**CANVAS vs. DNHS.**

As a reminder, CANVAS collected disorder measures on 501 block faces spread evenly around the city. DNHS audited about 4,000 blocks, clustered in 135 block groups.

The DNHS disorder measure, also termed ‘building quality’ was computed for each audited block group by counting the proportion of blocks within the block group for which each of three indicators of disorder was present, then performing a PCA on these 135 proportion measures and selecting the first component. The score for each block group was the linear combination of observed variables multiplied by component scores

**DNHS PCA results**

|  |  |
| --- | --- |
| Item | Component scores for component 1 |
| Broken/boarded up windows | .60 |
| Major repairs needed | .58 |
| Abandoned buildings | .56 |

% of variance explained by component 1: 86.2%

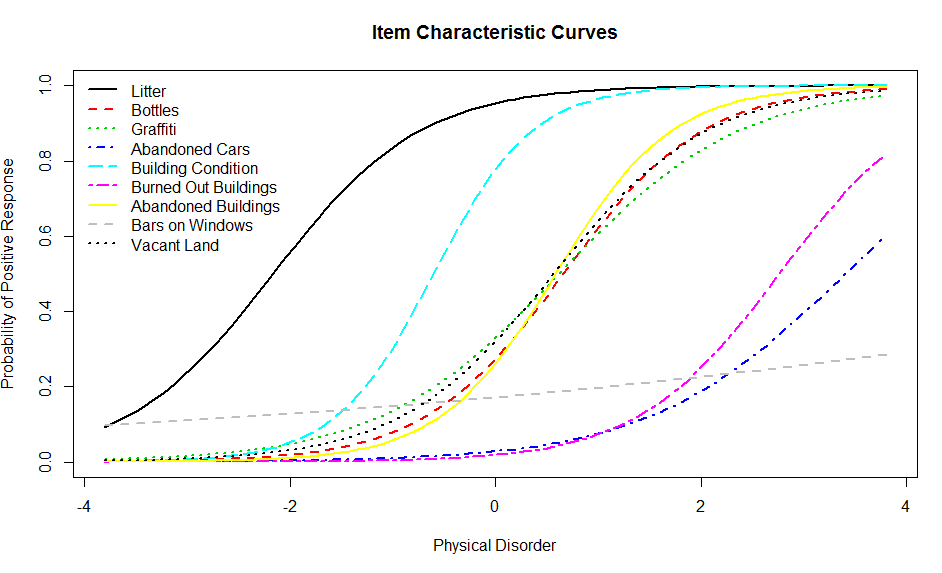
I tried to mimic this PCA approach with 3 roughly comparable items from CANVAS (I also compared to the gold-standard CANVAS measure (a 9-item IRT scale)):

**PCA of 3 similar items in CANVAS data**

|  |  |
| --- | --- |
| Item | Component scores for component 1 |
| Burned out buildings | .40 |
| Major repairs needed | .64 |
| Abandoned buildings | .65 |

% of variance explained by component 1: 53.9%

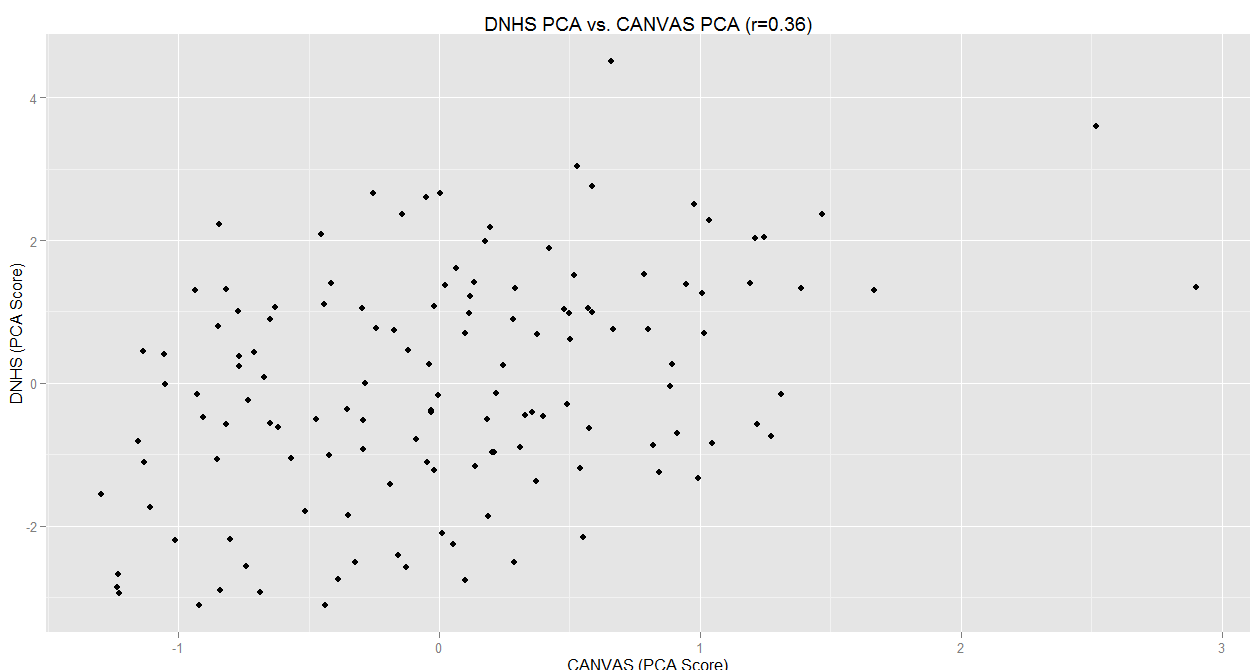
CANVAS’s less complete fit from component 1 may be a function of less variation in item scores – the CANVAS PCA was for 501 block faces, each with a score of 1 or 0 for each indicator. By comparison, DNHS had about 30 blocks within each block group, and so had a much wider range of scores

**CANVAS IRT**

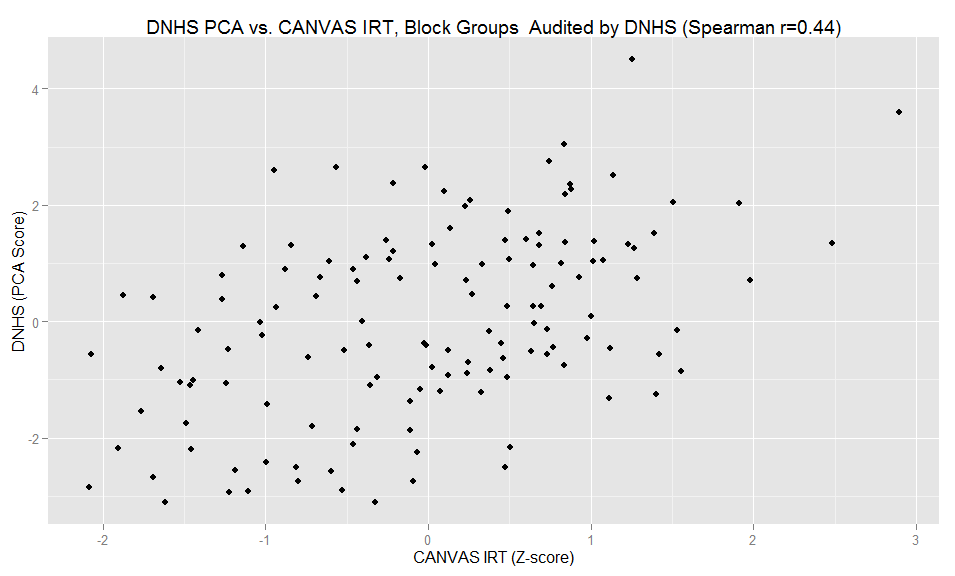
As a reminder, the IRT model we fit to CANVAS’s items was consistent with a uni-dimensional scale.

Not surprisingly, given the uni-dimensionality, more items led to more consistent estimation: for the IRT model, the reliability, computed as 1-(1/{Total Information}) was **0.92**.

**Predicted CANVAS PCA score at centroids of DNHS-audited block groups as compared to DNHS PCA score for the block group. (Used ordinary kriging for CANVAS PCA)**

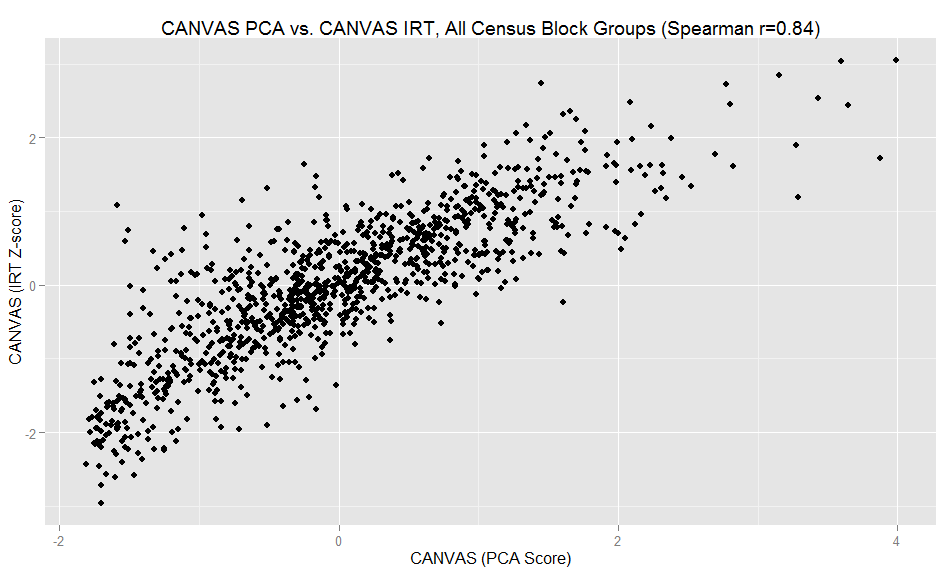


**DNHS PCA vs. CANVAS IRT Model, DNHS-audited block groups**



So even though the PCA measure was more directly comparable, the IRT measure had a better correlation with the DNHS measure, probably because the additional indicators add information lost in the PCA alone.

**But just to be clear, CANVAS IRT and PCA are much more highly correlated with each other than either is with DNHS:**



**So, how well do these measures predict?**

**Internal validation: Kriging Cross-validation**

Jackknife cross-validation (aka leave-one-out cross-validation) error estimates are computed by subtracting the observed estimate from the value estimated by kriging for that location from all other sampled locations. Lower numbers for RMSE and MAD are better

|  |  |  |  |
| --- | --- | --- | --- |
| Model | Number of x-validated locations | Root Mean Square Error | Median Absolute Deviation |
| DNHS | 135 | 1.54 | 1.09 |
| CANVAS (3-component PCA) | 501 | 1.06 | 0.56 |
| CANVAS (9-item IRT) | 501 | 0.72 | 0.49 |
| CANVAS PCA x-validating DNHS | 135 | 1.42 | 1.13 |

Conclusion: the CANVAS measure appears to be predicting at least as well as DNHS.

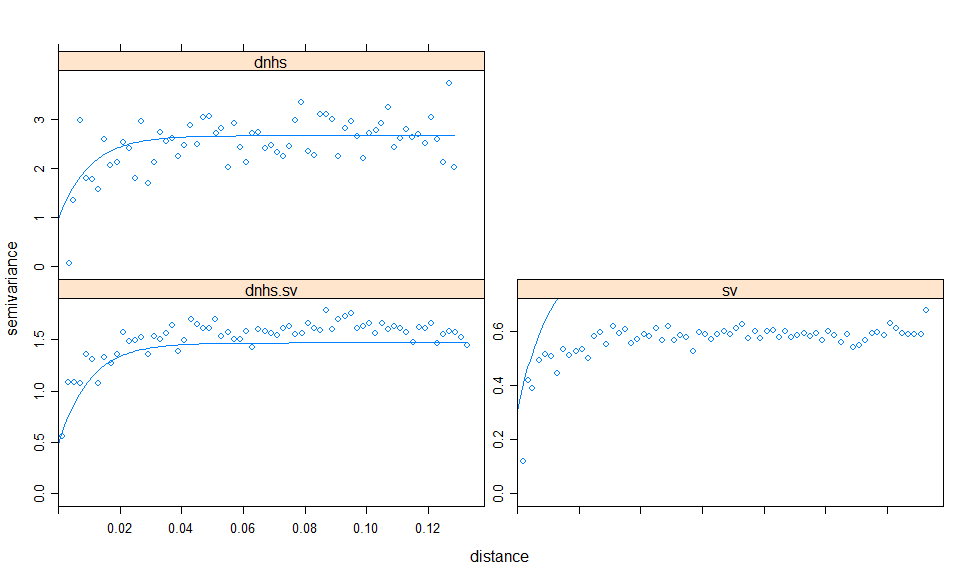
**Using CANVAS and DNHS: Co-kriging notes**

Mike suggested we investigate how much predictive improvement we might get by co-kriging, which essentially allows us to augment the predictions from a less complete sample with predictions from a more complete sample of something else. Here, we consider co-kriging the DNHS estimates using the Streetview IRT model as a co-variable.

A note on co-kriging vs. universal kriging/kriging with external drift

Co-kriging is appropriate to incorporate information from co-variables that can be spatially modeled and for which the value of the co-variable at the target location is unknown. (As in this case). By contrast, if we wanted to augment predictions of disorder using major/minor street status, borough, etc., then universal kriging would be more appropriate. Note that we could incorporate both spatially modeled co-variables and known predictors with universal co-kriging.

**Direct and cross-variograms with fitted functions:**



The fitted function looks good for the cross-variogram and the dnhs direct variogram, but off for the sv direct variogram. I believe this is a function of the requirement that the variogram suite never predict negative variance. But I don’t really understand it at this point – looking to Mike for help.

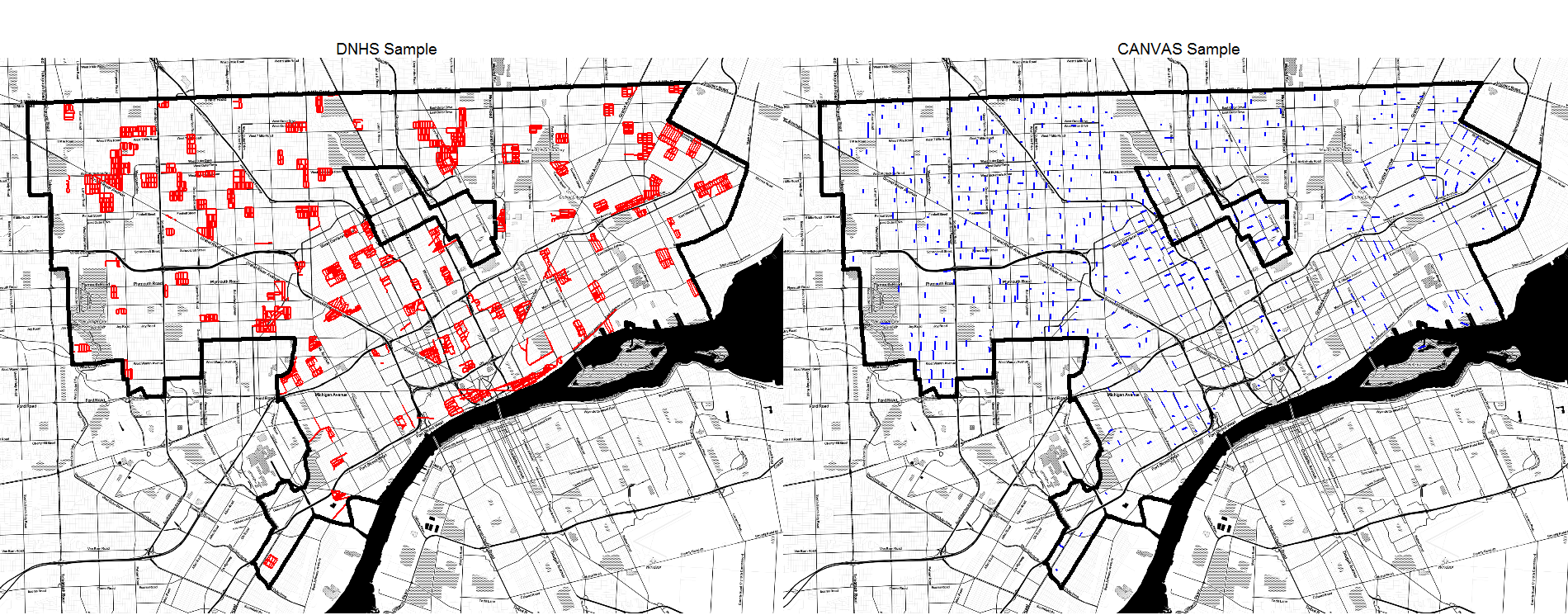
**Co-kriging Cross-validation**

|  |  |  |  |
| --- | --- | --- | --- |
| Model | Number of x-validated locations | Root Mean Square Error | Median Absolute Deviation |
| DNHS | 135 | 1.54 | 1.09 |
| CANVAS (3-component PCA) | 501 | 1.06 | 0.56 |
| CANVAS (9-item IRT) | 501 | 0.72 | 0.49 |
| CANVAS PCA x-validating DNHS | 135 | 1.42 | 1.13 |
| Co-Kriging DNHS with CANVAS IRT | 135 | 1.40 | 1.13 |
| Co-Kriging CANVAS IRT with DNHS | 501 | 0.71 | 0.47 |

Conclusion: either a) co-kriging does not substantially improve prediction or b) I did it wrong.

NOTES FROM PREVIOUS DISCUSSIONS OF DNHS AND CANVAS FOLLOW JUST FOR REFERENCE

**CANVAS and DNHS Samples**



Note that the CANVAS sample did not exclude Hamtramck and Highland Park, the ‘donut holes’ inside Detroit, whereas DNHS did.

Generally, DNHS assessed about 10x the number of blocks, but all sampled blocks are more geographically clustered. Other differences include:

1. DNHS audits were in 2008, whereas most CANVAS imagery (98% of segments) was from summer 2009
2. CANVAS assessed one side of the street whereas DNHS assessed both (I believe).

**Inter-rater reliability**

In spite of the different protocols, I wanted to take a look at inter-rater reliability. So assuming any pair of segments with end points within 20 meters matched, there were 52 street segments audited by both surveys. I ran some inter-rater reliability stats on the items included in both audits, which resulted in the following:

|  |  |
| --- | --- |
| Item assesses | Kappa |
| Boarded up or abandoned buildings | 0.43 |
| Repairs needed | 0.16\* |
| Vacant lots | 0.31 |
| Non-art graffiti | 0.02\*\* |
| Clean street | 0.08 |
| Abandoned cars | -0.06 |

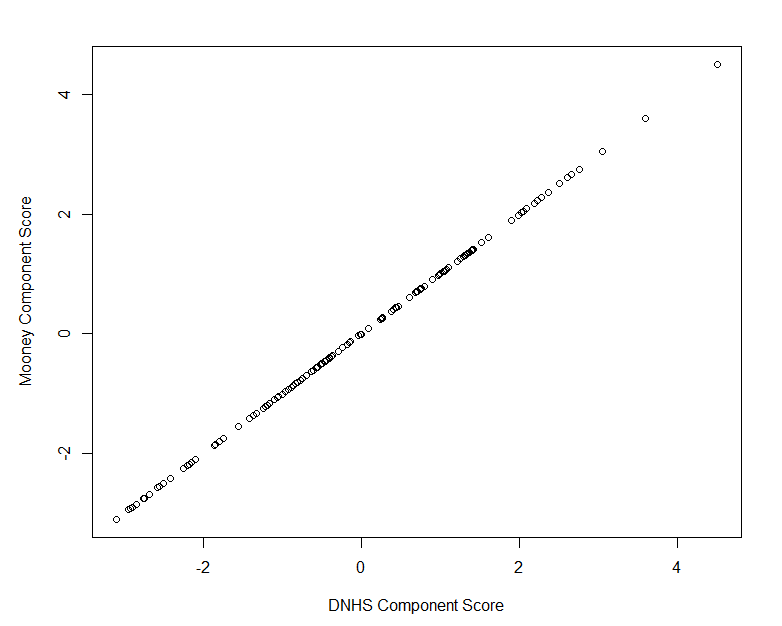
\* DNHS apparently had harsher criteria: P(DNHS|SV=yes)=1

\*\* SV included graffiti painted over

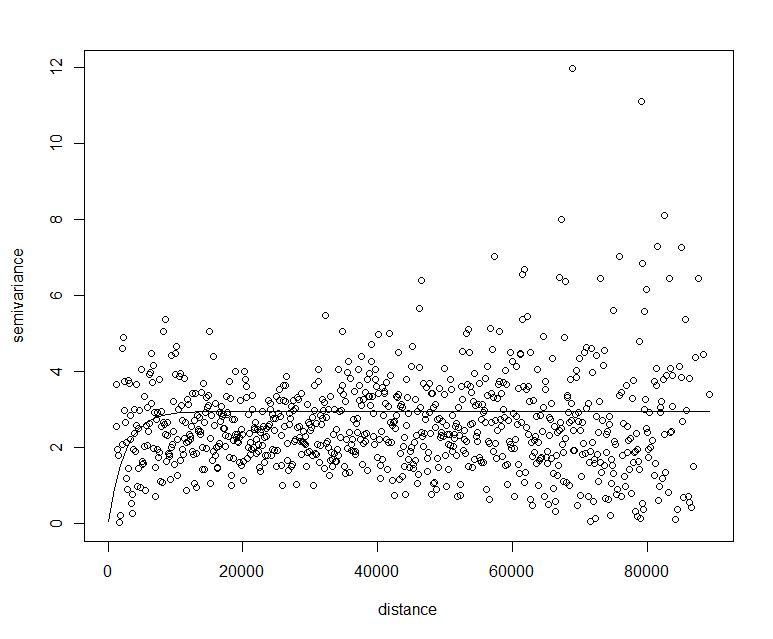
Note: SV audited only one side of the street, and a year elapsed between audited, so kappa may not be strictly appropriate.

**Kriged estimates**

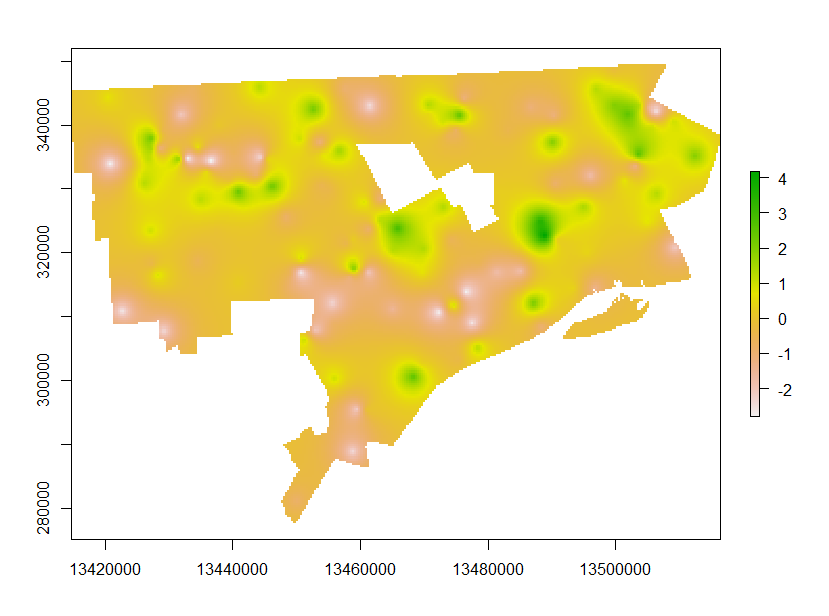
To get neighborhood estimates, DNHS estimated a neighborhood-level value from each cluster of audited segments, then performed a PCA to identify a component score for each neighborhood, then kriged that. I was able to replicate their PCA:



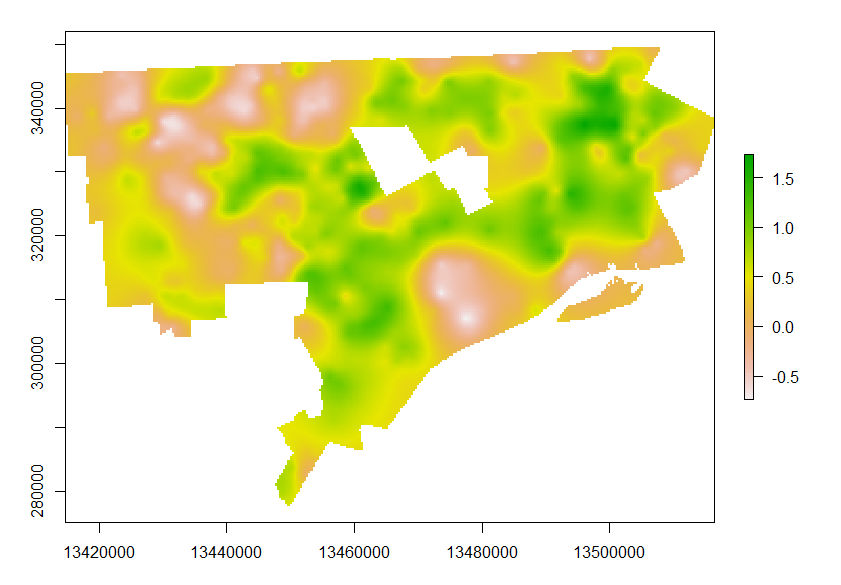
But I had to play with the bin size to get a variogram that showed much in the way of spatial autocorrelation

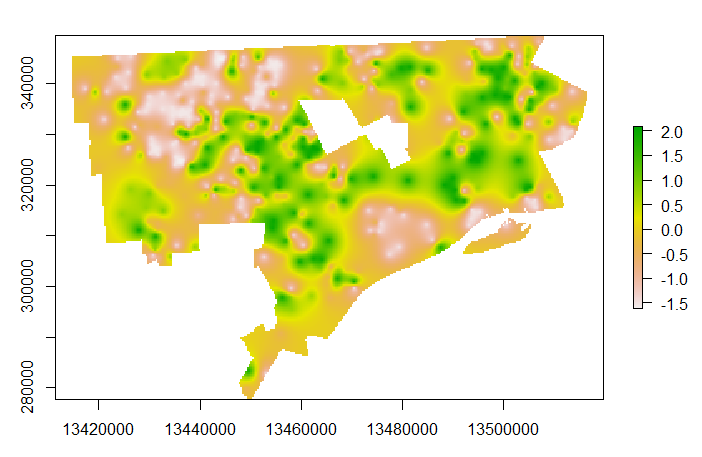


Using this, the DNHS view of disorder in Detroit looks like

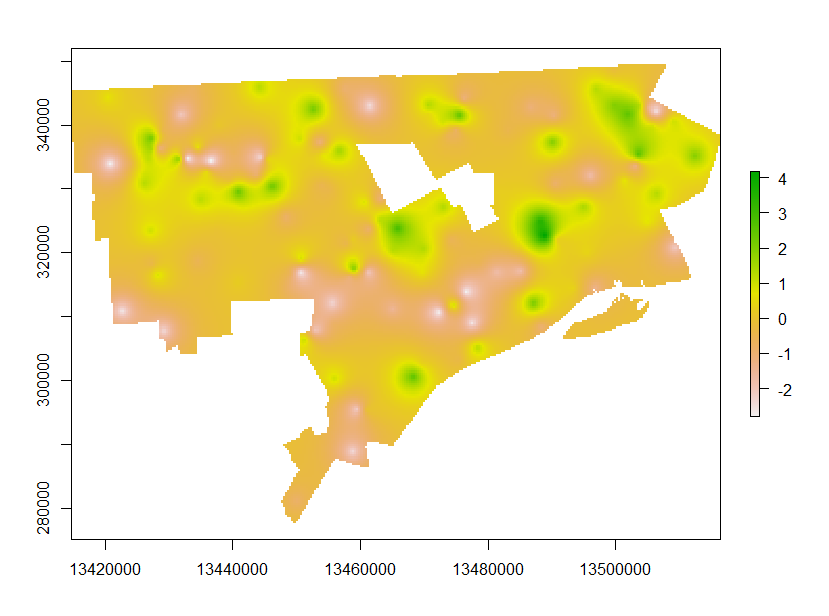


Whereas CANVAS looks like this:

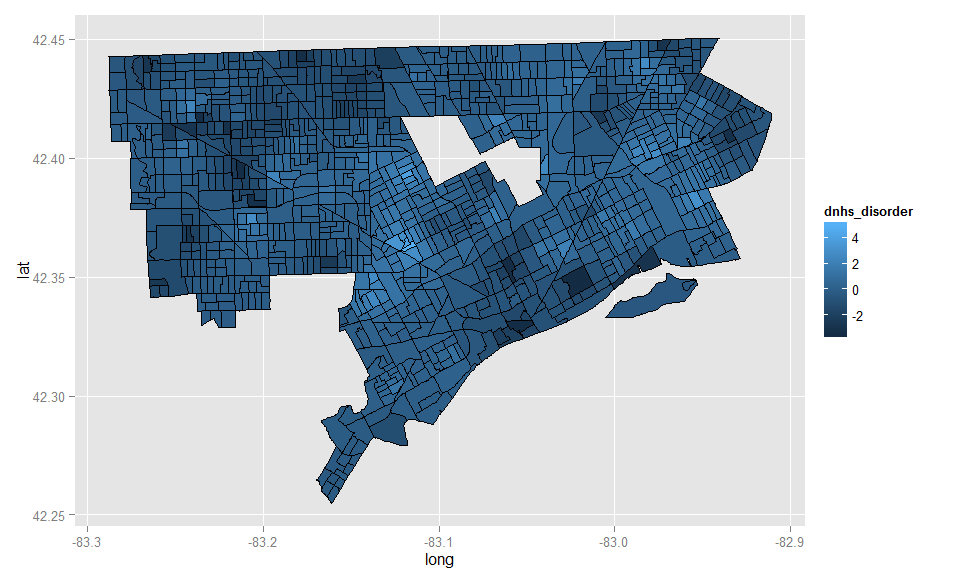
 Using a PCA of comparable measures, CANVAS’s building quality map would look like:



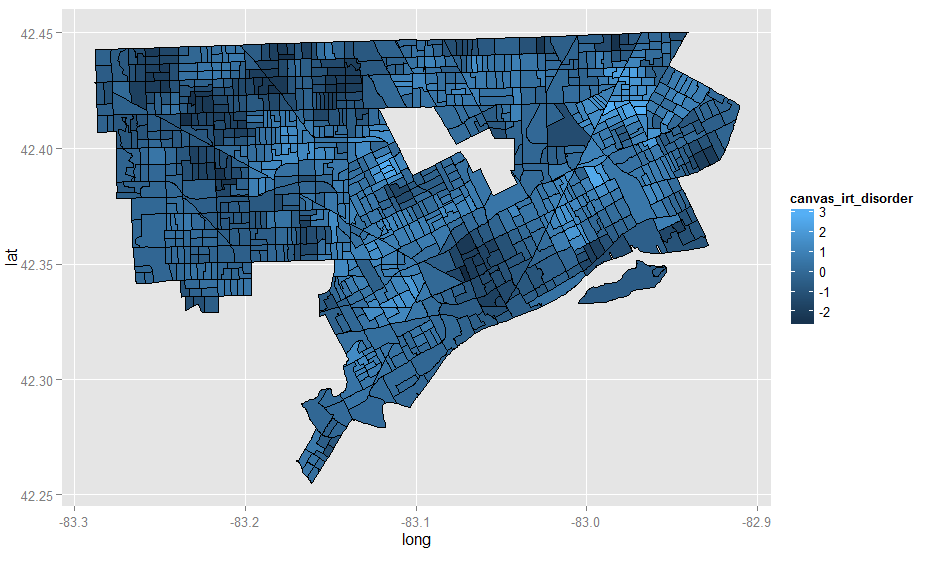
DNHS again:



DNHS BLOCK GROUP MAP



CANVAS IRT MAP



CANVAS PCA MAP

