Beyond Property Surveys

Modeling Vacancy Using Local Administrative Data

April Urban
Centre on Urban Poverty and Community Development
The Opportunity

Test the utility of machine-learning techniques on relevant community problem

Make additional use of comprehensive point-in-time survey

Community need and use potential

Routinized administrative data

Point-in-time property survey
What is machine learning?

As explained by a layman

TCI survey is a one-time measure of vacancy. How can we use other available data to help guess a property’s vacancy status?

Using TCI survey as “ground truth”

Comparing to other vacancy estimation methods (postal, REO status)
Have you had any OVV complaints in the last 2 years?

- No
  - Estimate not vacant
    - 94% correct

- Yes
  - What percentage of your taxes have you paid?
    - <1%
      - Estimate vacant
        - 79% correct
    - ≥1%
      - Estimate not vacant
        - 59% correct
Community Use

Single source of information for parcel-level vacancy proxy status (instead of postal vacancy OR vacancy survey OR water shut off)

CDCs that monitor vacancy properties can better target effort

Tested instrument to use when point-in-time survey is outdated
Additional Use

Usable as a measure of vacancy for other research work?

Ideas?
Contact

April Hirsh Urban
Center on Urban Poverty and Community Development
april.urban@case.edu
216-368-3390
Cleveland Property Survey

Project Description and Lessons Learned

Frank Ford
Senior Policy Advisor, Western Reserve Land Conservancy
fford@wrlandconservancy.org

October 27, 2016
**Property Vacancy**

How many?

Where is vacancy concentrated?

Establish baseline for tracking over time.

*Note: US Postal Data alternative.*

**Property Condition**

Rating by condition (A, B, C, D, F)

Number and concentration by condition?
158,000 parcels were surveyed over 16 weeks

Residential
Commercial
Industrial

Structures as well as vacant land
Survey Team

16 Surveyors

Plus 1 Team Leader

Surveyors worked in pairs

Surveyors got one week of training

Identifying indicators of vacancy

Applying the condition rating scale

In Classroom and in the Field

Project management overseen by Western Reserve Land Conservancy staff
Survey Rating Scale – Sample Conditions

A. Excellent
- No visible signs of deterioration
- Well maintained and cared for
- New construction/renovation
- Historic detailing, unique

B. Good
- Needs basic improvements
- Minor painting
- Removal of weeds
- Cleaning

C. Fair
- Some cracking of brick or wood
- Major painting required
- Deteriorated cornice
- Crumbling concrete

D. Deteriorated
- Major cracking or brick, wood rotting
- Broken or missing windows
- Missing brick and siding
- Open holes

E. Unsafe/Hazard
- House is open and a shell
- Can see through completely
- House ransacked and filled with trash
- In danger of collapse
- Immediate safety hazard to neighborhood
Survey software loaded on IPad Mini Tablets

Licensed from ESRI (at beginning)

Licensed from Loveland Technologies (later)

Average survey 3 minutes per house

Check off items on IPad

Vacancy

Specific Itemized Conditions

Rating A-B-C-D-F

Take photo of house on IPad

Data uploaded to server

Surveyors stay on sidewalk
How The Survey Was Organized

One neighborhood at a time

Coordinated in advance with local community organizations

Teams given maps and assigned to work in sub-areas, census tracts or voting precincts
Two Forms of Quality Control

1. Weekly random samples of each surveyor’s work. Compare photo to rating for consistency.

2. All Properties that were found “vacant” and rated “D” or “F”
   – Photo was checked for accuracy and consistency.
Survey Cost

$138,000 Labor
$45,000 GIS Services
$16,000 Equipment & Software
$58,000 Report and Publication
$42,000 Admin

$299,000 Total
Special Considerations

Consider the season – winter may limit survey opportunities, but ice melt on roof and steam from vents are good indications of occupancy.

Winter may be safest in high crime areas.

Summer provides the greatest survey opportunities, but may be less safe; safest times will be early in the AM and mid day; much more street activity in later afternoon.

E.g., taking a photo of a house with a half dozen people hanging out on the porch. Make a note of those and come back another day.
Limitations of An Exterior Survey With Respect to Vacancy

When compared to US Postal vacancy data, survey vacancy varied in an interesting way.

In the most distressed neighborhoods surveyors tended to find slightly higher vacancy than reported by the US Postal data.

In the most stable neighborhoods surveyors tended to find slightly lower vacancy than
Limitations of An Exterior Survey
With Respect to Rating Conditions

Surveyors on the sidewalk cannot see wide-open or missing doors in the rear.

Cannot see conditions on the inside of the house.

Surveyors found 3,800 vacant Ds and Fs.

But we learned later that 1,400 other homes rated A, B and C had already been condemned by the City for severe deterioration on the inside.
How the Survey is Being Used
Issue or Problem

Cleveland Building and Housing Department has limited resources to inspect all problem properties.

Many of the most distressed properties have not yet been identified, inspected or condemned.

Solution

The Cleveland survey identified the vacant Ds and Fs; City can efficiently prioritize its resources on these.
Minimize Negative Consequences From Bulk Tax Lien Sales

Issue or Problem

Tax lien buyers bid on a portfolio of liens on thousands of properties; when abandoned properties are discovered, the buyer often chooses to not foreclose and these remain in limbo with a clouded title.

Solution

The Cleveland survey enables county officials to identify and remove vacant structures from the tax lien sale portfolio; these are redirected to a special Board of Revision tax foreclosure, and conveyed to our county land bank.
Issue or Problem

Low median sale prices in our weak real estate market can make the renovation of the most distressed houses financially infeasible.

Solution

The Cleveland survey enables housing and community development officials to identify the vacant homes that are still in relatively good condition and more viable for redevelopment.
Issue or Problem

Communities with limited resources are often operating “blind”, i.e. where are the vacant homes; where are the vacancies that are most distressed; what’s the extent of our problem and how much will it cost to fix it?

Solution

The Cleveland survey enables housing and community development officials to develop a realistic appraisal of the resources they will need to address the vacant property problem.
Neighborhood Typology

Issue or Problem

Several years ago Cleveland developed a neighborhood typology that would assist the City in identifying neighborhoods based on trends, e.g. “stable”, “in decline”, “tipping”, etc., but it was based on limited property condition data.

Solution

The Cleveland survey now enables planners to “ground-truth” the typology with more accurate data.

The survey provides a baseline for monitoring trends over time.
For Additional Information

Surveys of Ohio cities conducted by the Western Reserve Land Conservancy can be accessed here:

http://www.wrlandconservancy.org/publications-by-type/special-publications/property-inventory-reports/
Property data
Cleveland/Cuyahoga County, OH

Michael Schramm
Director of Information Technology and Research, Cuyahoga Land Bank
Research Associate, Center on Urban Poverty and Community Development,
Case Western Reserve University
Cuyahoga Land Bank

Doors opened in June 2009

Non profit, government purposed community improvement corporation

Three engines that make Ohio County Land Reutilization Corporations effective:

Dedicated funding stream (penalty and interest on collected delinquent property taxes)

Enhanced tax foreclosure laws

Outside of government (can transact on a dime)
Cuyahoga Land Bank production

Properties transacted – 6,818
Properties disposed – 4,567
Properties in inventory – 2,251
Properties demolished – 4,656
Facilitated renovations – 1,397
Center on Urban Poverty and Community Development

Founded in 1988
Our focus:
Neighborhood as a fundamental interface between individuals and systemic forces that drive opportunities.

Specific Goals:
- Through local engagement aim to build knowledge of what works in policy and practice
- Through national partnerships aim to bring community to the forefront of efforts to address social disadvantage (founding partner in National Neighborhood Indicators Partnership housed at the Urban Institute)

A creative team of faculty, staff, and students with diverse set of skills
Community Driven Partnerships
The Center’s areas of work

Indicator data at the neighborhood level – public portal

Integrated data at the individual level (persons and properties)

Research and evaluation projects focused on children, families, and neighborhoods
NEO CANDO

Started in 1992, evolved over time

Data focused at neighborhood level and below
Overview of Center Data

**PERSONAL RECORDS**
- Birth & death certificates
- Child welfare records
- Juvenile court fillings
- Preschool & childcare records
- Home visiting records
- Public school student records
- Public assistance monthly eligibility
- Homeless intake & service records
- County jail admission & discharge
- Lead screenings
- Electronic health records
- Public Housing

**PLACE-BASED RECORDS**
- Property characteristics
- Sales transactions
- Deed transfers
- Foreclosure fillings
- Tax delinquencies
- Postal vacancies
- Lank bank properties
- Parcel shape files
- Crime reports
- Housing code enforcement

**CHILD System**
- Probabilistic matching & individual record linkage
- Random identifier assigned
- Secure/private

**NEO CANDO**
- Open tabulated data by block group, census track, & special geographic units (e.g. SPAs, County Districts, Wards, etc.)
- Analysts download data (CSV files) — e.g. researchers, students, agency personnel, media, consultants
- Display on Google maps

**NST Application**
- Parcel level data and customized tools for community development sector
- Trained users view customized display
- Users can upload their own data & see it in display
- Filtering, tabulation, reporting tools

**External apps**
- Customized applications; API enabled.

DUA’s govern use
IRB oversight
Highly secure data transfer & management

- IRB certified staff produce de-identified datasets for approved research.
- Researchers analyze re-authorized data to produce insightful applications.
There’s gotta be a better way!
NST web app

Data system of parcel-level data for all parcels in Cuyahoga County, OH

Much of the data is updated weekly

Data is searchable, filterable, sortable, downloadable

All data are georeferenced into local geographies and target areas

Quick mapping functions

User-added data allowed
NST web app

Houses data relevant for neighborhood stabilization activities by city, county, and community development entities:

- Property transfers
- Tax records
- Foreclosure filings
- Sheriff’s sales
- Cuyahoga Land Bank records
- Cleveland Land Bank records
- Cleveland Building and Housing administrative data
- Cleveland Housing Court administrative data
- Geographic information
- Others
Ways data are acquired

• Email
• FTP
• DVD/CD
• XML Stream
• Dropbox
• Google Drive
• Screen Scraping
NST web app – create data from data

Vacant lot proxy

County bldg value

Cleveland demo

Private demo permit

Cuyahoga Land Bank demo

Cuyahoga County Demo Program Demo
Cuyahoga Land Bank – NEO CANDO – NST Data

Trajectory Tiers

Tier 1 – UNDER CONTROL
• In Cuyahoga Land Bank Inventory
• Properties pending transfer to Cuyahoga Land Bank
• In Municipal Land Bank Inventory
• State Forfeiture

Tier 2 – ABOUT TO BE UNDER CONTROL
• Nuisance Demo (Municipal/Cuyahoga Land Bank)
• Tax foreclosure Affidavit to Municipal Land Bank
• Tax foreclosure Affidavit to Cuyahoga Land Bank
• Fannie Mae
• HUD

Tier 3 – CAN STEER TO PRODUCTIVE USE
• Tax foreclosure affidavit – not sent to a Land Bank
• Tax foreclosure (not dismissed)
• Tax delinquency

Tier 4 – OTHER VACANT/BLIGHTED
• Mortgage foreclosure (not dismissed)
• Bank Owned (not HUD or Fannie Mae)
• Other vacant lot
Boundaries can be dynamically dissolved based on tier and queried. For example, tiers 1, 2, 3.
Community Development Uses of Data

Tax foreclosure candidate identification – Cuyahoga Land Bank

Researching buyers – Cuyahoga Land Bank

Code Enforcement Partnership – City of Cleveland

Trend analysis – VAPAC

Community Control – Housing Court
Tax foreclosure candidate project

20,000 tax foreclosure eligible properties in Cuyahoga County that aren’t already in foreclosure

County Gov’t has resources/capacity to file 4,000 new cases per year

Which candidates would be most desirable to land banks (vacant lots → City of Cleveland, vacant structures → Cuyahoga Land Bank)
Using NST for tax foreclosure batching

First filter → Fast track eligible and is NOT already in tax foreclosure system (vacant lot/vacant structure)

Second filter → eliminate properties that do not meet land bank acquisition criteria (commercial/industrial/large apartment bldgs)

Third filter → properties land banks want target areas (NSP2, CDC, Economic Development)

vacant structures in tipping point neighborhoods/suburbs
Property Profile System vs NEO CANDO/NST

NEO CANDO/NST are planning tools and visualization tools that mash together various public data sources for planning and outreach.

PPS – Land Bank Property Management Database

Cuyahoga Land Bank and currently used by 7 land banks to track inventory, attach pictures, create documents, email ticklers, etc.
Data Driven Policy Research

Recent study of Cleveland kindergarteners that used CHILD to predict kindergarten readiness based on early childhood exposures and housing.

We built analysis datasets used in studies conducted by Dynometrics.

Two demolition studies - Demolition raises property values and lowers foreclosure rate.

Renovation study with Cleveland Neighborhood Progress.
We have a bunch of data

What else should we do with it?????