks in Table 8, above. In general, the blocks with the highest density tend to be blocks with many multi-unit structures and are usually in large urban areas and the blocks with the lowest density tend to be in more rural areas.

Table 10: Housing Unit Consistency Rates by Density of Housing Units in Block

	Consistency Rate (s.e.)
Low (<150 housing units/mile^2)	91% (7%)
Medium (150-320 housing units/mile^2)	87% (5%)
High (>320 housing units/mile^2)	82% (5%)

Most often, blocks that contain mobile homes are blocks with trailer parks, so the number of mobile homes in a block is just as meaningful as the presence or absence of mobile homes in a block when calculating HU consistency rates. The table below shows that HUs in blocks without any mobile homes had an 84 percent consistency rate and those in blocks with mobile homes had an 83 percent consistency rate.

Table 11: Housing Unit Consistency Rates by Presence of Mobile Homes in Block

	Consistency Rate (s.e.)
Blocks with no mobile homes	84% (5%)
Blocks with mobile homes	83% (8%)

4.2 Housing Unit Level Results by Block

It is interesting to look at the consistency rate by block. Overall, the weighted average consistency rate by block was 78 percent (standard error of 7 percent). This means that when listing the HUs in the blocks in our sample, the two FRs agreed about 78 percent of the time in a block as to whether or not a HU in that block is valid, based upon the unit status given to that HU and their actions overall upon the HU. As can be seen in the graph below of consistency rates by block, the consistency rates of the individual blocks ranged widely within the 201 blocks listed and ordered by block consistency rate, with 37 blocks having a 100 percent consistency rate and 6 blocks having a zero percent consistency rate. The six blocks with a zero percent consistency rate all had one valid housing unit sent out for listing.

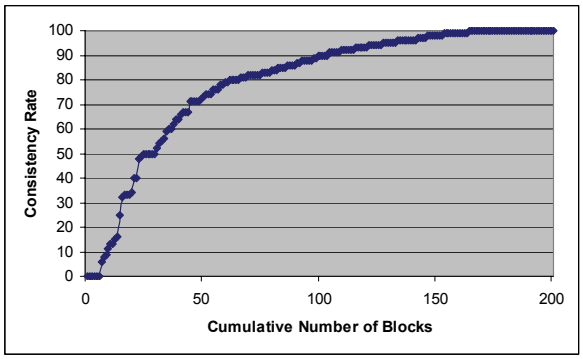

Figure 2: Housing Unit Consistency Rates by Block (in percent)

4.2.1 Housing Unit Consistency Rates by Block by Variables of Interest

How challenging it is to list a block may depend upon the characteristics of the block. In order to create frequency tables, the consistency rates for the blocks were grouped into four categories[11]. Looking at certain variables of interest about the blocks, our data[12] shows that there is a difference in block level consistency within each of the three tables, based upon a chi-square test at the 0.15 significance level.

As stated above, a block is considered to be in a permit-issuing area if the block is covered by a governmental entity that requires building permits to be taken out for all new residential construction. Around 69 percent of the blocks were in permit issuing areas. Table 12, below, shows that almost 100 percent of the blocks that had consistency rates under 75 percent were in permit issuing areas.

Table 12: Block Consistency by Permit Issuing Status

		Permit Issuing Status		
		Non-Permit Issuing	**Permit Issuing**	**Total**
Consistency Rate (s.e.)	**100%**	47,326 (39,196)	406,352(207,478)	453,678 (211,128)
	91% – 99%	68,204 (45,214)	143,305 (71,110)	211,508 (84,267)
	75% – 90%	19,003 (6,188)	222,245 (116,245)	241,248 (116,409)
	0% - 74%	7,060 (2,962)	362,145 (118,505)	369,206 (118,542)
	Total	141,593 (38,503)	1,134,047 (137,950)	1,275,641 (143,129)

The presence of multiunits in the block does potentially affect the block consistency results of the listing. Table 13 shows that almost 45 percent of the blocks that did not have any multi-units were 100 percent consistent and about 16 percent of the blocks had some multi-units were 100 percent consistent.

[11] The four categories are based upon the weighted quantiles and are: 0 to 74 percent consistent, 75 to 90 percent consistent, 91 to 99 percent consistent and 100 percent consistent

[12] When both weighted and unweighted 2x2 frequency tables were produced for each of these variables, they each had high chi-squares, likelihood ration chi-squares and mantel-haenszel chi-squares with associated p-valued < 0.15.

Table 13: Block Consistency Rates by Presence of Multiunits

		Presence of Multiunits in Block		
		No Multiunits	**Some/All Multiunits**	**Total**
Consistency Rate (s.e.)	**100%**	390,335 (213,763)	63,344 (44,244)	453,678 (211,128)
	91% – 99%	107,924 (48,066)	103,585 (69,358)	211,508 (84,267)
	75% – 90%	175,808 (113,064)	65,440 (28,569)	241,248 (116,409)
	0% - 74%	208,320 (88,012)	160,886 (71,869)	369,206 (118,542)
	Total	882,386 (189,593)	393,254 (84,551)	1,275,641 (143,129)

There are multiple ways to calculate the size of a block. Some common ways include block size in square miles, number of HUs in block, and block density in HUs/mile^2. Table 14 looks at the block consistency rates by number of HUs on the MAF prior to listing. For this table, small is zero to six HU/block (18 percent of unweighted blocks), average is 7 to 159 HU/block (47 percent), and large is 160 or greater HU/block (34 percent). For the small blocks, less than one-quarter were 100 percent consistent and almost three-quarters had consistency rates below 75 percent. For the large blocks, more than half had consistency rates greater than 90 percent. It should be noted that given some of the small cell sizes, we are not able to detect differences in the estimates.

Table 14: Block Consistency Rates by Block Size

		Number of HUs Sent Out for Listing			
		Small	**Average**	**Large**	**Total**
Consistency Rate (s.e.)	**100%**	42,462 (25,660)	407,233 (209,579)	3,983 (2,351)	453,678 (211,128)
	91% – 99%	0 (0)	183,556 (84,042)	27,954 (7,679)	211,508 (84,267)
	75% – 90%	54 (54)	217,993 (115,826)	23,201 (13,938)	241,248 (116,409)
	0% - 74%	154,184 (78,600)	211,353 (80,841)	3,669 (1,671)	369,206 (118,542)
	Total	196,701 (81,851)	1,020,133 (153,500)	58,807 (15,138)	1,275,641 (143,129)

4.2.2 Regression Analysis of Block Level Consistency Rates

When running regression analysis of the block level consistency rates, no model was found that fit the data well. However, a few important variables did stand out as predicting lower consistency rates, including: the number of valid ACS HUs on MAF that were not in Census and not on the latest DSF, the number of valid ACS HUs that were single-unit addresses on MAF, being in Western region (of four Census regions), being a rural block, and having a higher population at the time of Census 2000.

4.4 Other Measures

Although the focus of this paper has been on the consistency rate measurement, we would like to briefly mention four other methods of examining inter-rater agreement: error rate, Cohen's kappa coefficient, gross difference rate, and comparison of validity counts.

4.2.1 Error Rate

Our focus has been on consistency rates, but it is important to realize that inconsistency is different from error. If both FRs were incorrect, then the results would be consistent; consistently wrong, but consistent with each other. Therefore, it may also be of interest to calculate the error rates. Under simple independency assumptions, percent inconsistent = 2*e*(1-e), where e is the error rate.

As can be seen in Table 15 below, the consistency rate for all records is about 84 percent, which means about 16 percent inconsistent, or an error rate of about 9 percent. If we are just looking at records that were previously valid under ACS filter rules, the consistency rate is about 85 percent, which means about 15 percent inconsistent, or an error rate about 8.2 percent. If we are just looking at records that were previously invalid under ACS filter rules, the consistency rate is about 72 percent, which means about 28 percent inconsistent, or an error rate of about 17 percent.

Table 15: Error Rates

	Consistency Rate	Inconsistency Rate	Error Rate
All Records	84%	16%	9.0%
Previously eligible for ACS	85%	15%	8.2%
Previously ineligible for ACS	72%	28%	17.0%

4.2.2 Kappa Coefficient

Cohen's kappa coefficient is a measure used for evaluating agreement between two raters who have classified items into mutually exclusive categories, taking into account the probability agreement by chance. Although a larger kappa describes stronger agreement between the raters, the exact interpretation of kappa is somewhat subjective and there is currently no uniform method being used. There are some experts who are skeptical about its use, but it can be used to give an indication of inter-rater agreement (Gwet, 2001). As the table below shows, the kappa coefficient for the overall HU level data is 0.47. According to Landis and Koch (1977), a kappa of 0.47 can be interpreted by saying that, after accounting for chance, there is moderate agreement between the FRs. However, as we look at other variables for the HU data, we find that based upon the Landis and Koch interpretation of kappa, there was substantial agreement between the FRs for single units and for rural units, only fair agreement between the FRs for previously invalid units and for trailers, and just slight agreement units in blocks made up completely of trailers.

Table 16: Kappa Coefficient for Housing Unit Level Data

	Cohen's Kappa	Interpretation
Overall	0.47	Moderate agreement
Housing Units in Rural Blocks	0.69	Substantial agreement
Single Housing Units	0.63	Substantial agreement
Previously Invalid Units	0.25	Fair agreement
Trailers	0.23	Fair agreement
Housing Units in 100% Multiunit Block	0.05	Slight agreement

4.2.3 Gross Difference Rate

The gross difference rate (Hansen, Hurwitz, and Pritzker, 1964) is an estimate of simple response variance between the two listers, as can be seen in the formula below, where g is the gross difference rate and, in this

case, depending upon whether the listers agree, the squared term has a value of one (disagree) or zero (agree).

$$g={}^{1}/_{n}\,\Sigma(Y_{i1t} - Y_{i2t})^2$$

In this case, the weighted gross difference rate is 0.16, which means that if the measurements are independent and identically distributed, 0.08 is an unbiased estimate of the simple response variance. Although this estimate is interesting, it cannot be considered reliable because, as stated in the limitations section, we cannot assume that the measurements of HU consistency were independent.

4.2.4 Raw Validity Counts

A benefit of looking at the raw count of valid units returned from listing is that this count includes all HUs that the FRs considered valid, including adds. This also points to the idea that it is possible that the FRs disagreed on an individual HU basis, but they could have balanced each other out and still come out with similar counts of valid HUs for the block.

Looking at block level counts of valid HUs (including adds) shows how consistent the two FRs were in the total count of valid units for each of the 201 blocks returned from the listing. Figure 3 shows the counts of valid units returned by block, for each of the two FRs. Although there are certainly blocks with outliers, there are not many blocks where the valid counts from the two FRs differ greatly, as can be seen by the few blocks plotted further from the black trend line. Because these counts include added units, it is possible that these outlier blocks were due to the FR adding units that are beyond the block boundaries.

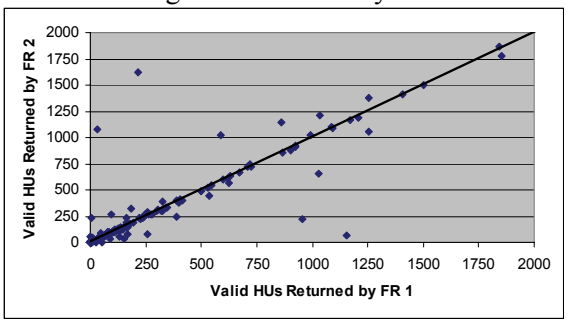

Figure 3: Block Level Counts of Valid Units Returned from Listing

5. Conclusion

In conclusion, when listing growth blocks, the two FRs agreed about 84 percent of the time as to whether or not an HU was valid. There are many variables involved, both at the HU and block level. However, by looking at various breakdowns of the data, it is possible to see that differences in consistency do exist, especially for HUs that were trailers, previously invalid, in multi-units or in blocks made up completely of multi-units. For blocks, the consistency rate varied by the presence of multiunts in the block and block size (as represented by number of HUs in the block). Although the gross difference rate and the raw validity counts show that the variance might not be so great, both the kappa coefficient and the error rate indicate that inter-rater agreement was not ideal. This points to two conclusions. First, that although no direct relationship was found to explain all the variation in the responses of the listing, the relationships between different variables and consistency do imply that there may be certain characteristics of HUs and blocks that contribute to listing difficulties. Second, this is definitely an area that has room for improvement. Area listing is not perfect and could benefit from further research as we try to ascertain what really does contribute to listing difficulties.

References

Gwet, K. (2001), *Statistical Tables for Inter-Rater Agreement*. (Gaithersburg : StatAxis Publishing).

Hansen, M., Hurwitz, W.,and Pritzker, L. (1964), The estimation and interpretation of gross differences and the simple response variance, In C.R. Rao with D.B. Lahiri, K.R. Nair, P.Pant, and S.S. Shrikhande, eds., Contributions to Statistics, Presented to Professor P.C. Mahalonibis on the Occasion of his 70th Birthday. Oxford, England, pp. 11-136.

Landis, J.R. and Koch, G. G. (1977), "The measurement of observer agreement for categorical data", *Biometrics*. Vol. 33, pp. 159-174.

Loudermilk, C., and Li, M., (2009), *A National Evaluation of Coverage for a Sampling Frame Based on the Master Address File (MAF)*, Proceedings of the American Statistical Association.