Lifting Up What Works

California Unincorporated:

Mapping Disadvantaged Communities in the San Joaquin Valley

TECHNICAL GUIDE



In partnership with California Rural Legal Assistance, Inc. and California Rural Legal Assistance Foundation PolicyLink is a national research and action institute advancing economic and social equity by Lifting Up What Works[®].

Find this report online at www.policylink.org.

©2013 by PolicyLink All rights reserved.

Design by: Leslie Yang

COVER PHOTOS COURTESY OF: Hector Gutierrez.

s-Madera Ranchos

California Unincorporated: Mapping Disadvantaged Communities in the San Joaquin Valley

In partnership with California Rural Legal Assistance, Inc. and California Rural Legal Assistance Foundation

// TECHNICAL GUIDE



This technical guide will help those interested in replicating this model in their area; it provides a step-by-step detailed account of technical mapping methods and additional sources for data.

This project used Geographic Information Systems (GIS) technology to identify unincorporated communities that are small, densely populated, low-income, and, in some cases, unmapped. These are places that are more likely to have poor quality infrastructure or service provision. This research was developed using the Geographical Information System ESRI ArcView for ArcMap 9.3.1 with the Spatial Analyst 9.3 extension running on a Windows machine with an Intel Core 2 Quad CPU @ 2.40GHz with 2.00 GB of RAM.

The following indicators were used to help identify disadvantaged unincorporated communities (DUCs), defined as disproportionately low-income places that are densely settled and not within city limits.

Methods

This research used three basic parts to identify disadvantaged unincorporated areas:

- 1. Unincorporated Status. Because unincorporated areas are those outside of city boundaries, we used boundary shapefiles from cities, counties, or the U.S. Census Bureau to determine incorporation status.
- 2. Parcel Density. For the purposes of this research, we were more concerned with small communities that resemble the density of suburban and urban communities and less concerned with less populated, less dense rural areas. Parcels that are small and close together are typically more urban. To determine the density of communities, we used **parcel** shapefiles.
- 3. Low income. As our focus is on addressing the unique infrastructure challenges faced by low-income communities, this research used U.S. Census block group data to compare the income status of local households relative to households across the state.

Beyond the three basic layers, we applied additional indicators to further verify that places highlighted by the analysis are indeed underserved communities. We used data on land use and aerial photography, for example, to help filter out agricultural land, undeveloped land, and resort communities. The following section presents more detail on how a variety of datasets were used and the challenges that they present.

1. Determining Unincorporated Status

Unincorporated places are areas outside of city boundaries. This research starts by identifying the city boundaries and determining which communities are outside of city boundaries. In California, there are several sources for city boundary shapefiles. Most city governments, usually through their planning departments, maintain shapefiles of city boundaries, but often counties and Local Area Formation Commissions (LAFCOs) also have copies. These agencies may also have **sphere of influence** boundaries which show "the probable physical boundaries and service area"¹ of a city (or special district) in the future.

The U.S. Census Bureau also tracks unincorporated areas that are closely settled and calls them **Census Designated Places** (CDPs). According to the U.S. Census, CDPs are named, unincorporated communities that generally contain a mix of residential, commercial, and retail areas similar to those found in incorporated places of similar sizes. Map 1 shows CDPs located in Fresno County, shaded in deep yellow, as an example. (Data for all maps in this guide are derived from PolicyLink analysis of 2000 U.S. census, county parcel data, and Google Earth.)

¹ California Government Code Section 56076.

PolicyLink



MAP 1: Fresno County Showing City Boundaries, Spheres of Influence, and Census Designated Places

2. Determining Unincorporated Status and Density

For an approximation of neighborhood density, we calculated the number of parcels per square mile. The goal of this step was to find clusters of development outside of city limits that are similar in density to existing Census Designated Places.

Parcel densities were calculated using the centroid (or middle point) of each parcel. As Map 2 illustrates, each point on the map represents a parcel. Areas with many dots close together indicate places with greater densities of housing or commercial development. (A description of how centroids were created is on page 12.)

Because this research focuses on unincorporated areas, only parcel centroids outside of the city limits were selected. Centroids that were not in agricultural land were also selected; the yellow area in Map 3 illustrates areas considered agricultural. The Farmland Monitoring and Mapping Program (FMMP) data are updated every two years to track how land is currently being used. Categories of "developed," "rural residential," and "agricultural" helped to refine which parcels were of interest.

Map 4 shows the centroids that are outside of city limits and not in land categorized as agricultural. From these remaining centroids, we calculated parcel density using the spatial analyst kernel density tool. (See page 12 for the kernel density settings.)

MAP 2: Sample of Parcel Centroids



MAP 3: Centroids with City Limits and Agricultural Land







PolicyLink

This research focuses on places where the density is similar to that of cities or existing Census Designated Places.

For reference, we calculated densities of a few CDPs to determine a minimum threshold value for community density. This calculation was based on the developed part(s) of each CDP. We did not use the entire area because CDP boundaries are often larger than the developed part(s) of the community. Unincorporated areas that were at least as dense as current CDPs (approximately 250 parcels per square mile) were selected. Map 5 shows the resulting density calculation for unincorporated areas; Map 6 shows only the areas that are above 250 parcels/sq. mi.

Some very small and more rural communities were filtered out of this model, but could be included in the future.² Using a density threshold value lower than 250 parcels per square mile only picked up moderately developed areas that didn't appear to be communities. This method does overlook densely developed areas that are built within one un-subdivided parcel, like mobile home and trailer parks, but we would need to apply density calculations to a layer with buildings and structures in order to identify these types of communities. We do not currently have a data layer of this type that covers the entire San Joaquin Valley.

MAP 5: Density Calculation for Centroids



MAP 6: Areas with Density Greater than Threshold



² Selecting parcels within census blocks with a population greater than zero can be used at this point or later in the model to filter out subdivided, undeveloped land.

3. Approximating Low-Income Status

As mentioned previously, the communities that face significant infrastructure challenges are often lowincome. Early on, we tested several measurements and thresholds to approximate low-income status: poverty rates above 30 percent, poverty rates above 40 percent for children, median family income, median household income compared to the county's, and median household income compared to that of the state's. The indicator that seemed most inclusive of the communities we were interested in was median household income below 80 percent of the state's median household income. This is a benchmark used in several state-level infrastructure funding programs that target lowincome communities, including the Safe Drinking Water State Revolving Loan Fund and the Storm Water Management Program. In 2000, the median household income of the state of California was \$47,493, so any census block group with a median income of less than \$37,994 was included in our analysis.

As Map 7 shows, block groups can be large in size in rural areas (like the eastern edge of Fresno County) because there are fewer people living there. There are several limitations to this data. It is old data; there were several shifts in local and national economies between 1999 and now that undoubtedly affected families in the San Joaquin Valley for the worse. Secondly, the U.S. Census Bureau has historically undercounted rural populations, people of color, and those who are not native English speakers.³ Lastly, the size of a census block group is often much larger than the small communities of concern. For example, a census block group may include many wealthy households as well as a cluster of low-income families. This block group will have a resulting high median income and may be excluded from our model. One example of this, illustrated in Map 8, is the small community of Drummond and Jensen on the southeast edge of the City of Fresno. We know from our work that the community is low-income, but is part of a census block group that has a median household income greater than 80 percent of the state's median.

Map 7 also shows all of the block groups mapped by percentage income compared to the state.

MAP 7: Census Block Groups in Fresno County



MAP 8: Low-Income Community in Higher-Income Block Group



MAP 9: Low-Income Block Groups



³ Ed Kissam and Ilene Jacobs. Census Undercount and Immigrant Integration into Rural California Life. 1999. UC Davis <u>http://migration.ucdavis.edu/cf/more.php?id=112_0_2_0</u>.

PolicyLink

The next step, illustrated in Map 9, was to select only those block groups where median household income was less than 80 percent of the state's. Refer to page 13 for a more detailed methodology on how we identified low-income block groups were identified.

Combining the Data

As Map 10 shows, after locating low-income communities, we layered the results from the density calculation shown in orange. Map 11 shows the next step of selecting only the dense areas that intersected with the low-income areas.

Visual Inspection

Areas that were, by visual inspection with aerial imagery (and Google's street view), obviously new and/or not low income were removed. These are most commonly found in areas at the edge of cities which had been agricultural fields during the 1990 census but have since been developed (Map 12). Areas less than threequarters of an acre in size were also removed as they were too small, often containing only one or two houses, if any.

Layering the data and excluding places based on visual inspection resulted in 525 distinct areas across the San Joaquin Valley. These areas do not have clear boundaries because the density calculations result in amorphous spots that are not bounded by parcels or major roads. A closer inspection of aerial maps can often help identify more specific neighborhood boundaries.

The next section describes the census data used to approximate the population, the demographics, and more specifics about the income for the places identified.

MAP 10: Low-Income Areas with High Density Overlay



MAP 11: Selection of Overlapping Areas



MAP 12: Visual Inspection



4. Estimating the Population and Demographics of Disadvantaged Unincorporated Communities

Population, demographics, and the number of lowincome households were calculated using census blocks and then intersected with the boundaries of DUCs. The result was an estimated value based on the area of each DUC. Where a DUC boundary extended across two or more census blocks with different density values, estimates were calculated for each portion and aggregated by the ID number of the DUC. The population and demographic estimates were largely based off of a methodology created by Dhayna Elizabeth Elias and Rebecca Baran-Rees, Master of City Planning students at Cornell University.⁴

⁴ These estimates were developed as part of the "Introduction to GIS for Planners" course taught at Cornell University by Associate Professor Stephan Schmidt.

Model Requirements

The following pages describe in more detail the sequence of formatting and processing data in ArcView 9.3 for this model. As with many GIS procedures, there are several courses of action available to achieve the same results. The following methods produced consistent results for repeated runs of this model.

This model was developed and run using ArcView 9.3.1 with Spatial Analyst on a Windows XP machine with an Intel Quad CPU, 2.4 GHz processor, and 2.0 Gb RAM.

Gathering and Preparing Data for the Model

Data were collected and organized prior to running the model. Data providers include online resources, County Planning departments, LAFCOs, and project partners.

Required Data for the Model

- Parcel Geometry (no attributes required)
- City Limits
- City Spheres of Influence (SOIs)
- 2000 Census Income Data, by Census Block Group
- Study Area Boundary shapefile

Additional Datasets Used

- Farmland Mapping and Monitoring Program (FMMP)
- Census Designated Places (CDP)
- Geographic Names Information System (GNIS) Place Names
- Census Blocks for Population Value
- Aerial Imagery for Study Area

Data Descriptions and Preparation

Parcels – Parcel geometry shapefiles (without attributes) were gathered for the study area to calculate density of settled land, independent of incorporation status. *Create parcel centroids and merge into one study areawide parcel centroid shapefile.* (see instructions below)

City Limits – The most current city limits shapefiles were obtained from county planning departments and LAFCOs to indicate incorporated status. U.S. census data were used where county specific coverage was unavailable. *Merge all City Limits data into one study areawide City Limits shapefile.*

City Spheres of Influence (SOI) – SOI shapefiles from County Planning Departments and Local Agency Formation Commissions were used to determine if communities were in the urban fringe. In counties where SOI coverage was not available, a one-mile buffer of city limits was used as a proxy. *Merge all SOI data into one study areawide SOI shapefile.*

Census Income Data by Block Group – Median household income for each block group (attribute P053001 in the census) was compared to statewide median household income. (For 2000, this was \$47,493.)

Study Area Boundary – The eight-county study area (San Joaquin Valley) was dissolved to one shapefile.

Farmland Mapping & Monitoring Program

(FMMP) – FMMP produces maps and statistical data used for analyzing California's agricultural resources. The FMMP classifies types of land including 'Prime Farmland' and 'Grazing Land', among several others. The maps are updated every two years with the use of a computer mapping system, aerial imagery, public review, and field reconnaissance which is helpful for accuracy. http://www.conservation.ca.gov/dlrp/FMMP/Pages/ Index.aspx.

Agricultural lands from this dataset were used to filter out completely undeveloped areas from the model. The FMMP data is derived from lowaltitude aerial imagery and represents what is "on the ground" versus what is surveyed (parcels) or automatically interpreted from satellite imagery. This is a very useful dataset that may not be available in other areas of the country.

Census Designated Places (CDPs) – According to the U.S. Census Bureau, a CDP is a densely settled concentration of population that is not within an incorporated place, but is locally identified by a name. CDPs were used to compare community density, determine if an unincorporated area had been previously identified, and gather existing place name information.

Geographic Names Information System

(GNIS) – A point shapefile of current and historic place names available from the USGS. Used for community identification where applicable.

Census Blocks – According to the U.S. Census, a block is an "area bounded on all sides by visible features, such as streets, roads, streams, and railroad tracks, and by invisible boundaries, such as city, town, township, and county limits, property lines, and short, imaginary extensions of streets and roads. Generally, census blocks are small in area; for example, a block bounded by city streets. However, census blocks in remote areas may be large and irregular and contain many square miles." Census blocks provided population data (not income data) and was used to remove areas with no population from the calculation.

Aerial Imagery – <u>ESRI's World Imagery</u> layer was used as a visual reference while running and checking the model. It is hosted online and is less memory intensive than storing and navigating other high-resolution aerial imagery. <u>http://www.esri.com/</u> software/arcgis/arcgisonline/standard-maps.html.

Running the Model

Creating Centroids from Parcel Shapefiles

Parcel centroids were used to calculate the density of settlements across the study area. The following method worked to create a study areawide parcel centroid (point) shapefile:

- Add each parcel polygon shapefile to an ArcMap window and remove the projection from the data frame, keeping WGS 1984 coordinate system.
- Working in an unprojected coordinate system, create x/y values using the "Calculate Geometry" function in the attribute table. These coordinates should be created in new columns (ie: "X_coord" and "Y_coord") with a field type of "double," a precision value of 10, and a scale value of 5, with units set to decimal degrees.
- In a new ArcMap window, add the database file that has the newly calculated X and Y coordinates.
- Set a relevant projection for the geographic region of the study area (in this example, it was CA State Plane IV, Feet).
- Create X/Y events from the new coordinate values from each dbf.
- Save each as a new shapefile. Use the same coordinate system as the data frame.
- Add newly projected centroid point shapefiles to view.
- Merge into one study areawide shapefile.

Filtering Parcel Centroids before Density Calculations

Removing extraneous parcel centroids before running density calculations yields clearer results. Parcels inside city limits, those in census blocks with zero population, and those on completely undeveloped land were excluded from density calculations.

- Using the compiled city limits shapefile, select only centroid points that are outside of city limits (including areas in the middle of, but not contained by, city limits).
- If applicable, remove centroid points that are within census blocks of zero population.
- If applicable, select only centroid points that are within developed land areas. We used the Farmland Mapping and Monitoring Program data for this step. The FMMP includes two "developed" land type categories (Urban & Built Up Land [D], and Rural Residential [R]) to verify that areas of high parcel density are developed. Areas with high parcel density that are not developed indicate "empty" subdivisions, usually far from incorporated areas or infrastructure services ("paper subdivisions").

Calculating Parcel Density in Spatial Analyst

Concentrations of parcels were identified by calculating density across the study area. Instead of a visual dot density represented by parcel centroids, density calculation yields a continuous raster or grid with cell values reflecting parcel density per square mile. This format is useful in further model steps.

- Using the spatial analyst extension, open the kernel density tool (found in the Arc Toolbox: Spatial Analyst Tools > Density > Kernel Density).
- Input the prepared parcel centroid shapefile and use the following settings: No population field value, cell size = 100, search radius = 1000, units = sq mi (these settings yielded reasonable densities for our study. See "Discussion of Parcel Density Values" below).

- Areas within city limits and other excluded lands (if applicable) should be set to zero. (The Kernel density calculation can expand grid values into these excluded areas from adjacent parcel centroid clusters. Alternatively, excluded areas can be masked during the model run).
- Make a grid where city polygons have values greater than zero.
- Make a grid where the complete study area has a value of zero.
- Use the mosaic function to combine these two resulting grids. Do not average the values and enter the cities grid on top first.
- Use raster calculator with the following statement:

setnull mosaic_cities_studyarea> 0 (kerneldensity_grid, no data)

(This yields a grid with density values only outside of cities and excluded lands)

 From this grid, select parcel density values > 250 (yields a 0/1 grid where 1 = parcel density > 250 parcels per sq mi.

Discussion of Parcel Density Values

To compare and calibrate our parcel density threshold of 250 parcels per square mile, parcel densities of Census Designated Places (CDPs) were calculated to verify an appropriate density threshold (parcels / sq mi) when selecting communities in unincorporated areas. Data used in this process were Fresno county parcel centroids, Fresno County CDP boundaries, and Fresno FMMP (Farmland Mapping and Monitoring) data from 2008.

Use in SJV Communities Model

The FMMP data were used to identify developed areas within CDPs. (CDP boundaries are often larger than the actual community and can include areas of farmland or forest.) Urban / Built-Up (D) and Rural Residential (R) land cover types were used to select parcel centroids within existing Fresno County CDPs. Parcel density was calculated using the Kernel density tool (cell size = 100, search radius = 1000, units = sq mi). The results were clipped to the (FMMP) developed areas within CDPs. A visual check concluded that a majority of developed areas within CDPs had a parcel density greater than 250 parcels / sq. mi. Developed areas within CDPs that had less than 250 parcels / sq. mi. were more varied in their development type, i.e., ranchette, warehouse, etc.

The initial choice of 250 parcels / sq. mi. appears to be a reasonable threshold when selecting community clusters from across the study area. Where clusters had been identified in CDPs, the boundaries defined for areas greater than 250 parcels /sq. mi. contained a majority of the developed area in that CDP.

Selecting Low-income Block Groups

Income data from the 2000 census were available at the Block Group Level. We chose areas at or below 80 percent of the State Median Household Income to identify low-income, or disadvantaged, areas of our study.

- From the Census Block Group shapefile, make a grid (from 2000 Census Block Group shapefile attribute 53001 or, Median Household Income)
 0.001 cell size was the appropriate value for this model run.
- Make another study-area wide grid with the value set to the state's Median Household Income.
- In Raster Calculator, divide **53001** grid by the **State Median Household Income grid.**
- From the resulting point grid, calculate areas with a value <= 0.8 as a new grid (which will have values of 0 and 1, where 1 = areas less than or equal to 80% of the State's Median Household Income)

Identifying Low-Income, High-Density Areas

Areas of high parcel density outside of city limits were intersected with those of low-income to show potentially disadvantaged unincorporated communities.

- In Raster Calculator, add the high density and low-income grids to identify overlapping areas ("High Density GRID " + "Low-income GRID") = 2 (this yields only areas of overlap).
- Convert the resulting raster to a shapefile using Arc toolbox > conversion tools > from Raster
 > Raster to Polygon. Do not use the "Simplify Polygon" option.
- Select gridcode = 1 (ensure that this is the lowincome, high density area represented in the GRID), and save as a new shapefile.

Additional Filters

An additional combination of selections were used as a final check of the identified communities.

- Areas that are (by visual inspection with aerial/ satellite imagery) obviously new and/or not low income were deleted. These are most commonly found in areas at the edge of cities which were "low-income" agricultural fields during the 1990 census and have since been developed.
- Areas that were less than three-quarters of an acre in size. These are too small to be considered communities and often contain only one or two houses, if any, as shown in Map 13.
- Any obvious narrow "slivers" that were a result of layer alignment issues (i.e., along city limits and census tract overlaps) were deleted. These are most often artifacts of registration inconsistencies between data sources.

Classifying Areas as Fringe, Legacy, or Island

The following method was used to classify the identified areas according to their spatial relationship with incorporated areas. We termed this attribute as "Placement."

- Using "Select by Location," areas outside of City Spheres of Influence (SOIs) were selected and classified as Legacy.
- Using "Select by Location," areas inside of City SOIs (or within 1 mile of City Limits where

MAP 13: Identifying Areas Smaller than 3/4 acre



City SOIs were unavailable) were selected and classified as **Fringe**.

• Areas completely inside City Limits (visual selection) were classified as **Island**.

Naming

We spatially joined the identified communities to the CDP and GNIS datasets to add names to the places that were identified. Community names were also collected in hard-copy form from participants of a July 2010 workshop hosted by California Rural Legal Assistance in Fresno. These were digitized and attributed where applicable. Communities without a name listed in these sources maintained the unique numeric ID assigned by ArcMap.

Checking Results

While a comprehensive field check was not in the scope of this work, we used Google Street View and Google Earth to look at identified communities. Collaboration across researchers was possible by use of a Google Earth KMZ of the model results. Notes and corrections were added to the community names in Google Earth and then incorporated to the GIS layer.

Calculating Estimated Population, Demographics, and Low-income Households

Estimating Population and Demographics for DUCs

Final population and demographics of the modeled DUCs were calculated as follows:

- Pull census blocks with population and race.
- People per-square-mile and race per-squaremile were calculated for each block (All, Non-Hispanic Black, Non-Hispanic Asian, Non-Hispanic White, and Non-Hispanic Other including American Indian, Pacific Islander, Other, two or more races).
- Using the union function, join this Census Block shapefile with the modeled DUC polygon shapefile.
- Calculate the square mileage for area of each DUC / Block union segment.
- The estimated population of each segment was calculated by multiplying the block value (people-per-sqmi) by the sqmi of unioned segment.
- Population and demographic values of each segment were aggregated by DUC_ID in SPSS.
- Join the aggregated table to the original DUC shapefile.
- As noted previously, small DUCs within large rural census blocks yielded lower than expected population estimates due to the low population density of the block.

Estimating Number of Low-income Households in DUCs

For a comparison of low-income households in the county, unincorporated areas, and DUCs, the following method was used to calculate the number of households within DUCs:

- Pull census block groups with number of households per income range.
- We wanted the total number of households that were at or under 80 percent of the California Median Household Income value, \$37,994 (\$38,000).
- Add the number of households in each range under \$38,000 to obtain the total number of low-income households in the block group.
- Calculate the area of each block group in square miles.
- Calculate the number of low-income households per square mile and within the block group.
- Using the union function, join this block group shapefile with the modeled DUC shapefile.
- Calculate the square miles for the area of each DUC / Block Group segment.
- Multiply the block group value (people-per-sqmi) by the sqmi of the unioned segment. Aggregate all household estimates of each segment in SPSS.
- Join the aggregated table to the original DUC shapefile.

Bonadelle Ranchos-Madera Ranchos

Authors' Biographies

Chione Flegal is an associate director with PolicyLink. She works to ensure that infrastructure and land use policy promote economic, social, and environmental equity. She leads the organization's policy work in California, as well as efforts to address disparities faced by low-income unincorporated communities.

Solana Rice is a senior associate with PolicyLink. As a researcher and policy advocate, she works to advance opportunity for low-income communities and communities of color by promoting equitable solutions to community development through wealth building, small business development, and other economically focused strategies.

Jake Mann is a temporary research assistant with PolicyLink supporting the mapping and analysis for the Community Equity Initiative.

Jennifer Tran is a program associate with PolicyLink. Her work focuses on equitable economic growth, conducting research, data analysis, and GIS mapping.

Acknowledgments

The Community Equity Initiative would like to thank the many people who have made this work possible. Foremost, we would like to acknowledge the following for their continued commitment and dedication to improving California's disadvantaged unincorporated communities:

Martha Guzman-Aceves Brian Augusta Erika Bernabei Kara Brodfueherer Juan Carlos Cancino Cara Carrillo Amparo Cid Maria Sofia Corona Olivia Faz Niva Flor Verónica Garibay Rubén Lizardo Emily Long Laura Massie Cristina Mendez Enid Picart Walter Ramirez Ruby Renteria Victor Rubin Phoebe Seaton Esmeralda Soria Elica Vafaie Ariana Zeno

Thank you to the following who contributed content and research expertise:

Carolina Balazs Rebecca Baran-Rees John Capitman Dhanya Elizabeth Elias Ann Moss Joyner

Thanks to the PolicyLink editorial and production teams.

Finally, this work would not have been possible without the generous support of:

The James Irvine Foundation The California Endowment The Open Society Foundations The Women's Foundation of California The Kresge Foundation Lifting Up What Works®



Headquarters:

1438 Webster Street Suite 303 Oakland, CA 94612 t 510 663-2333 f 510 663-9684

Communications:

55 West 39th Street 11th Floor New York, NY 10018 t 212 629-9570 f 212 768-2350

www.policylink.org

©2013 by PolicyLink All rights reserved.