Unhappy Metros: Panel Evidence

Adam Okulicz-Kozaryn*

Rutgers - Camden

Saturday 13th November, 2021   14:46

We study the effect of living in a metropolitan area on life satisfaction (subjective wellbeing). The literature agrees that inhabitants of metropolitan areas tend to be less satisfied with their lives than inhabitants of smaller settlements in the developed world. This is the first study using longitudinal dataset to test the unhappy metro hypothesis. Using the 2009-2019 US Panel Study of Income Dynamics (PSID), we find support for the corss-sectional findings: metros are less happy than nonmetros.

Urbanism is not just a built environment, but a way of life (Wirth 1938). Urbanism affects humans in multiple and profound ways, indeed urbanism is arguably the most significant disruption of human habitat in our species history (Okulicz-Kozaryn 2015). World is urbanizing at an astonishing pace–urban population will explode from .75b in 1950 in to 6.75b in 2050 (`https://population.un.org/wup`)–6 billion people put in urban areas over just 100 years.

At the same time an agreement has emerged that aside from, or even instead of, traditional development measures such as GDP and HDI, it is useful to measure human development with Subjective Wellbeing measures (Stiglitz et al. 2009, Diener 2009).

The urban-rural happiness gradient states that happiness raises from its lowest in largest cities to highest in smallest places, little towns, villages, and open country. The evidence of urban-rural happiness gradient is overwhelming–urban unhappiness is common in the developed world (Okulicz-Kozaryn and Valente 2021).

Yet all studies to date are cross sectional and panel evidence is missing. Few studies that use panel data do not actually test the urban unhappiness hypothesis.

White et al. (2013b) and White et al. (2013a) use British panel (BHPS) but test green space (such as gardens, parks, and proximity to coast), not size of a place. Similarly, Alcock et al. (2014) is a panel (BHPS) but also examining green space, not size of a place.

Hoogerbrugge and Burger (????)

using BHPS examines size of a place, but rather than testing urban unhappiness, again, tests open green space hypothesis. The size of a place cutoff is at 10,000 people or even 3,000 people for Scottland. Hence, much of the places above the cutoff, such as large villages and small towns are not really "urban." They are lacking defining features of urbanness: size, density, and heterogeneity Wirth (1938). The build envirnment in villages or small towns lacks tall bildings, urban transit, airports, etc. Way of life in such places is not urban, lacking shaloowness transitriness, etc The way of life lacks transitoriness, etc these words from city book.

cite when metropolis is too big and urgan unhappiness is common–ideally it should be a gradient, and if necessarily a binary disctinction then it is several hundred thousand, not 3 or 10 thoudsand as in Hoogerbrugge and Burger (????).

Rehdanz and Maddison (2008) uses a German panel dataset (GSOEP), properly defining urban rural happiness gradient with multiple cutoffs including at several hundred thousand, but without panel modelling techniques such as fixed or random effects.

1 Data and model

We use 2009-2019 psid from `psidonline.isr.umich.edu`, the reason being that swb question only started in 2009. from the family files we only retain household reference person (or head as it used to be called). This is the same practice as in (Brown and Gathergood 2019).

the swb question reads: Please think about your life as a whole. How satisfied are you with it? Are you completely satisfied, very satisfied, somewhat satisfied, not very satisfied, or not at all satisfied? on scale from 1 (low) to 5 (high).

and the key independent variable is metro as defined in table 1.

metro	beale rural-urban code	description
1	1	Metro: Counties in metro areas of 1 million population or more
1	2	Metro: Counties in metro areas of 250,000 to 1 million population
1	3	Metro: Counties in metro areas of fewer than 250,000 population
0	4	Nonmetro: Urban population of 20,000 or more, adjacent to a metro area
0	5	Nonmetro: Urban population of 20,000 or more, not adjacent to a metro area
0	6	Nonmetro: Urban population of 2,500 to 19,999, adjacent to a metro area
0	7	Nonmetro: Urban population of 2,500 to 19,999, not adjacent to a metro area
0	8	Nonmetro: Completely rural or less than 2,500 urban population, adjacent to a metro area
0	9	Nonmetro: Completely rural or less than 2,500 urban population, not adjacent to a metro area

Table 1: metro variable: Metropolitan/Non-metropolitan Indicator: This indicator is derived from the 2013 Beale-Ross Rural-Urban Continuum Codes published by USDA based on matches to the FIPS state and county codes: 1. Metropolitan area (Beale-Ross Code ER775923= 1-3); 0. Non-metropolitan area (Beale-Ross Code ER775923= 4-9). Each county in the U.S. is assigned one of the 9 codes.

Summary statistics are shown in Supplementary Online Material (SOM).

There are 3 vars that not only predict SWB, but also are likely to be confounded with metro: race, political views, and religiosity–yet, they are irrelevant in fixed effects model as they are mostly constant over short period of time such as that considered here.

panel structure description of metro variable is shown in the appendix yes one problem is that tehre is little variability in urb-rur

controls are set in table **??** var_des in SOM

in controls we follow Brown and Gathergood (2019)

it is impostant to control for employment status–it both predicts swb, especially unmeployment has lasting negative effect on swb, and it also correlates with urbanicity as cities have more employment opportunities (e.g., O'Sullivan 2009).

boilerplate on linear models, no need for categorical dependent variable modeling, elaborate

a standard fixed effects model is given by:

$$SWB_{it} = \gamma METRO_{it} X_{it}\beta + \alpha_i + u_{it} \tag{1}$$

Where, $METRO_{it}$ is a metro dummy for person i at time t. γ is the main coefficient of interest. α_i (i=1...n) is the unknown intercept for each person(n person-specific intercepts). SWB is the dependent variable, where i = person and t = wave. X_{it} is a vector of control variables. β is the vector of coefficients for control variables. u_{it} is the error term.

A limitation of fixed effects model is that there are just 6 waves and tehre is limited variablility across time and hence time-invariant variables cannot be estimated. metro which is not that often changed does change but not very often—discuss that des sta from som.

2 Resuklts

by wave each cross section just have in app and refer briefly, and re; in body only fe

Table 2: FE regressions of SWB.

	a1	a2	a3	a4	a5
metro	0.01	-0.04*	-0.03*	-0.04**	-0.04*
age		0.02***	0.02***	0.01***	0.00
age sq		-0.00**	-0.00	-0.00	-0.00
last year total family income		0.00*	0.00	0.00	0.00
unemployed		-0.18***	-0.18***	-0.16***	-0.16***
male		0.27	0.21	0.07	0.08
health		0.13***	0.13***	0.10***	0.10***
kids			-0.01	-0.01	-0.01
college			-0.08*	-0.07	-0.07
married			0.18***	0.17***	0.17***
family unit size			0.04***	0.03***	0.03***
distress				-0.05***	-0.05***
constant	3.71***	2.37***	2.45***	2.90***	3.60***
state and year dummies	no	no	no	no	yes
N	37567	37489	36285	36142	36142

*** p<0.01, ** p<0.05, * p<0.1

nice size effect like a third or half of step in 1-5 hea, like going third or half way from poor health to fair health for instance

3 Conclusion and discussion

This is the first panel data investigation of metro-nonmetro happiness gap. The results confirm crosssectional evidence of urban unhappiness.

Future research can improve in a number of ways. Use finer classfication than binary metro-nonmetro. As more waves become available, it will be possible to estimate SWB from moving across urbanicity.

Regressions coefficients on `metro` are not significant without controling for predictors of SWB, so it is important to adjust the metro non-metro happiness gap with happiness predictors, unlike in Burger et al. (2020).[1]

Supplementary Online Material (SOM)

[note: this section will NOT be a part of the final version of the manuscript, but will be available online instead]

---

[1] Burger et al. (2020) also uses faulty Gallup data as elaborated in Okulicz-Kozaryn and Valente (2021)–in general, one should steer away from Gallup happiness data–Gallup charges $30,000 for access (per one year), clearly "happiness industry", not happiness research Davies (2015).

3.1 Variables' Definitions

<p align="center">**Table 3:** Variable definitions.</p>

name	description
swb	"Please think about your life as a whole. How satisfied are you with it? Are you completely satisfied, very satisfied, somewhat satisfied, not very satisfied, or not at all satisfied?" 1 (lo) - 5 (hi)
metro	"Metropolitan/Non-metropolitan Indicator. This indicator is derived from the 2013 Beale-Ross Rural-Urban Continuum Codes published by USDA based on matches to the FIPS state and county codes." 1 Metropolitan area (Beale-Ross Code ER775923= 1-3) 0 Non-metropolitan area (Beale-Ross Code ER775923= 4-9)
age	age
age sq	age squared
last year total family income	last year total family income
unemployed	EMPLOYMENT STATUS-1ST MENTION; We would like to know about what you do – are you working now, looking for work, retired, keeping house, a student, or what?–FIRST MENTION; 1="Looking for work, unemployed", 0 otherwhise
male	gender
health	"Now I have a few questions about your health. Would you say your health in general is excellent, very good, good, fair, or poor?" 1 (poor) to 5 (excellent)
kids	"Number of Persons Now in the FU Under 18 Years of Age"
college	"Did (you/he/she) attend college?" 1='yes', 0='no'
married	"Are you married, widowed, divorced, separated, or have you never been married?" 1='married'; 0 otherwhise
family unit size	Number of Persons in FU at the Time of the Interview
white	"What is (your/his/her) race? (Are you/Is [he/she]) white, black, American Indian, Alaska Native, Asian, Native Hawaiian or other Pacific Islander?–FIRST MENTION" 1='white', 0 otherwhise
distress	The K-6 Non-Specific Psychological Distress Scale

3.2 Summary statistics

```
        id:  2, 3, ..., 14365                              n =        10108
        yr:  2009, 2011, ..., 2019                         T =            6
             Delta(yr) = 1 unit
             Span(yr)  = 11 periods
             (id*yr uniquely identifies each observation)

Distribution of T_i:   min      5%     25%       50%      75%      95%      max
                         1       1       2         4        6        6        6

     Freq.  Percent    Cum. |  Pattern*
  -------------------------+----------
     3179    31.45   31.45 |  111111
      723     7.15   38.60 |  ....11
      672     6.65   45.25 |  .....1
      548     5.42   50.67 |  ...111
      505     5.00   55.67 |  ..1111
      502     4.97   60.64 |  1.....
      481     4.76   65.39 |  .11111
      480     4.75   70.14 |  111...
      450     4.45   74.59 |  11....
     2568    25.41  100.00 |  (other patterns)
  -------------------------+----------
    10108   100.00         |  XXXXXX
  ----------------------------------------
*Each column represents 2 periods.
```

```
Variable         |      Mean   Std. Dev.      Min      Max |    Observations
-----------------+-------------------------------------------+----------------
swb     overall  |  3.718723   .8759134        1        5 |   N =     37767
        between  |             .7136601        1        5 |   n =     10091
```

```
          within |              .5603667    .5187227    6.552056 | T-bar = 3.74264
                 |                                              |
met       overall |  .7801749    .4141335           0          1 |     N =    37730
          between |              .385049            0          1 |     n =    10073
          within |              .1878218   -.0531584    1.613508 | T-bar = 3.74566
                 |                                              |
age       overall |  44.85923    16.82858          16         99 |     N =    37928
          between |              17.23457           17         99 |     n =    10107
          within |              2.911229    37.19257    51.60923 | T-bar = 3.75265
                 |                                              |
age2      overall |  2295.544    1698.311          256       9801 |     N =    37928
          between |              1728.178          289       9801 |     n =    10107
          within |              285.6776    1268.044    3368.044 | T-bar = 3.75265
                 |                                              |
inc       overall |  61242.84    81095.25           0    3316000 |     N =    37912
          between |              66126.72           0    1883797 |     n =    10108
          within |              39658.52   -937554.6    2052160 | T-bar = 3.75069
                 |                                              |
une       overall |  .0907629    .2872754           0          1 |     N =    37923
          between |              .2197753           0          1 |     n =    10108
          within |              .2116276   -.7425705    .9240962 | T-bar = 3.75178
                 |                                              |
male      overall |  .5491432    .4975856           0          1 |     N =    37930
          between |              .4994373           0          1 |     n =    10108
          within |              .0058544   -.2508568    1.049143 | T-bar = 3.75247
                 |                                              |
hea       overall |  3.436374    1.046857           1          5 |     N =    37862
          between |              .92175             1          5 |     n =    10100
          within |              .5731675    .4363742    6.603041 | T-bar = 3.74871
                 |                                              |
kid       overall |  .6846296    1.119852           0         11 |     N =    37930
          between |              1.061457           0         11 |     n =    10108
          within |              .4720193   -3.148704    5.18463 | T-bar = 3.75247
                 |                                              |
col       overall |  .6264205    .4837605           0          1 |     N =    36608
          between |              .4803487           0          1 |     n =     9674
          within |              .0744979   -.2069129    1.459754 | T-bar = 3.78416
                 |                                              |
mar       overall |  .2802942    .4491489           0          1 |     N =    37928
          between |              .412416            0          1 |     n =    10107
          within |              .1586708   -.5530391    1.113628 | T-bar = 3.75265
                 |                                              |
nFU       overall |  2.281413    1.412387           1         14 |     N =    37930
          between |              1.317328           1         13 |     n =    10108
          within |              .615575    -3.468587    8.081413 | T-bar = 3.75247
                 |                                              |
whi       overall |  .5256386    .4993489           0          1 |     N =    37697
          between |              .4985538           0          1 |     n =    10038
          within |              .0213918   -.2743614    1.192305 | T-bar = 3.75543
                 |                                              |
k         overall |  3.608904    4.151942           0         24 |     N =    37689
          between |              3.629813           0         24 |     n =    10083
          within |              2.339743   -10.05776    22.6089 | T-bar = 3.73788

(obs=5.00  ,55  8.00  )
```

	swb	met	age	age2	inc	une	male	hea	kid	col	mar	nFU	whi
swb	1.00												
met	-0.08	1.00											
age	0.09	-0.05	1.00										
age2	0.09	-0.06	0.98	1.00									
inc	0.13	0.06	0.10	0.06	1.00								
une	-0.12	0.02	-0.19	-0.18	-0.13	1.00							
male	0.09	-0.06	-0.01	-0.03	0.29	-0.02	1.00						
hea	0.27	0.02	-0.24	-0.23	0.20	-0.01	0.15	1.00					
kid	-0.01	0.02	-0.29	-0.30	0.01	0.09	-0.10	0.05	1.00				
col	0.04	0.08	-0.08	-0.09	0.24	-0.12	0.06	0.17	-0.04	1.00			
mar	0.20	-0.06	0.17	0.14	0.43	-0.12	0.55	0.13	0.10	0.14	1.00		
nFU	0.04	0.01	-0.16	-0.19	0.16	0.05	0.08	0.06	0.86	-0.03	0.35	1.00	
whi	0.09	-0.19	0.16	0.17	0.26	-0.16	0.26	0.12	-0.16	0.19	0.29	-0.09	1.00
k	-0.37	-0.00	-0.12	-0.12	-0.17	0.12	-0.14	-0.30	0.04	-0.10	-0.18	-0.02	-0.09

(note: k row also shows 1.00 in final column)

```
(obs=6,294)
```

	swb	met	age	age2	inc	une	male	hea	kid	col	mar	nFU	whi
swb	1.0000												
met	-0.0233	1.0000											
age	0.0701	-0.0482	1.0000										

```
     age2 |   0.0704  -0.0482   0.9850   1.0000
      inc |   0.1675   0.0793   0.0604   0.0308   1.0000
      une |  -0.0931   0.0118  -0.1420  -0.1345  -0.1376   1.0000
     male |   0.0619  -0.0174  -0.0316  -0.0407   0.2897  -0.0217   1.0000
      hea |   0.3035   0.0486  -0.1854  -0.1776   0.2135  -0.0393   0.1189   1.0000
      kid |   0.0501  -0.0036  -0.2768  -0.2929   0.0566   0.0246  -0.0676   0.0553   1.0000
      col |  -0.0072   0.1033  -0.0532  -0.0633   0.2552  -0.1330   0.0317   0.1209  -0.0449   1.0000
      mar |   0.2059  -0.0181   0.1509   0.1293   0.4559  -0.1102   0.5009   0.1282   0.1246   0.1368   1.0000
      nFU |   0.1095  -0.0064  -0.1650  -0.1919   0.2014  -0.0005   0.0708   0.0593   0.8656  -0.0266   0.3526   1.0000
      whi |   0.0585  -0.1388   0.1392   0.1496   0.2572  -0.1132   0.2024   0.0756  -0.0782   0.1494   0.2623  -0.0181   1.0000
        k |  -0.3863  -0.0110  -0.1877  -0.1813  -0.1506   0.1098  -0.1076  -0.3022   0.0164  -0.0556  -0.1855  -0.0381  -0.0161   1.00
```

3.3   Panel Structure of Metro Variable

xttab met

```
            |        Overall        |       Between        |   Within
        met |    Freq.   Percent     |    Freq.   Percent   |   Percent
   ---------+-----------------------------------------------------------
    Inap.:  |     8294     21.98     |     2947     29.26   |    77.39
    Metropol|    29436     78.02     |     8362     83.01   |    93.19
   ---------+-----------------------------------------------------------
      Total |    37730    100.00     |    11309    112.27   |    89.07
                                      (n = 10073)
```

xtsum met

```
Variable         |      Mean    Std. Dev.       Min        Max |    Observations
-----------------+--------------------------------------------+----------------
met      overall | .7801749    .4141335          0          1 |    N =   37730
         between |             .385049            0          1 |    n =   10073
         within  |             .1878218   -.0531584   1.613508 |  T-bar = 3.74566
```

3.4   Single Year And Pooled Results V FE Results

TODO check and cp writeup to discussion; also do mention in results or data that RE and single year and pooled in SOM, adn cp footnote about retaining head only frollowing gordon

Time invariant person-level characteristics, such as personality traits do matter–the metro unhappiness disadvantage is only about half in fixed effects model v single-year or pooled data.

Table 4: Regressions of SWB: FE v OLS 2015

	c1-FE	c1-2015	c2-FE	c2-2015	c3-FE	c3-2015	c4-FE	c4-2015	c5-FE	c5-2015
metro	0.01	-0.11***	-0.04*	-0.12***	-0.03*	-0.10***	-0.04**	-0.09***	-0.04*	-0.08***
age			0.02***	-0.00	0.02***	-0.01**	0.01***	-0.01***	0.00	-0.01***
age sq			-0.00**	0.00**	-0.00	0.00***	-0.00	0.00***	-0.00	0.00***
last year total family income			0.00*	0.00***	0.00	0.00**	0.00	0.00*	0.00	0.00*
unemployed			-0.18***	-0.23***	-0.18***	-0.24***	-0.16***	-0.18***	-0.16***	-0.18***
male			0.27	0.08***	0.21	-0.05*	0.07	-0.06**	0.08	-0.06**
health			0.13***	0.26***	0.13***	0.26***	0.10***	0.18***	0.10***	0.18***
kids					-0.01	-0.04*	-0.01	-0.03	-0.01	-0.03
college					-0.08*	-0.12***	-0.07	-0.13***	-0.07	-0.13***
married					0.18***	0.28***	0.17***	0.24***	0.17***	0.24***
family unit size					0.04***	0.06***	0.03***	0.05***	0.03***	0.05***
distress							-0.05***	-0.06***	-0.05***	-0.06***
constant	3.71***	3.82***	2.37***	2.80***	2.45***	2.93***	2.90***	3.63***	3.60***	3.74***
state and year dummies	no	no	no	no	no	no	no	no	yes	yes
N	37567	6256	37489	6250	36285	6092	36142	6068	36142	6068

*** p<0.01, ** p<0.05, * p<0.1;
robust std err (OLS)

6

Table 5: Regressions of SWB: FE v OLS Pooled 2015-2019.

	d1-FE	d1-09-19	d2-FE	d2-09-19	d3-FE	d3-09-19	d4-FE	d4-09-19	d5-FE	d5-09-19
metro	0.01	-0.11***	-0.04*	-0.12***	-0.03*	-0.10***	-0.04**	-0.09***	-0.04*	-0.08***
age			0.02***	-0.00	0.02***	-0.01**	0.01***	-0.01***	0.00	-0.01***
age sq			-0.00**	0.00**	-0.00	0.00***	-0.00	0.00***	-0.00	0.00***
last year total family income			0.00*	0.00***	0.00	0.00**	0.00	0.00*	0.00	0.00*
unemployed			-0.18***	-0.23***	-0.18***	-0.24***	-0.16***	-0.18***	-0.16***	-0.18***
male			0.27	0.08***	0.21	-0.05*	0.07	-0.06**	0.08	-0.06**
health			0.13***	0.26***	0.13***	0.26***	0.10***	0.18***	0.10***	0.18***
kids					-0.01	-0.04*	-0.01	-0.03	-0.01	-0.03
college					-0.08*	-0.12***	-0.07	-0.13***	-0.07	-0.13***
married					0.18***	0.28***	0.17***	0.24***	0.17***	0.24***
family unit size					0.04***	0.06***	0.03***	0.05***	0.03***	0.05***
distress							-0.05***	-0.06***	-0.05***	-0.06***
constant	3.71***	3.82***	2.37***	2.80***	2.45***	2.93***	2.90***	3.63***	3.60***	3.74***
state and year dummies	no	no	no	no	no	no	no	no	yes	yes
N	37567	6256	37489	6250	36285	6092	36142	6068	36142	6068

*** $p<0.01$, ** $p<0.05$, * $p<0.1$;
robust std err (OLS)

3.5 Random Effects

Table 6: RE regressions of SWB.

	b1	b2	b3	b4	b5
metro	-0.04***	-0.05***	-0.04***	-0.05***	-0.06***
age		0.00	-0.01***	-0.01***	-0.01***
age sq		0.00***	0.00***	0.00***	0.00***
last year total family income		0.00***	0.00***	0.00***	0.00***
unemployed		-0.22***	-0.22***	-0.19***	-0.19***
male		0.07***	-0.05***	-0.08***	-0.08***
health		0.19***	0.19***	0.14***	0.14***
kids			-0.02**	-0.02**	-0.02**
college			-0.06***	-0.08***	-0.08***
married			0.27***	0.24***	0.25***
family unit size			0.04***	0.04***	0.04***
distress				-0.05***	-0.05***
constant	3.74***	2.92***	3.03***	3.59***	3.68***
state and year dummies	no	no	no	no	yes
N	37567	37489	36285	36142	36142

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

References

ALCOCK, I., M. P. WHITE, B. W. WHEELER, L. E. FLEMING, AND M. H. DEPLEDGE (2014): "Longitudinal effects on mental health of moving to greener and less green urban areas," *Environmental science & technology*, 48, 1247–1255.

BROWN, G. D. A. AND J. GATHERGOOD (2019): "Consumption Changes, Not Income Changes, Predict Changes in Subjective Well-Being," *Social Psychological and Personality Science*, 0, 1948550619835215.

BURGER, M. J., P. S. MORRISON, M. HENDRIKS, AND M. M. HOOGERBRUGGE (2020): "Urban-Rural Happiness Differentials across the World," *World Happiness Report*.

DAVIES, W. (2015): *The Happiness Industry: How the Government and Big Business Sold us Well-Being*, Verso Books.

DIENER, E. (2009): *Well-being for public policy*, Oxford University Press, New York NY.

HOOGERBRUGGE, M. AND M. BURGER (????): "Selective Migration and Urban-Rural Differences in Subjective Well-being: Evidence from the United Kingdom'," *Urban Studies*.

Okulicz-Kozaryn, A. (2015): *Happiness and Place. Why Life is Better Outside of the City.*, Palgrave Macmillan, New York NY.

Okulicz-Kozaryn, A. and R. R. Valente (2021): "Urban unhappiness is common," *Cities*, 103368.

O'Sullivan, A. (2009): *Urban economics*, McGraw-Hill.

Rehdanz, K. and D. Maddison (2008): "Local environmental quality and life-satisfaction in Germany," *Ecological economics*, 64, 787–797.

Stiglitz, J., A. Sen, and J. Fitoussi (2009): "Report by the Commission on the measurement of economic performance and social progress," *Available at www.stiglitz-sen-fitoussi.fr*.

White, M. P., I. Alcock, B. W. Wheeler, and M. H. Depledge (2013a): "Coastal proximity, health and well-being: Results from a longitudinal panel survey," *Health & Place.*

——— (2013b): "Would You Be Happier Living in a Greener Urban Area? A Fixed-Effects Analysis of Panel Data," *Psychological science*, 24, 920–928.

Wirth, L. (1938): "Urbanism as a Way of Life," *American Journal of Sociology*, 44, 1–24.