HCV and PBS8 by Chicago communities

June 9, 2022

This is the notebook with which I demonstrate the distribution of two affordable housing programs: Housing Choice Vouchers (HCV) and Project-Based Section 8 (PBS8) in Chicago. I start with a dataset from HUD containing information on HCV and PBS8 data in each census tract in Illinois, a dataset that converts each census tract to the community area to which it belongs, and a shapefile for Chicago community areas. I process and merge these datasets to graph the number of occupied units for HCV and PBS8 in each community area in Chicago.

[1]:
```
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
import numpy as np
%matplotlib inline
```

[2]:
```
assIL = pd.read_csv('All Assisted Housing in IL by census tract.csv')
```

[3]:
```
# Filter out Cook County data
asscook0 = assIL[assIL.Name.str.contains('Cook County')]
# Eliminate all negative entries
d1 = asscook0.select_dtypes(include = 'number').copy()
d1.where(d1>0, 0, inplace=True)
d2 = asscook0.select_dtypes(exclude = 'number')
asscook = pd.concat([d2, d1], axis=1)
```

[4]:
```
# Convert entries under "Codes" into integers
def replace_func(x):
    return int(x[2:-1])
a=asscook.Code.map(replace_func)
asscook = asscook.assign(Code=a)
```

[6]:
```
# Read the dataset converting census tracts to community areas in Chicago
conv = pd.read_excel('/Users/bling/Documents/21-22/Housing policy/
 ↪Census_Tracts_in_Chicago_Community_Areas.xlsx',␣
 ↪sheet_name='TractsCommunityAreas')
conv.columns
```

[6]:
```
Index(['Tract', 'Label', 'CommunityAreaNumber', 'CommunityAreaName'],
dtype='object')
```

```
[7]: # Rename the asscook column to prepare for merge
     asscook.rename({'Code':'Tract'}, axis ='columns', inplace=True)
```

```
[9]: # Merge the two datasets to get the affordable housing data for each tract in␣
      ↪each community area in Chicago
     ahchi = asscook.merge(conv, how = 'inner', on ='Tract')
```

```
[10]: ahchi.columns
```

```
[10]: Index(['Program label', 'Sub-program', 'Name', 'Tract',
              'Congressional District', 'CBSA', 'PLACE', 'Latitude', 'Longitude',
              'State', 'PHA Total Units', 'HA category', 'Summary level', 'Program',
              'Subsidized units available', '% Occupied', '# Reported', '% Reported',
              'Average months since report', '% moved in past year',
              'Number of people per unit', 'Number of people: total',
              'Average Family Expenditure per month ($$)',
              'Average HUD Expenditure per month ($$)', 'Household income per year',
              'Household income per year per person', '% $1 - $4,999',
              '% $5,000 - $9,999', '% $10,000 - $14,999', '% $15,000 - $19,999',
              '% $20,000 or more',
              '% Households where wages are major source of income',
              '% Households where welfare is major source of income',
              '% Households with other major sources of income',
              '% of local median (Household income)', '% very low income',
              '% extremely low income', '% 2+ adults with children',
              '% 1 adult with children', '% female head',
              '% female head with children',
              '% with disability, among Head, Spouse, Co-head, aged 61 years or less',
              '% with disability, among Head, Spouse, Co-head, aged 62 years or older',
              '% with disability, among all persons in households',
              '% 24 years or less (Head or spouse)',
              '% 25 to 49 years (Head or spouse)', '% 51 to 60 (Head or spouse)',
              '% 62 or more (Head or spouse)', '% 85 or more (Head or spouse)',
              '% Minority', '%Black Non-Hispanic', ' %Black Hispanic',
              '%Native American Non-Hispanic',
              '%Asian or Pacific Islander Non-Hispanic', '% Hispanic',
              'Average months on waiting list', 'Average months since moved in',
              '% with utility allowance', 'Average utility allowance $$',
              '% 0 - 1 bedrooms:', '% 2 bedrooms', '% 3+ bedrooms', '% Overhoused',
              '% in poverty (Census tract)', '% minority (Census tract)',
              '% single family owners (Census tract)', '%White Hispanic',
              '%White Non-Hispanic', '% Multiple Race', '%Other Hispanic', 'Label',
              'CommunityAreaNumber', 'CommunityAreaName'],
             dtype='object')
```

```
[11]: key_vars = ['Program label','Tract','Program','Subsidized units available', '%␣
       ↪Occupied',
          '% moved in past year','Number of people per unit', 'Number of people:
 ↪ total',
          'CommunityAreaNumber', 'CommunityAreaName']
ahchi_num = ahchi[key_vars]
```

[12]:
```
# Calculate the product of two columns (occupancy rate & number of available␣
 ↪units) to get occupied units
ahchi_num['Occupied Units']=ahchi_num['Subsidized units available']*ahchi['%␣
 ↪Occupied']/100
ahchi_num['Occupied Units']=ahchi_num['Occupied Units'].round(decimals = 0)
```

```
/var/folders/5g/vyw42tkj7f32hbm585d3jxzr0000gn/T/ipykernel_712/77969010.py:1:
SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-
docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
  ahchi_num['Occupied Units']=ahchi_num['Subsidized units available']*ahchi['%
Occupied']/100
/var/folders/5g/vyw42tkj7f32hbm585d3jxzr0000gn/T/ipykernel_712/77969010.py:2:
SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-
docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
  ahchi_num['Occupied Units']=ahchi_num['Occupied Units'].round(decimals = 0)
```

[13]:
```
# Filter out data for HCV
ahchi_num_hcv = ahchi_num.loc[ahchi_num['Program label']=='="Housing Choice␣
 ↪Vouchers"']
ahchi_num_hcv.drop (columns = 'Program label')
```

[15]:
```
# Which census tract saw the biggest influx (highest % moved in past year)?
ahchi_num_hcv.nlargest(5,'% moved in past year')
```

[15]:
```
                  Program label        Tract  Program  \
352   ="Housing Choice Vouchers"  17031081000        3
33    ="Housing Choice Vouchers"  17031010702        3
710   ="Housing Choice Vouchers"  17031221500        3
1989  ="Housing Choice Vouchers"  17031833000        3
1568  ="Housing Choice Vouchers"  17031610300        3

      Subsidized units available  % Occupied  % moved in past year  \
352                          406          91                    43
33                            68          91                    30
```

```
710                          168          91                    28
1989                         116          91                    28
1568                          13          91                    27

      Number of people per unit  Number of people: total  CommunityAreaNumber  \
352                         1.1                      385                    8
33                          1.7                       99                    1
710                         1.3                      192                   22
1989                        1.1                      110                   28
1568                        2.4                       26                   61

     CommunityAreaName  Occupied Units
352    Near North Side           369.0
33         Rogers Park            62.0
710       Logan Square           153.0
1989    Near West Side           106.0
1568          New City            12.0
```

[16]:
```
# Sum up all occupied HCV units by community area
hcv_com=ahchi_num_hcv.groupby('CommunityAreaNumber').sum()
hcv_com=hcv_com.drop(columns=['Tract','Program','% Occupied','% moved in past␣
 ↪year',
                             'Number of people per unit'])
```

[21]:
```
# Import the shapefile for community areas in Chicago
import geopandas
com_map= geopandas.read_file('/Users/bling/Documents/21-22/Housing policy/
 ↪Geographic data/Boundaries - Community Areas (current)/
 ↪geo_export_8de711f3-84a7-4e58-9b48-fa4037567404.shp')
```

[22]:
```
# Convert the area numbers to integers, rename the column
com_map['area_numbe']=pd.to_numeric(com_map['area_numbe'])
com_map.rename({'area_numbe':'CommunityAreaNumber'}, axis='columns',␣
 ↪inplace=True)
```

[23]:
```
# Merge the community area map with the dataset containing data for each census␣
 ↪tract
map_w_hcv = com_map.merge(hcv_com, how = 'inner', on ='CommunityAreaNumber')
```

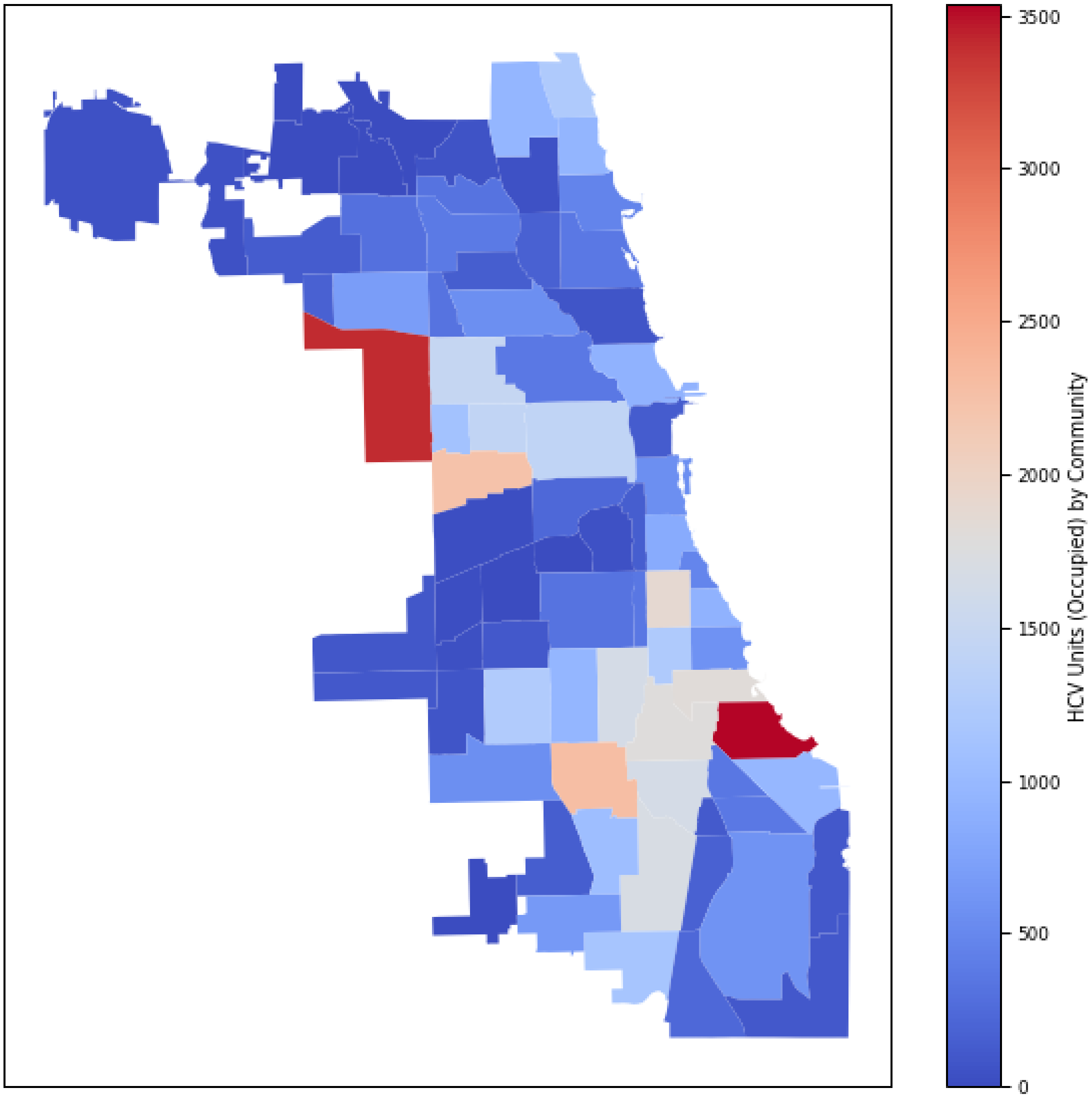
[24]:
```
# top 5 community areas with the largest number of HCV occupied units
map_w_hcv.nlargest(5,'Occupied Units')
```

[24]:
```
    area area_num_1  CommunityAreaNumber  comarea  comarea_id  \
38   0.0         43                   43      0.0         0.0
25   0.0         25                   25      0.0         0.0
69   0.0         71                   71      0.0         0.0
29   0.0         29                   29      0.0         0.0
```

```
3    0.0           38                    38      0.0         0.0

          community  perimeter    shape_area     shape_len  \
38      SOUTH SHORE        0.0  8.181272e+07  44249.646117
25           AUSTIN        0.0  1.992542e+08  75226.474917
69   AUBURN GRESHAM        0.0  1.050654e+08  46757.721716
29   NORTH LAWNDALE        0.0  8.948742e+07  44959.459663
3   GRAND BOULEVARD        0.0  4.849250e+07  28196.837157

                                             geometry  \
38  POLYGON ((-87.54398 41.75515, -87.54400 41.755…
25  POLYGON ((-87.78942 41.91751, -87.78927 41.917…
69  POLYGON ((-87.63990 41.75615, -87.63990 41.755…
29  POLYGON ((-87.72024 41.86987, -87.71965 41.869…
3   POLYGON ((-87.60671 41.81681, -87.60670 41.816…

    Subsidized units available  Number of people: total  Occupied Units
38                        3889                     6410          3539.0
25                        3748                     6952          3412.0
69                        2508                     5253          2282.0
29                        2455                     5405          2228.0
3                         2071                     3428          1884.0
```

[31]:

```
sns.set_context('paper')
ax1 = map_w_hcv.plot('Occupied Units',legend=True, cmap='coolwarm',
               legend_kwds={'label':'HCV Units (Occupied) by Community'},
              figsize=(10,10))
ax1.set_yticks([])
ax1.set_xticks([])
fig = plt.gcf()
fig.savefig('/Users/bling/Documents/my_first_figure.pdf')
```



```
[32]: pbs8 = ahchi_num.loc[ahchi_num['Program label']=='="Project Based Section 8"']
      pbs8.drop(columns=['Program label','Tract','Program','Number of people per␣
       ↪unit'])
```

```
[32]:       Subsidized units available  % Occupied  % moved in past year  \
      3                            304          94                     9
      14                            56         100                    16
      20                           171          98                     7
      28                            41         100                    10
      34                            25          91                     9
      …                             …           …                     …
      2236                          28          96                     7
      2240                          31          97                     3
```

```
2253                          10           0                     0
2257                          98         100                     3
2262                         168          96                     7

      Number of people: total  CommunityAreaNumber CommunityAreaName  \
3                         579                    1       Rogers Park
14                         63                    1       Rogers Park
20                        183                    1       Rogers Park
28                         46                    1       Rogers Park
34                         55                    1       Rogers Park
…                         …                    …               …
2236                       66                   29    North Lawndale
2240                       87                   29    North Lawndale
2253                        0                   38  Grand Boulevard
2257                      102                    5      North Center
2262                      266                   42          Woodlawn

      Occupied Units
3              286.0
14              56.0
20             168.0
28              41.0
34              23.0
…                …
2236            27.0
2240            30.0
2253             0.0
2257            98.0
2262           161.0

[284 rows x 7 columns]
```

[34]:

```
pbs8.nlargest(5, 'Number of people per unit')
# It looks like that projects in Grand Boulevard and Woodlawn tend to be a bit␣
 ↪overcrowded
# because the average number of people per unit exceeds 3 people.
```

[34]:

```
                   Program label        Tract  Program  \
2063  ="Project Based Section 8"  17031836000        5
1262  ="Project Based Section 8"  17031420800        5
968   ="Project Based Section 8"  17031271300        5
2091  ="Project Based Section 8"  17031836900        5
2222  ="Project Based Section 8"  17031842500        5

      Subsidized units available  % Occupied  % moved in past year  \
2063                          11          92                    17
1262                          11          95                     0
```

```
968                           53          91                    26
2091                          85          91                    11
2222                         694          96                    15

      Number of people per unit  Number of people: total  CommunityAreaNumber  \
2063                        3.9                       47                   38
1262                        3.5                       39                   42
968                         3.2                      160                   27
2091                        3.1                      251                   27
2222                        3.0                     1992                   69

           CommunityAreaName  Occupied Units
2063         Grand Boulevard            10.0
1262                Woodlawn            10.0
968       East Garfield Park            48.0
2091      East Garfield Park            77.0
2222  Greater Grand Crossing           666.0
```

[35]:

```
# Where are projects that are under-utilized?
pbs8.loc[(pbs8['% Occupied']<80)&(pbs8['Subsidized units available']>10)]
# In West Town, Kenwood, Woodlawn, and Logan Square, the available units tend to␣
 ↪< 80% utilized.
# Notably, in Woodlawn there are 291 subsidized units available, with only 71%␣
 ↪being utilized
```

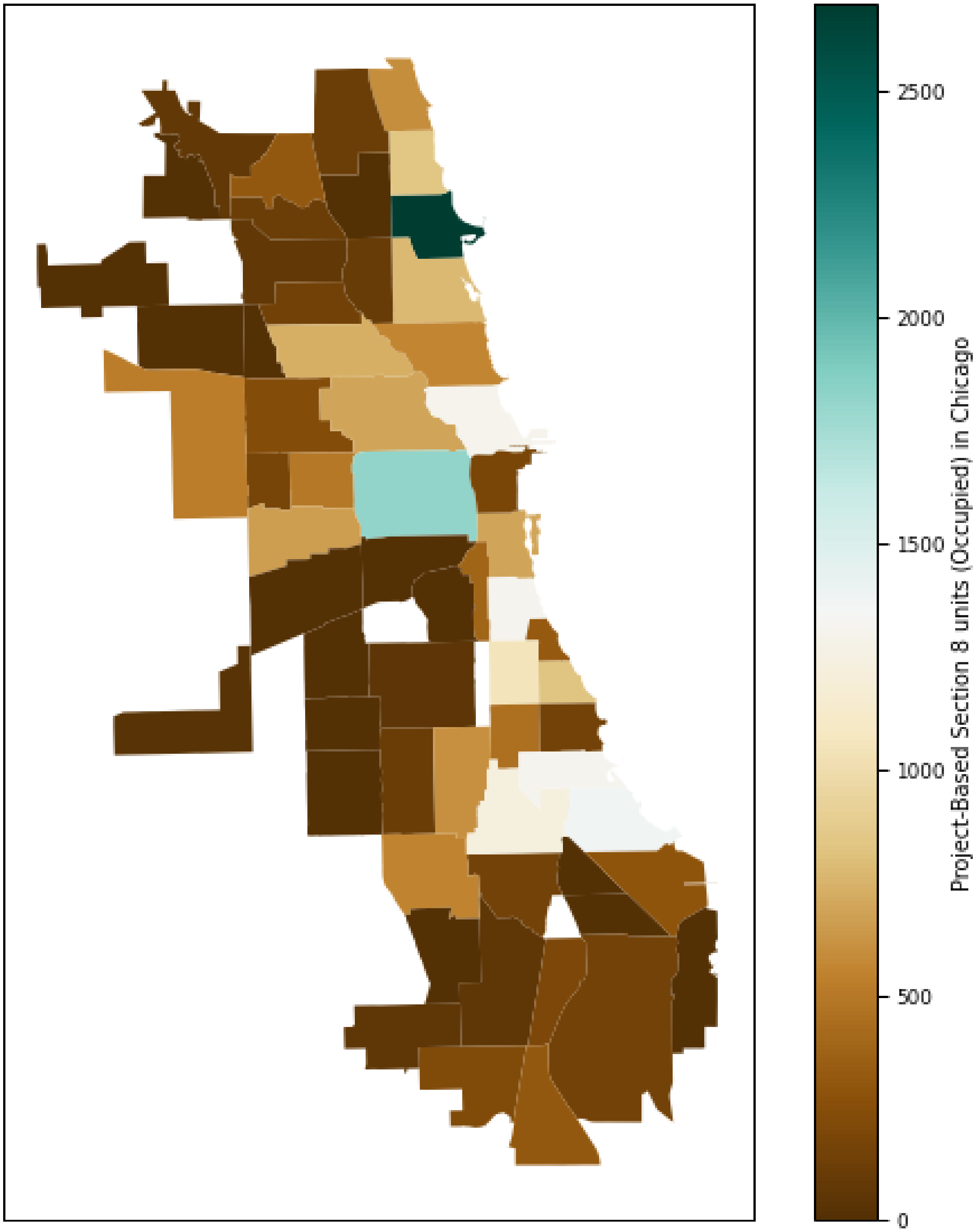
[35]:

```
                   Program label        Tract  Program  \
800   ="Project Based Section 8"  17031241100        5
835   ="Project Based Section 8"  17031242500        5
1187  ="Project Based Section 8"  17031390500        5
1251  ="Project Based Section 8"  17031420500        5
1981  ="Project Based Section 8"  17031832400        5

      Subsidized units available  % Occupied  % moved in past year  \
800                           11           0                     0
835                           12           0                     0
1187                          48          74                     3
1251                         291          71                     1
1981                          12           0                     0

      Number of people per unit  Number of people: total  CommunityAreaNumber  \
800                         0.0                        0                   24
835                         0.0                        0                   24
1187                        2.8                      108                   39
1251                        2.1                      376                   42
1981                        0.0                        0                   22

     CommunityAreaName  Occupied Units
```

```
800         West Town              0.0
835         West Town              0.0
1187          Kenwood             36.0
1251         Woodlawn            207.0
1981     Logan Square              0.0
```

```
[36]: # Merge the project-based-section 8 data with community map
      pbs8_num = pbs8.drop(columns=['Tract','Program','% moved in past year','%␣
       ↪Occupied',
                              'Number of people per unit']).
       ↪groupby('CommunityAreaNumber').sum()
      pbs8_w_map= com_map.merge(pbs8_num, on='CommunityAreaNumber')
```

```
[40]: sns.set_context('paper')
      ax2 = pbs8_w_map.plot('Occupied Units', legend=True,␣
       ↪figsize=(10,10),cmap='BrBG',
                            legend_kwds={'label':'Project-Based Section 8 units␣
       ↪(Occupied) in Chicago'})
      ax2.set_yticks([])
      ax2.set_xticks([])
      fig = plt.gcf()
      fig.savefig('/Users/bling/Documents/my_second_figure.pdf')
```


2500
2000
1500
1000
500
0
Project-Based Section 8 units (Occupied) in Chicago