Section 7 – Evaluation of Bias Report

INTRODUCTION

This report describes the results of an evaluation of the geocoding and mapping techniques for the Mapping Evaluation Follow-on Project: Legal Services in Georgia Updates and Refinements. The goals of the project are to assess the suitability and value of mapping techniques for legal services in Georgia. A series of maps and charts depicting variables relevant to legal aid services were provided for this project. For this evaluation of bias, a preview sampling of maps and charts were produced using techniques and methods developed from the previous mapping project in Georgia. In addition, a geocoding results dataset for statistical evaluation was provided. This evaluation of bias is to ensure that the cartographic techniques and geocoding methodology used is statistically robust, and accurately represents the data provided.

CONCLUSIONS

There is no significant bias in the three client database fields tested. The departures from the mean are relatively small. The highest departures from the mean were found in the "none" category (no ethnicity indicated) of the "ethnicity" field. This reflects the case records where in addition to no ethnicity provided, no address or zip code information is provided; therefore neither geocoding method is possible. These are cases where small amounts of data were collected to document these cases in general.

The excellent geocoding success rate eliminates the concern regarding non-geocoded participants. The bias towards geocode fails (NG) for participants with rural routes is negated by the small frequencies in that category. The spatial bias of 'enhanced' geocoding in rural areas has no bearing on the mapping results. The spatial bias only indicates the need for and success of the 'enhanced' technique to counteract the poor quality of address data in rural areas.

DISCUSSION OF FINDINGS

Geocoding Location Bias

Two methods were applied to geocode the grantee database: MapMarker (geocoding software) and an "enhanced" geocoding method, developed to assign case records not successfully geocoded by the MapMarker method to a Block Group based on the ZIP Code area. Before evaluation of any potential geocoding bias within the grantee database, each resulting geocoded case record was classified into five groups based on the geocoding results:

- S – Single address match (MapMarker)
- M – Multiple address candidate match (MapMarker)
- Z – successful match at the ZIP code centroid level (MapMarker)
- EN - ENhanced geocoding method
- NG – Not Geocoded by either method

To evaluate location bias, geocoding success was compared across the following address types: regular residential, rural route, and PO Box. Previous experience geocoding Atlanta Legal Aid

Society cases in Georgia demonstrated that rural locations will tend to be under-represented due to insufficient address information.

The application of the enhanced geocoding methodology has minimized this concern, but it is still a factor to consider. Tables 7.1 and 7.2 illustrate the percentages of geocoding match success by address type. Approximately 90% of the cases not successfully geocoded by either method are rural route addresses with no zip code provided. 15.2% of all rural route address types were not successfully matched by either geocoding method. But overall, a geocoding success rate of 96.27% out of 64,731 total cases indicates low potential for significant bias.

Table 7.1 – Percentage of Geocoding Success by Address Type

Address Type	S	M	Z	EN	NG
Regular residential	88.98	6.06	4.85		0.10
Rural Route	62.84	1.72	20.23		15.21
PO Box	90.53		5.30	3.79	0.38

Table 7.2 – Geocoding Success by Address Type
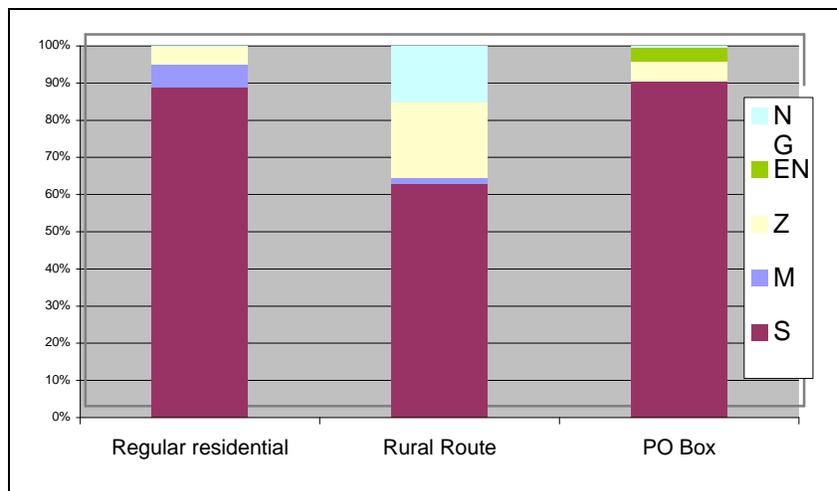
In addition to examination of bias by address type, percentages and departures from the mean of geocoding match levels were compared within three client database fields: problem code, closure code and ethnicity (Tables 7.4 – 7.12). Because there are a large number of problem and closure code categories, the top six categories (Table 7.3) were analyzed and a seventh category, ("other") incorporates all other problem and closure code records.

Table 7.3 – Problem and Closure Code Descriptions

Problem Code	Description
2	Collection
31	Custody/Visitation
32	Divorce/Separation/Annulment
37	Spouse Abuse
63	Landlord/Tenant (not Public housing)
95	Wills/Estates

Closure Code	Description
A	Counsel and Advise
B	Brief Services
G	Negotiated Settlement with Litigation
H	Administrative Agency Decision
I	Court Decision
E	Client Withdrew or Did Not Return

There is no significant bias in the three client database fields tested. Tables 7.4 through 7.12 illustrate these findings.

PROBLEM CODE

Table 7.4 – Problem Codes and the Percentage of Each Geocoding Assignment Category

Problem Code	S	M	Z	EN	NG
2	85.66	4.27	4.08	5.87	0.12
31	83.54	6.01	4.78	5.57	0.11
32	82.22	5.35	4.70	7.38	0.35
37	81.41	4.74	4.18	9.40	0.28
63	84.57	7.80	4.29	3.18	0.16
95	87.46	3.52	5.03	3.57	0.41
Other	83.70	5.83	4.90	5.30	0.26
Mean	84.08	5.36	4.57	5.75	0.24

For each problem code, the geocoding assignment category [S, M, Z, EN, NG] percentages sum across each row to 100%, and reflect al cases being analyzed for each problem code. The mean in the table above reflects the mean for all addresses geocoded in this project, and not the mean for just the project codes displayed.

Table 7.5 – Departure from Mean of Table 7.4 Data

Problem Code	S	M	Z	EN	NG
2	1.58	-1.09	-0.49	0.11	-0.12
31	-0.54	0.65	0.21	-0.18	-0.13
32	-1.86	-0.01	0.13	1.63	0.11
37	-2.67	-0.62	-0.39	3.64	0.04
63	0.49	2.44	-0.28	-2.57	-0.09
95	3.38	-1.84	0.47	-2.18	0.17
Other	-0.38	0.47	0.34	-0.45	0.02

To calculate the departure from the mean, the means in table 7.4 were subtracted from each of the assignment values in Table 7.4.

Table 7.6 – Geocoding Results by Problem Code
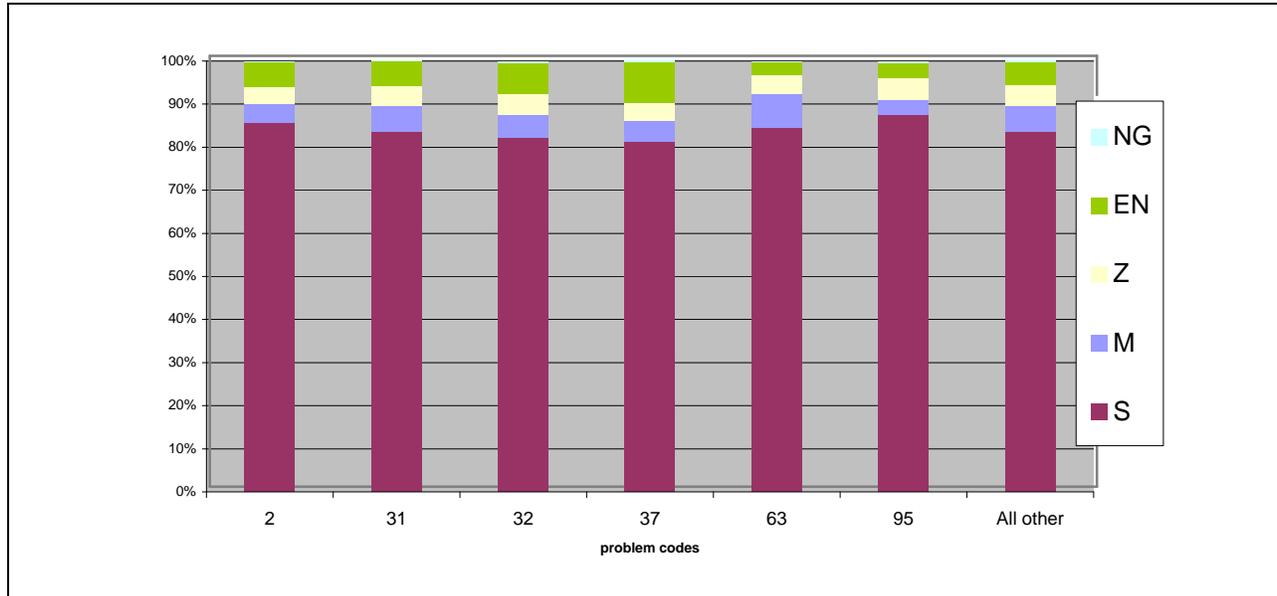
Note the similarity or consistency in the results.

CLOSURE CODE

Table 7.7 – Closure Codes and the Percentage of Each Geocoding Assignment Category

Closure Code	S	M	Z	EN	NG
A	83.73	5.78	4.77	5.48	0.24
B	86.06	5.32	4.88	3.51	0.23
G	80.74	5.22	4.79	8.82	0.43
H	79.18	5.36	4.40	10.57	0.48
I	79.93	5.51	5.15	9.11	0.30
K	82.33	5.39	3.38	8.38	0.52
Other	83.92	5.35	4.65	5.77	0.31
Mean	82.27	5.42	4.58	7.38	0.36

Table 7.8 – Departure from Mean of Table 7.7 Data

Closure Code	S	M	Z	EN	NG
A	1.46	0.36	0.20	-1.90	-0.12
B	3.79	-0.10	0.30	-3.86	-0.13
G	-1.53	-0.20	0.22	1.44	0.07
H	-3.09	-0.06	-0.17	3.19	0.12
I	-2.34	0.09	0.57	1.73	-0.06
K	0.06	-0.03	-1.20	1.00	0.16
Other	1.64	-0.07	0.08	-1.61	-0.05

Table 7.9 – Geocoding Results by Closure Code
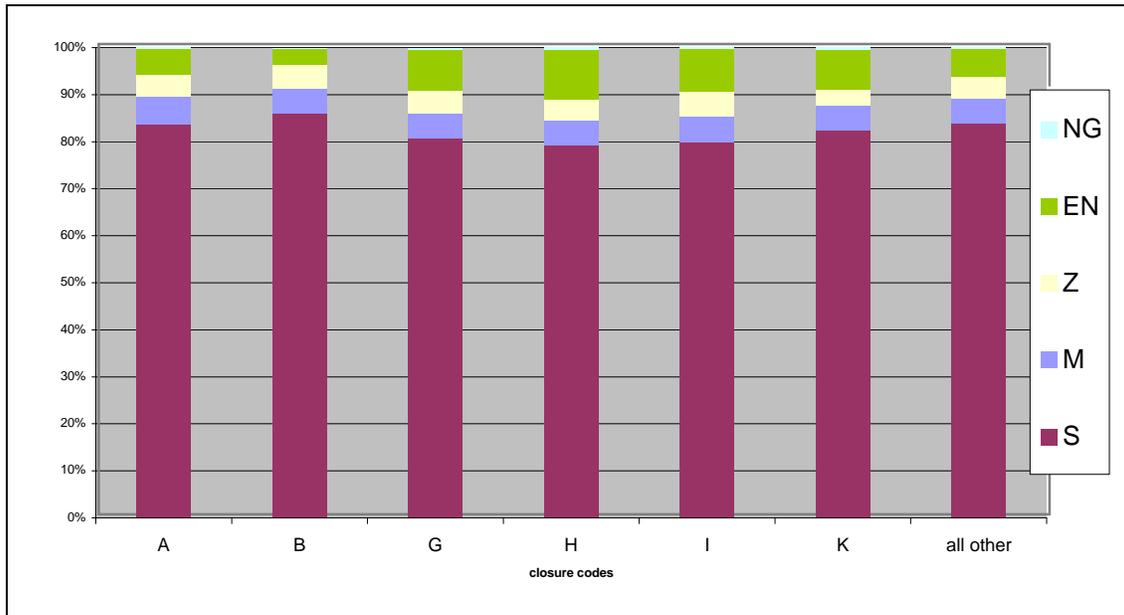
ETHNICITY

Table 7.10 – Ethnicity and the Percentage of Each Geocoding Assignment Category

Ethnicity	S	M	Z	EN	NG
Asian	84.03	7.56	5.04	3.36	0.00
Black	83.62	6.43	4.56	5.20	0.18
Hispanic	78.94	8.78	7.43	4.61	0.25
Native American	83.05	3.39	6.78	6.78	0.00
Other	89.01	2.42	4.67	3.48	0.41
White	83.54	4.43	4.73	6.92	0.38
None	94.55	0.00	5.45	0.00	0.00
Mean	85.25	4.72	5.52	4.34	0.18

Table 7.11 - Departure from Mean of Table 7.10 Data

| Ethnicity	S	M	Z	EN	NG
Asian	-1.22	2.85	-0.48	-0.97	-0.18
Black	-1.63	1.72	-0.96	0.87	0.01
Hispanic	-6.31	4.07	1.90	0.27	0.07
Native American	-2.20	-1.33	1.26	2.44	-0.18
Other	3.77	-2.29	-0.86	-0.86	0.24
White	-1.71	-0.29	-0.79	2.59	0.21
None	9.30	-4.72	-0.07	-4.34	-0.18

Table 7.12 – Geocoding Results by Ethnicity
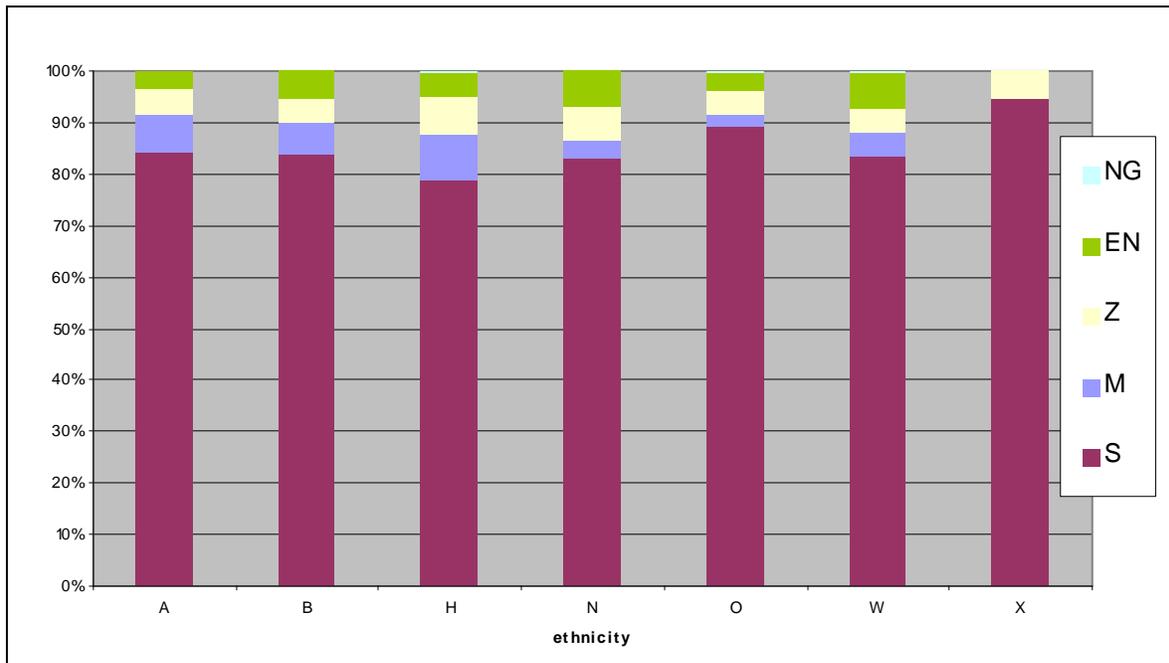
Figure 7.1 – Map of the distribution of total cases and percentage of total cases by Zip code geo-coded using Enhanced Geocoding Method in Georgia.