Surveying Community Stakeholders: Exploring the Viability of a National Sampling Frame

Barbara J. Robles, PhD*
Board of Governors of the Federal Reserve System, Barbara.j.robles@frb.gov

Abstract

This paper explores the viability of creating a national local development organization (LDO) and community stakeholder sampling frame. The motivation for assessing the feasibility of a national sampling frame stems from a data collection program currently piloted by the Community Affairs function of the Federal Reserve System. Using public domain administrative data from the Internal Revenue Service 990 business master file for tax exempt organizations and the Census of Governments (Census Bureau) for local public sector entities, a stratified sampling frame is considered. Coverage and further stratification questions in the context of LDOs serving low-to-moderate income locales are raised.

Keywords: (sampling frame, local development organizations (LDOs), community stakeholders and administrative data)

**The opinions expressed in this paper are those of the author and do not necessarily reflect the views of the Federal Reserve Board or its staff.*

1. Overview and Background

Community stakeholders are opinion makers, change agents and local leaders knowledgeable about current and emerging community economic conditions. Community economic development stakeholder surveys are often administered only at the local level by grant-funded foundation-university-public sector and non-profit organization partners. Some of these community surveys are called 'asset-mapping' or 'community needs assessments' and focus on limited respondent pools using questionnaires designed to capture very narrow community sector issues or services.

Local, place-based economic conditions can be assessed by identifying administrative data sources or registers that contain local, municipal, state and regional level information on local development organizations (LDOs). Once the universe of LDOs and community stakeholders is identified from publicly available administrative data and registers, one can stratify the target population by regional districts, organizational categories, and entity size to determine the potential respondent pool. Community stakeholder and local development organizations would include the following sectors: non-profit organizations, faith-based entities, and public sector agencies that provide education services, health services, financial, tax and legal services, housing services, transportation services, food banks and farmers' markets, youth development, cultural, arts, leisure, community development, and social services. The target population consists of those LDOs supporting local economic development sustainability and inclusive growth. We explore the feasibility of creating a sampling frame that would identify the universe of LDOs.

This paper provides an overview of information gathering initiatives across the Federal Reserve System that capture community conditions, economic inclusion indicators and emerging issues in predominantly low-to-moderate income (LMI) communities from 2009 to the present. During the financial crises, many community stakeholders serving LMI communities relayed anecdotal information from a variety of geographical locales and produced narratives from particular economic sectors about emerging issues. The problem with interpreting anecdotal information is that it often produces more questions than answers. For example, was the issue being described local or was there a geographical dispersion factor to consider? Was the story a one-time event or did the issue contain the seeds of a trend? In a rapidly changing economic environment such questions become crucial to fully understanding the significance of emerging conditions at the local level as well as being able to track these conditions over time and geographies.

The paper is organized as follows: (i) section two presents ongoing exploratory regional survey efforts for capturing current and emerging community economic development issues that parallel business expectations and outlook surveys, (ii) section three explores methodological approaches and identifies the potential data sources that exist in the public domain that may be used to create a national sampling frame for LDOs and community stakeholders. The paper concludes by identifying future research questions addressing emerging issues in identifying the universe of LDOs.

2. Summary of Community Data Initiative – Findings and Insights

The aftermath of the deep financial recession of 2007-2009 presented the community development research field with an array of opportunities directly impacting the task of information gathering at the local level. While most economic data is monitored at the micro and macro levels, the goal to ascertain regional variations in markets and economic conditions brought to light the challenge of capturing local community information. Such community information captures vital data on recovery capacity of particular regional markets and geographical areas. The community affairs function of the Federal Reserve System in the United States is uniquely organized[1] to capitalize on providing early information at the community level that creates a bridge between ongoing micro data surveys and the macro aggregates of particular economic sectors. At the same time, new technologies have emerged in the community development data collection space amplifying the opportunities and the challenges for monitoring emerging community development and economic inclusion issues.

Gathering primary community data was approached as an opportunity to initiate consistent fielding of quantitative and qualitative mixed-mode surveys that capture real-time information. This new source of standardized community stakeholder information complements the ongoing micro and macro data collection efforts of other agencies. Focusing on the data gaps brought to light by the Great Recession allowed us to implement a more systematic approach to collecting information on current local conditions and emerging challenges facing low- and moderate-income (LMI) communities. We identified the following outcomes to a collective, geographic specific information gathering effort:

a. Enables us to differentiate between single event anecdotes and trends over time;

b. Provides us with a measure of geographical dispersion for triangulating district data findings with national data monitoring results; and

c. Capitalizes on Reserve Banks' deep regional and district knowledge and community relationships.

Using web-based surveys as a primary data collection method, we minimized the time lag present in larger, more costly national surveys. In addition to using a real-time data gathering internet platform, we provided sampling rigor by employing both quantitative and qualitative mixed-methods in the survey design. Each reserve bank collected district specific information from local community stakeholders. Community stakeholder respondents were identified as local development organizations such as community based organizations (CBOs), non-governmental organizations (NGOs), community financial institutions, small businesses, faith based organizations, educational institutions, foundations and public sector agencies. To complement and independently triangulate district level findings, a national survey is administered as a deliberative sample using a mixed-method instrument. In addition to

[1] The Federal Reserve System consists of 12 geographically defined regional Reserve Banks and the Board of Governors of the Federal Reserve System located in Washington, DC.

triangulating the community level indicators resulting from Reserve Bank collection efforts, secondary sources of community level data from other government agency surveys[2] are used as comparative references.

Table 1. Federal Reserve System Community Conditions Data Collection Overview

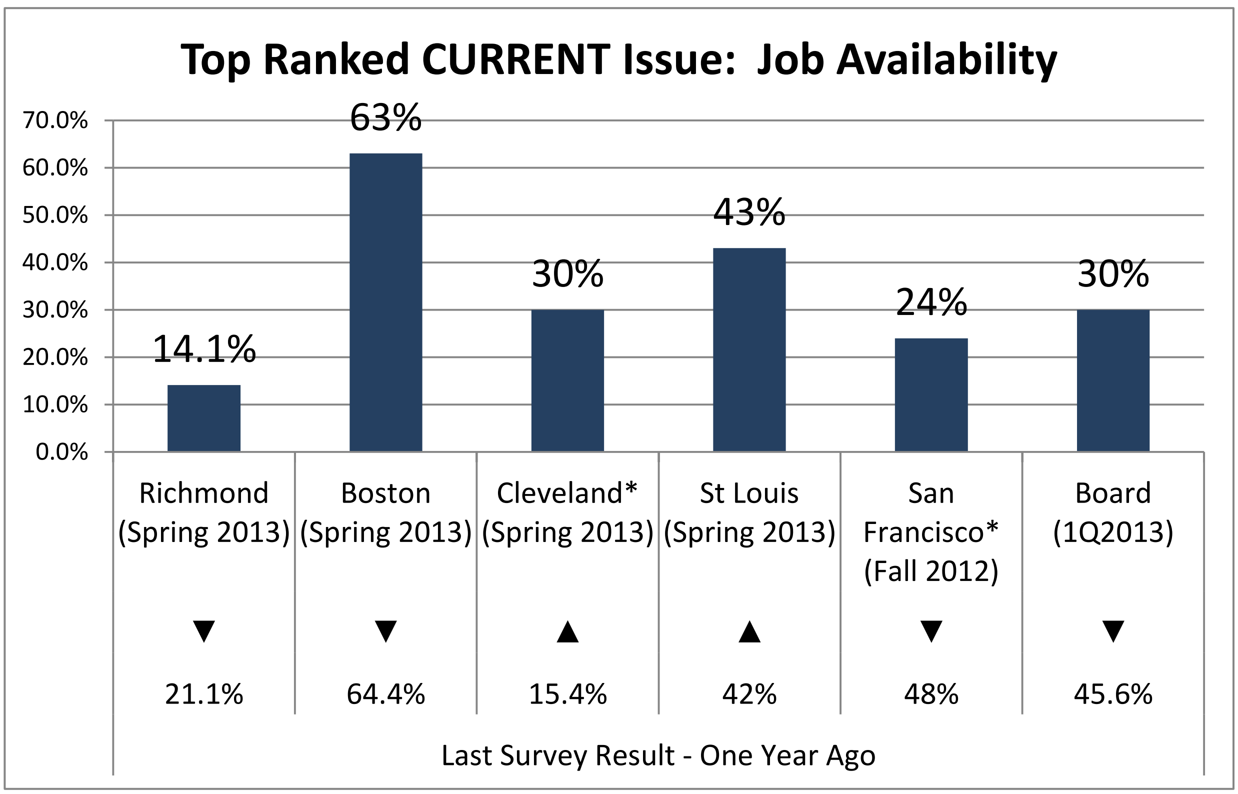
Respondent Sample	Publication	Instrument Type	Collection Frequency
Board Community Stakeholder Sample			
San Francisco FRB	Vantage Point	Ranking	Biannual
Richmond FRB	Community Pulse	Ranking	Biannual
St. Louis FRB	Community Development Outlook Survey	Ranking	Annual
Boston FRB	Community Outlook Survey	Combination	Annual
Cleveland FRB	Issues Insight Survey	Combination	Biannual
Minneapolis FRB	Ninth District Insight Report	Combination	Biannual
Non-Profit/NGO Stakeholder Sample			
Kansas City FRB	Low-to-Moderate Income (LMI) Index	Diffusion Index	Quarterly
Dallas FRB	Community Outlook Survey (COS)	Diffusion Index	Annual
Philadelphia FRB	Community Outlook Survey	Diffusion Index	Quarterly
Board of Governors	Community Indicators	Combination	Quarterly
Small Business Stakeholder Sample			
New York FRB	Small Business Survey		Annual
Atlanta FRB	Small Business Survey		Biannual

[2] The secondary sources employed for validation and tracking of community survey findings most commonly used are: the American Community Survey, American Housing Survey, Bureau of Labor Statistics, and the Longitudinal Employer-Household Dynamics.

Each Reserve Bank designed a survey instrument and identified particular community stakeholder respondents that best correspond with district specific issues and resources. Three types of polling instruments have emerged from the 12 Reserve Bank district community development research units: (i) diffusion indices,[3] (ii) top three ranked economic issues (current and emerging; See Figure 1) and (iii) small business establishment surveys (credit access and local economic conditions). All survey instruments also include an open-ended, unstructured question that allows for the capture of stakeholder concerns not addressed in the questionnaire. The survey instruments are consistently administered either using a quarterly, bi-annual or annual time frame.

The structured questions from all the surveys provide a quantitative assessment of local community economic development conditions across the Federal Reserve System. The unstructured information content generated by the open-ended text responses from the LDO and community stakeholder respondents allows for the application of thematic cluster and sentiment analysis revealing community concerns unique to the geographical district. At the same time, the results of the text analytics corroborate newly emerging community issues captured in the structured questions. By employing a mixed methods approach (using both quantitative and qualitative methodologies), we are able to monitor rapidly changing conditions on the ground while capturing newly emerging issues useful to regional and macroeconomic analysis. In addition to providing better information to policymakers and researchers, the capture of local community conditions provides useful information to low-to-moderate income communities and promotes economic inclusion.

Figure 1. Community Conditions across Selected Reserve Bank Districts – Percent of Respondents Ranking Job Availability as the Top Current Issue



[3] A diffusion index measures the degree of dispersion of change in the index. The amount of change is based on components of the index that have increased, decreased or remains unchanged within a specific time period. See the Monthly Labor Review, "Diffusion Indexes: a barometer of the economy," Vol.113, No.4 (April 1990), pp. 13-21 and Edmiston (2013) for a complete discussion of diffusion indexes.

3. Exploring Public Domain Data Sources and Identifying a National Local Development Organization (LDO) Sampling Frame

Exempt organizations and public sector agencies comprise the bulk of local development organizations and community stakeholder administrative data that is in the public domain. In 1995, the Internal Revenue Service (IRS) instituted a National Taxonomy of Exempt Entities (NTEE) creating 26 major sectors with corresponding sub-classifications (see Table 3). The administrative data (IRS Exempt Organizations Business Master File) contains address and zip code level information, NTEE codes, legal category, date of filing, income and asset size classification, and other variables. The data is processed on an ongoing basis and posted on the IRS website. The internet allows us to create the universe of all reporting exempt organizations with the IRS nationally.[4] The exempt organizations are often community based, local or with a regional reach. For our purposes, splitting the exempt organization database into the 12 reserve bank districts is the first step in identifying the exempt entity geographical coverage.

Table 2. Internal Revenue Service, 990 Exempt Organizations Business Master File by Reserve Bank Districts, (July 2013)

FRB Districts	IRS 990 Organizations	Census Pop 2012	Population to Exempt Organization Ratio
San Francisco	281,406	64,724,758	230
Chicago	193,467	34,472,443	178
Atlanta	180,522	44,393,689	246
Richmond	161,575	30,877,596	191
Minneapolis	109,544	9,094,002	83
Dallas	105,866	27,730,230	262
Cleveland	98,650	16,962,116	172
Boston	88,703	13,628,869	154
Kansas City	83,224	17,645,485	212
New York	72,788	26,570,167	365
Philadelphia	66,090	13,153,889	199
St Louis	63,718	14,660,796	230

Source: Internal Revenue Service, 990 Business Master File (http://www.irs.gov/uac/SOI-Tax-Stats-Exempt-Organizations-Business-Master-File-Extract-(EO-BMF)

The largest reserve bank district by geography and population is San Francisco and the smallest reserve bank district by population is Minneapolis although not the smallest by exempt organization coverage (see Table 2). The number of exempt organizations varies by geography, historical population migrations, economic booms and declines as well as changing migration patterns. The ratio of population-to-organization coverage is on average 210 but varies by population density and geography from a ratio of 83:1 to 365:1.

[4] Gørnbjerg (1989) and Gørnbjerg et al (2010) indicate that not all local development organizations and community stakeholders are contained in the IRS database due to a variety of issues: (i) under the revenue threshold for filing ($5K), (ii) congregation status, (iii) lack of capacity issues, (iv) affiliated with a registered organization or fiscal agent and (v) other statutory, technical or compliance reasons.

The asset size classification of the exempt organizations is allocated on a four tier frame: zero, small, medium and large, where zero is zero or less, small is between $1 and $99,999, medium falls between $100,000 and $4,999,999, and large is $5 million and over.

The final classification is by aggregated sector. Following the classification scheme of the United Nations (UN) for incorporation of non-profits into the systems of national accounts, I collapse the 26 IRS NTEE categories into the 11 categories used by the UN and include the stand-alone IRS category of Unknown (Z) to end with a total of 12 categories. We have identified12 Reserve Banks, 4 sizes of organizations and 12 LDO sectors by which to stratify and allocate the master list of exempt organizations.

Table 3. Collapsing National Taxonomy of Exempt Entities into United Nations Non-Profit Sector Classifications

IRS -NTEE Code & Description (26)	**UN Classifications (11)**	**Aggregated NTEE (12)**
A - Arts, Culture and Humanities		
B - Educational Institutions and Related Activities	Culture & Recreation	A+N
C - Environmental Quality, Protection and Beautification	Education & Research	B+O+U+V
D - Animal-Related	Health	E+F+G+H
E - Health – General and Rehabilitative	Social Services	K+P+W
F - Mental Health, Crisis Intervention	Environment	C+D
G - Diseases, Disorders, Medical Disciplines	Development & Housing	L+S
H - Medical Research	Law, advocacy & Politics	I+M+R
I - Crime, Legal-Related	Philanthropic Intermediaries & voluntarism promotion	T
J - Employment, Job-Related	International	Q
K - Food, Agriculture and Nutrition	Religion	X
L - Housing, Shelter	Business & Professional Assoc., unions	J+Y
M - Public Safety, Disaster Preparedness and Relief	Unknown	Z
N - Recreation, Sports, Leisure, Athletics		
O - Youth Development		
P - Human Services – Multipurpose and Other		
Q - International, Foreign Affairs and National Security		
R - Civil Rights, Social Action, Advocacy		
S - Community Improvement, Capacity Building		
T - Philanthropy, Voluntarism and Grant-making Foundations		
U - Science and Technology Research Institutes, Services		
V - Social Science Research Institutes, Services		
W - Public, Society Benefit – Multipurpose and Other		
X - Religion-Related, Spiritual Development		
Y - Mutual/Membership Benefit Organizations, Other		
Z - Unknown		

Source: Internal Revenue Service, 990 Business Master File (2013) (http://www.irs.gov/uac/SOI-Tax-Stats-Exempt-Organizations-Business-Master-File-Extract-(EO-BMF) ; the Handbook on Non-Profit Institutions in the System of National Accounts, United Nations (2003) (http://unstats.un.org/unsd/publication/seriesf/seriesf_91e.pdf).

Data from the local public sector entities are collected every five years by the Bureau of the Census. These data are listed under three separate classifications: general purpose, special district and school districts. General purpose governments are composed of local county, municipal, and township government entities. County governments provide general government activities in geographic locales. Municipal governments are sub-county general purpose governments providing general services for a defined population and area. Consolidated city-county governments are treated as municipal governments. Township governments are sub-county general purpose governments providing general services for areas without regard to population concentrations (Bureau of the Census, 2006).

Special district governments are established to provide only one or a limited number of designated services (functions) and have sufficient administrative and fiscal autonomy to qualify as independent governments (Bureau of the Census, 2006).

School district governments provide public elementary, secondary and/or higher education services and have sufficient administrative and fiscal autonomy to qualify as independent governments. They do not include school systems that are "dependent" on a county, municipal, township, or state government (Bureau of the Census, 2006).

Table 4. Public Sector Entities Master File, Census of Governments, Bureau of the Census, 2012

FRBs	General Purpose	School Districts	Special Districts	All Public Sector Entities	Ratio of Public Sector Entity to Population in FRB
Atlanta	2363	518	2748	5629	7887
Boston	1592	656	1358	3606	3779
Chicago	7229	2114	4533	13876	2484
Cleveland	3457	947	1582	5986	2834
Dallas	1714	1159	2486	5359	5175
Kansas City	4918	1605	7517	14040	1257
Minneapolis	6727	1165	2836	10728	848
New York	1986	1011	1382	4379	6068
Philadelphia	2023	539	1572	4134	3182
Richmond	1728	133	1214	3075	10041
San Francisco	1949	1961	6526	10436	6202
St Louis	3189	1037	3397	7623	1923
Totals	38875	12845	37151	88871	

Source: Census of Governments Master Files (July 2012), http://www.census.gov/govs/cog/

The Census of Government master files contain addresses and geographical locations as well as web-site information which allows for an additional (non-overlapping) sampling frame to be employed for the public sector entities. The stratification process for these entities would depend upon population for general purpose governments, enrollments for school districts, and service area size for special districts.

4. Future Challenges and Remaining Questions

The unique mission of the Community Affairs function of the Federal Reserve System provides a multidisciplinary space by which to pilot a variety of survey and data collection techniques: web-based quantitative, qualitative, and mixed-methods surveys. As our data results continue to accrue over time, we are able to ascertain a variety of fine-tuning mechanisms that yield more methodological rigor by incorporating emerging data visualization and spatial mapping techniques. Data mapping and spatial analysis promises to increase our knowledge of geographical dispersion and correlated spatial events with or without seasonal fluctuations. We continue to monitor and apply various triangulation methodologies to provide verification and independent validation of our results. By creating a stratified sampling frame that relies on administrative data in the public domain such as the 990 Internal Revenue Service business master files and the Census Bureau government sector master files, we are able to explore the best stratification categories to employ. Several methodological and research questions remain:

1-How to manage and mitigate the inherent skewed distribution of urban-rural population ratios by local development organizations (LDOs)?

2-Does including a correction for geographical variation and density across FRB districts accounting for area/spatial disparities among the FRBs help ensure sufficient coverage?

3-Would identification of particular zip codes via census tract aggregation allow us to also oversample in distressed economic areas?

We continue to explore various methodological avenues contributing to simplifying the stratification protocol. We are also assessing state and local level registers for local development organizations that might contribute to enhanced coverage.

5. References

Bureau of the Census. "Classification Manual: Covering the activities of the Federal, state, and local governments," (October 2006) http://www2.census.gov/govs/pubs/classification/2006_classification_manual.pdf

Das, Marcel et al, (2011), Social and Behavioral Research and the Internet, Volume in the Series of European Association of Methodology (Routledge, Taylor & Francis Group: New York, NY)

Dattalo, P. (2010), Strategies to Approximate Random Sampling and Assignment, (Oxford University Press, Inc.: New York, NY)

Dillman, D. et al, (2009), Internet, Mail, and Mixed-Mode Surveys: The Tailored Design Method, 3rd Edition, (John Wiley & Sons, Inc.: Hoboken, NJ)

Edmiston, K., (2013), "The Low- and Moderate-Income Population in Recession and Recovery: Results From a New Survey," Kansas City Federal Reserve Bank, Economic Review, First Quarter, (http://www.kc.frb.org/publicat/econrev/pdf/13q1Edmiston.pdf)

Getz, P. M & M. G. Ulmer, "Diffusion Indexes: a barometer of the economy," Monthly Labor Review, Vol.113, No.4 (April 1990), pp. 13-21

Governor Daniel K. Tarullo (2009), "The Community Affairs Function," speech delivered at the Interagency Community Affairs Conference, held in Arlington, Virginia, August 25, (http://www.federalreserve.gov/newsevents/speech/tarullo20090825a.htm)

Gørnbjerg, K. A., "Developing a Universe of Nonprofit Organizations: Methodological Considerations," Nonprofit and Voluntary Sector Quarterly, Vol. 18, No.1, (Spring 1989), pp. 63-80

Gørnbjerg, K. A., H. K. Liu, and T. H. Pollack, "Incorporated bu Not IRS-Registerd: Exploring the (dark) Grey Fringes of the Nonprofit Universe," Nonprofit and Voluntary Sector Quarterly, Vol. 39, No. 5, (2010), pp. 925-945

Snijkers, Ger; Haraldsen, Gustav; Jones, Jacqui; Willimack, Diane (2013-07-01). Designing and Conducting Business Surveys (Wiley Series in Survey Methodology) (Kindle Location 4414). Wiley. Kindle Edition.