The Art of the Deal

Identifying and Visualizing Differences Between Property Tax Levied and Services Rendered in Montgomery County, Maryland

Sofi Sinozich

2019-12-11

Contents

Motivation

NEED MORE HERE

CHECK WHETHER 1-YEAR OR 5-YEAR INCOME ESTIMATES

Montgomery County, Maryland is one of the richest counties in the United States (U.S. Census Bureau, n.d.), but much of that wealth is concentrated in the southwest of the county, with affluence generally decreasing aross the east-west axis (and to some degree, south-north). Correspondingly, neighborhoods in the southwest tend to pay disproportionately more in taxes (Lublin 2019), which are then redistributed across the county based on need.

But despite this redistribution, there remains a significant east-west divide in quality of life and services.

In theory, there is a happy medium - somewhere where the tax burden is most in-proportion to the government services and resources available, as well as other quality-of-life factors.

For this project, I focused primarily on property taxes, due primarily to data availability and interest. Income taxes and property taxes make up the bulk of county revenues, with income and property more or less in equal share (Montgomery County Department of Finance n.d.). While income taxes are typically confidential, many localities (including Montgomery County) publish databases of taxes levied on each individual property as public record. These data were filtered to include only those tagged as a primary residence to best represent property taxes of homeowners.

Another difference from income taxes is that property taxes have the additional wrinkle of not always being levied on the most current value of the home. In addition to typical assessments done at the homeowner's request for various purposes, the Maryland government re-assess properties on a three-year rolling schedule ("Real Property" n.d.). This allows for homes to appreciate in value - due to structural improvements or neighborhood trends - without a corresponding property tax increase for up to three years, which opens up more opportunities for discrepancies between property taxes and the general "value" of the area or home.

Data

Data were collected from a variety of sources to cover a wide range of factors that determine a neighborhood's value.

U.S. Census Bureau: * Shapefiles for Montgomery County tracts * Median home values (ACS) * Population (ACS)

Open Data DC: * WMATA Metrobus stops * WMATA Metro stations

Maryland iMAP: * Broadband internet information (up/down speeds, providers per census block) * MVA and VEIP facilities * Municipal, county and state police station locations * Municipal and county fire station locations * MARC train stops

dataMontgomery: * Local hospitals (Montgomery County and neighboring areas) * Montgomery County RideOn bus stops * Bikeways * Publicly reservable spaces (community facilities) * Early voting centers * Polling places * Property taxes

Montgomery County Public Schools (MCPS): * High school boundaries * Academic scores and demographic information by high school * Reported crimes

Most data were already in GeoJSON or similar format such that they could easily be aggregated by census tract.

```
bikeways <- st_read("https://data.montgomerycountymd.gov/api/geospatial/icc2-ppee?method=export&format=
  st_transform(crs=4269)
# Intersect with the tract information to get the total mileage within each tract
bikeways %<>% st_intersection(mocotracts,.) %>%
  mutate(distance = st_length(geometry)) %>%
  group_by(TRACTCE) %>% summarize(totaldistance = sum(distance)) %>%
  st_drop_geometry # Dropping geometry for simplicity
```

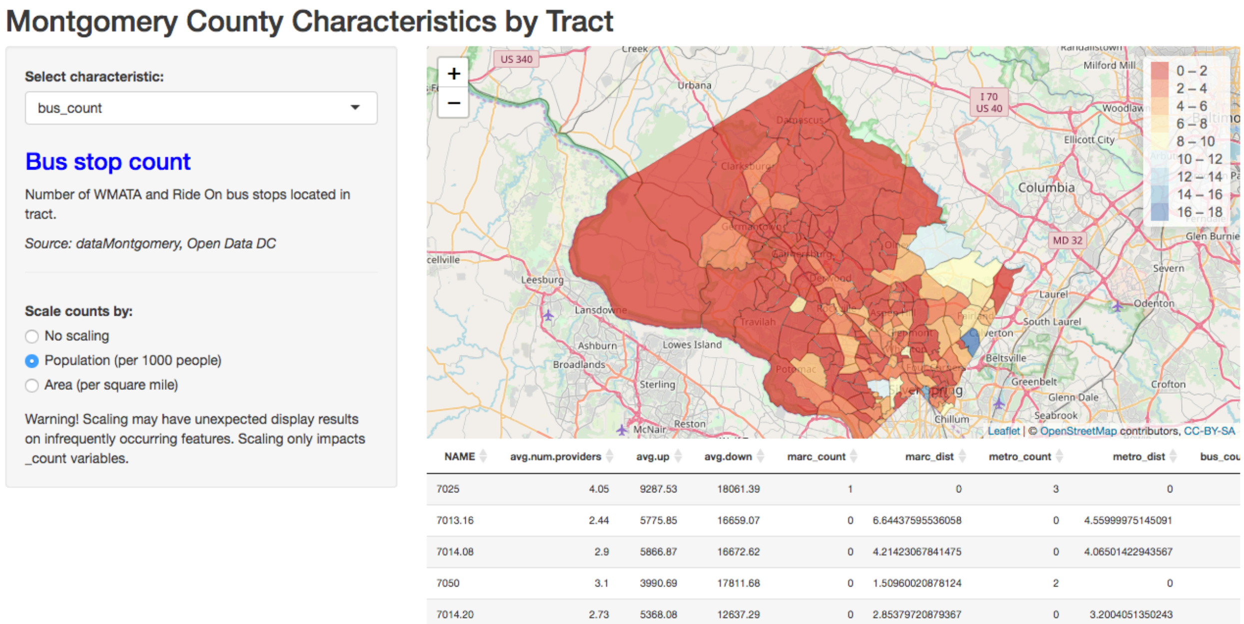
Others required more data processing. For example, MCPS does not (as of writing) provide school-level academic information in a machine-readable format, but does so in a large PDF report (Schools at a Glance). This required parsing the PDF and manually identifying the character locations of the desired fields.

```
saag_full <-
  pdf_text("https://www.montgomeryschoolsmd.org/departments/regulatoryaccountability/glance/currentyear/
# Isolate the high schools
saag_hs <- saag_full[370:421]
# Remove Thomas Edison since non-trad HS
saag_hs <- saag_hs[-c(13,14)]

# Very long regexes here will overrun the margins
hsp1<-
  str_match_all(string=saag_hs,pattern=regex("^ +([A-Za-z- \\.]+High School)[\\s\\S]+Enrollment = ([0-9
hsp2<-
  str_match_all(string=saag_hs,pattern=regex("Total\\nSchool\\s+([0-9]{3})\\s+([0-9]{3})\\s+([0-9]{3,4})
mcpshs <-
  hsp1[seq(1,length(saag_hs),2)] %>% sapply(.,function(x) x[2:7]) %>% t %>%
  cbind(hsp2[seq(2,length(saag_hs),2)] %>% sapply(.,function(x) x[2:5]) %>% t)
colnames(mcpshs)<-
  c("name","enrollment","esol","farms","gradrate","umdreq","satv","satm","satt","apib")
```

This munging eventually returns the desired data table. (Sample of two schools for demonstration purposes.)

Once collected, data were mapped and explored first ad-hoc and later using an interactive dashboard.

	name	enrollment	esol	farms	gradrate	umdreq	satv	satm	satt	apib
1	Bethesda-Chevy Chase High School	2,102	6.7	11.1	94.4	46.8	634	616	1249	73.5
2	Montgomery Blair High School	3,083	18.0	36.3	86.2	31.5	573	569	1142	53.3
3	James Hubert Blake High School	1,624	5.0	34.5	90.2	39.0	535	501	1036	41.8
4	Winston Churchill High School	2,123	5.0	5.0	95.0	71.5	624	628	1252	78.2
5	Clarksburg High School	2,148	8.0	26.8	91.5	42.3	548	541	1089	37.1
6	Damascus High School	1,271	5.0	14.6	94.0	45.2	574	564	1138	48.2



One major consideration was how to scale count data - by population or area. In most cases, the population scaling both made more sense conceptually and resulted in a more differentiated distribution. The interactive plot made it much easier to see visually what differences this would make per tract.

The "value score"

Each tract was assigned a summary "value score" based on four factors. These factors were determined largely by ad-hoc conceptual grouping rather than a rigorous method, but also with the goal of creating coherent subscores rather than grouping aspects that might be statistically similar but difficult to parse. Each subscore was rescaled from 0 to 100.

Subscore: Schools

Tracts were assigned a schools subscore on the basis of four academic indicators: 1. Total average SAT score (verbal + math) 2. % of graduates scoring here than a 3 (AP) or 4 (IB) 3. % of students meeting University of Maryland entrance requirements 4. Graduation rate

Since tract boundaries do not match school boundaries, school data was taken as the average of all of the schools who serve the tract. Scores scaled for the percentage of English for Speakers of Other Languages (ESOL) students and those eligible for free and reduced price school meals (FARMS) were also created as a way to benefit schools that work with more challenging populations.

```
mocodata %<>% mutate(score_school = ((satt/160 + apib + umdreq + gradrate)/4) %>% rescale(to=c(0,100)),
                     score_school_esol = (score_school/(1-esol/100)) %>% rescale(to=c(0,100)),
                     score_school_farms = (score_school/(1-farms/100)) %>% rescale(to=c(0,100)))
```

Subscore: Crime

Subscore: Services

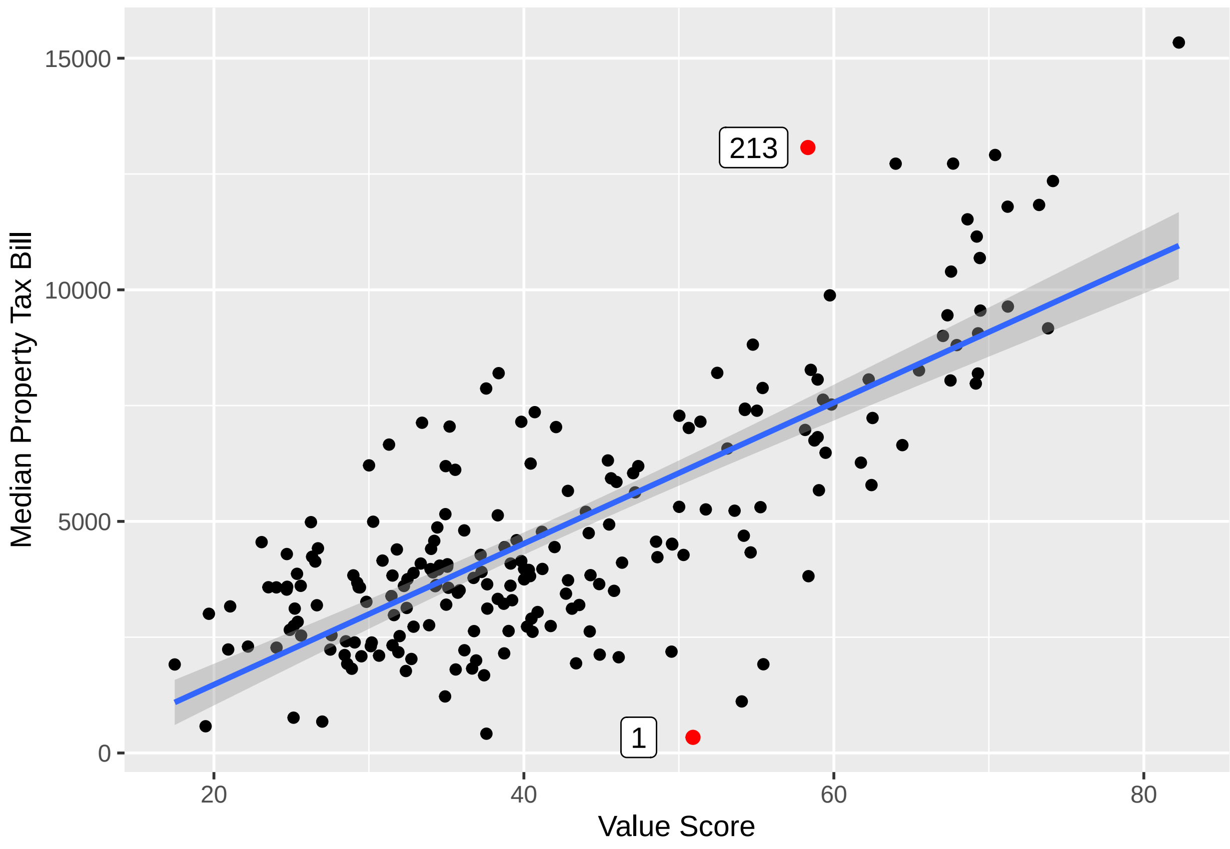
Subscore: Home values

Ranking the tracts on value

Simple linear regression of the median property tax bill for principal residences in the tract against on the value score of the tract (mean bill total was also tested and produced similar results).

(As previously noted, two tracts were not rated due to WHY)

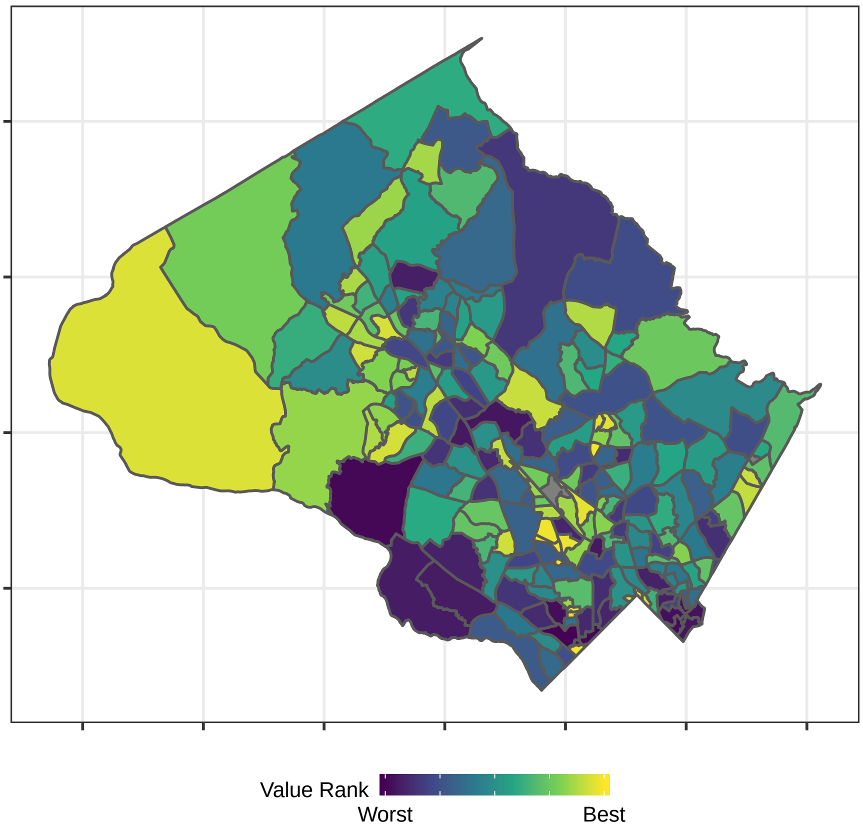
This was accomplished by ranking the tracts on the basis of the signed residuals - higher residuals indicate higher taxes levied than the value score would predict, while negative residuals indicate lower taxes levied than predicted by the score.

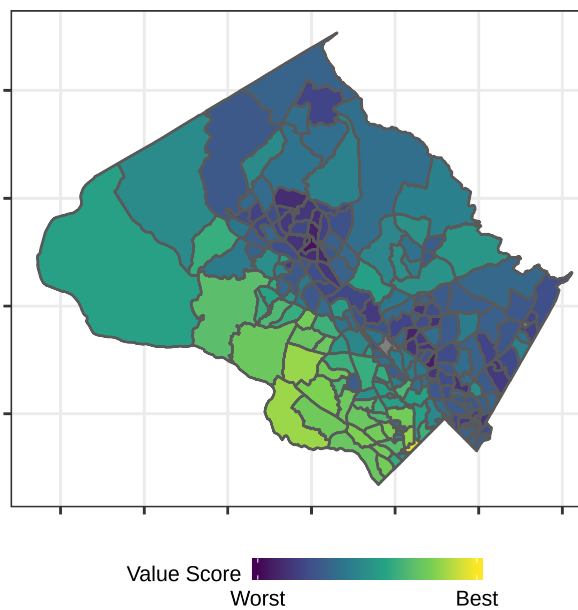



```
valrankplot <- mocodata %>% ggplot + geom_sf(aes(fill = valuerankrev)) +
  scale_fill_viridis_c("Value Rank",limits = c(1,213),breaks = seq(1,213,53), labels = c("Worst","","",
  theme_bw()+
  theme(legend.position = "bottom", legend.title = element_text(size = 8,vjust = .8),legend.text = eleme

valscoreplot <- mocodata %>% ggplot + geom_sf(aes(fill = valuescore)) +
```

```
  scale_fill_viridis_c("Value Score",limits = c(15,85),breaks = c(15,85),labels = c("Worst","Best")) +
  theme_bw()+
  theme(legend.position = "bottom", legend.title = element_text(size = 8,vjust = .8),legend.text = eleme

proptaxplot <- mocodata %>% ggplot + geom_sf(aes(fill = bill_total_median)) +
  scale_fill_viridis_c("Median Property Tax Bill",limits = c(300,15400),breaks = c(300,15400),labels = 
  theme_bw()+
  theme(legend.position = "bottom", legend.title = element_text(size=8,vjust = .8), legend.text = eleme
```

Adjusting to property tax levels severely penalizes the wealthiest (and highest-tax-paying) tracts in the southwest, as one can see from the raw value score and property tax distributions.





But aside from the western-most part of the southwest, there appear to be many opportunities for “good value” across the county, including just to the east, where the worst ranked tract (a sliver just above Friendship Heights) is just down the road from the best ranked tract (a similarly small area within downtown Bethesda). Though both of these tracts fall in the we

(While each tract technically has a name assigned by the Census Bureau, these names are not meaningful outside of being unique identifiers.)

Predicting the value score

(DEMOS STUFF)

```
# What happens here depends on the specific project
```

```
# What happens here depends on the specific project
```

Analysis

This section presents the main results, such as (for example) stats and graphs that show relationships, model results and/or clustering, PCA, etc.

```
# What happens here depends on the specific project
```

```
# What happens here depends on the specific project
```

```
# What happens here depends on the specific project
```

Discussion

This section summarizes the results and may briefly outline advantages and limitations of the work presented.

Tracts do not correspond to neighborhood boundaries

The notion of value is of course quite opinionated, thus the addition of the adjustable subscore weights. But even the data collection process required subjective decisionmaking. For many, the availability of public transport, for example, is not a major consideration in where they decide to live, as they are perfectly happy walking, biking, or driving to any place they need to go.

Most difficult are intangibles - aspects like how clean a neighborhood is or how friendly the neighbors are. These community aspects may be an important part of why people choose to live where they do, but are not often measured in any systematic way.

Future Directions

This project could easily be extended in a variety of ways.

1. Improve value scoring system The value scoring system was primarily based on items that were of personal interest as well as were readily available from free public datasources. It has not been validated theoretically or empirically. There are undoubtably many more variables that could be added to the scoring model (and some that could be removed) and the base weights could be adjusted to better match conventional norms about the value of a neighborhood. More subjective assessments could also be incorporated into the score, though these would likely need to be sourced from a purpose-built survey or other data-collection process, as data of this type is not frequently available at the tract level.
2. Create a more full-featured dashboard The current interactive provides a few different options and some level of user-customizability, but could easily incorporate, for example, more mapping layers (presenting the source data, for example) and the ability to locate a specific address. Aesthetic improvements could also aid usability, particularly for those less familiar with computer interfaces.

References

Lublin, David. 2019. "Yes to Racial Equity and Social Justice. No on This Legislation." *The Seventh State*. http://www.theseventhstate.com/?p=12208.

Montgomery County Department of Finance. n.d. "December 2018 Revenue Update and Selected Economic Indicators." Accessed December 10, 2019. https://www.montgomerycountymd.gov/Finance/Resources/Files/data/economic/CCBriefing121118.pdf.

"Real Property." n.d. Accessed December 10, 2019. https://dat.mayland.gov/realproperty/Pages/default.aspx.

U.S. Census Bureau. n.d. "Income in the Past 12 Months (in 2018 Inflation-Adjusted Dollars)." *2018 American Community Survey (1-Year Estimates)*.