

# Unhappy Metros: US Panel Data Evidence

Adam Okulicz-Kozaryn\*

Rutgers - Camden

Wednesday 3<sup>rd</sup> November, 2021 18:34

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 cross-sectional findings: metros are less happy than nonmetros.

SUBJECTIVE WELLBEING (SWB), LIFE SATISFACTION, HAPPINESS, PSID, XXX TODO ADD TO EBIB AS KEYWORD PAPER-CODE-NAME AND TAG WITH EBIB KEYWORDS

Urbanism is not just built environment but a way of life (Wirth 1938) affecting in humans in multiple and profound ways, indeed urbanism is arguably the most significant disruption of human habitat in our species history MY CITY BOOK. World is urbanizing at an astonishing pace –that recent stat from .6 to 6b in 100 years or sth from recent paper! 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 SEN AND DIENER BOOK FOR PUB POLICY MAKING

The evidence of urban-rural happiness gradient is overwhelming. urban unhappiness is common and some morrison stuff and couple others like 10 from that boilerplate in recent cities article about least happy places around the world specific cities but guess 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.

Yet all studies to date are cross sectional and panel evidence is missing. Few studies that use panel data do not actually test the gradient. Hoogerbrugge and Burger (????) use inadequate measurement, the urban-rural cutoff is at extremely low population size of 10,000 people or even 3,000 people for Scotland. Large villages and small towns are not really “urban” lacking defining features of urbanness: size, density, and heterogeneity Wirth (1938). The build environment in villages or small towns lacks tall buildings, urban transit, airports, etc. Way of life in such places is not urban, lacking shalowness transitoriness, etc The way of life lacks transitoriness, etc these words from city book. cite when metropolis is too big and urban unhappiness is common—ideally it should be a gradient, and if necessarily a binary distinction then it is several hundred thousand, not 3 or 10 thousand as in Hoogerbrugge and Burger (????).

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. Rehdanz and Maddison (2008) uses a German panel dataset (GSOEP), 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 0).

---

\*EMAIL: adam.okulicz.kozaryn@gmail.com

I thank Gordon D. A. Brown for sharing STATA code. All mistakes are mine.

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 constant over time.

panel structure description of metro variable is shown in the appendix yes one problem is that there is little variability in urb-rur controls are set in table ?? var\_des in SOM

in controls we follow my cities when metropolis is too big (make sure we do):

race! cities more minorities and minorities less happy (berry US swb paper)

liberal democrat, rep con—cities more liberal/democrat and democrats less happy my jap paper

it is important to control for employment status—it both predicts swb, especially unemployment 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:

$$SWB_{it} = \gamma METRO_{it} X_{it} \beta + \alpha_i + u_{it} \quad (1)$$

Where,  $METRO_{it}$  is a metro dummy for person  $i$  at time  $t$ .  $\gamma$  is the main coefficient of interest.  $\alpha_i$  ( $i=1\dots n$ ) 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.  $\beta$  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 there is limited variability across time and hence time-invariant variables cannot be estimated, notably on metro which is not that often changed

## 2 Results

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 hea

### 3 Conclusion and discussion

## 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]

### 3.1 Variables' Definitions

**Table 3:** Variable definitions.

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 otherwise
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 otherwise
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 otherwise
distress	The K-6 Non-Specific Psychological Distress Scale

### 3.2 Summary statistics

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

n = 10108  
 T = 6

Distribution of T\_i: min 1 5% 1 25% 2 50% 4 75% 6 95% 6 max 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

	swb	met	age	age2	inc	une	male
swb	1.0000						
met	-0.0777*	1.0000					
age	0.0845*	-0.0424*	1.0000				
age2	0.0862*	-0.0494*	0.9821*	1.0000			
inc	0.1292*	0.0637*	0.0918*	0.0497*	1.0000		
une	-0.1247*	0.0175	-0.1885*	-0.1795*	-0.1272*	1.0000	
male	0.0846*	-0.0588*	-0.0125	-0.0256	0.2901*	-0.0173	1.0000
hea	0.2610*	0.0263*	-0.2471*	-0.2404*	0.2028*	-0.0052	0.1445*
kid	-0.0030	0.0154	-0.2874*	-0.3019*	0.0129	0.0904*	-0.0911*
col	0.0432*	0.0780*	-0.0847*	-0.0955*	0.2438*	-0.1242*	0.0619*
mar	0.2020*	-0.0566*	0.1662*	0.1341*	0.4214*	-0.1160*	0.5524*
nFU	0.0456*	0.0159	-0.1561*	-0.1877*	0.1589*	0.0429*	0.0868*
whi	0.0889*	-0.1813*	0.1524*	0.1609*	0.2491*	-0.1519*	0.2460*
k	-0.3653*	-0.0041	-0.1151*	-0.1069*	-0.1761*	0.1175*	-0.1473*
	hea	kid	col	mar	nFU	whi	k
hea	1.0000						
kid	0.0502*	1.0000					
col	0.1741*	-0.0370*	1.0000				
mar	0.1270*	0.1009*	0.1451*	1.0000			
nFU	0.0532*	0.8620*	-0.0247	0.3508*	1.0000		
whi	0.1129*	-0.1561*	0.1912*	0.2796*	-0.0857*	1.0000	
k	-0.3066*	0.0306*	-0.1049*	-0.1843*	-0.0308*	-0.0836*	1.0000

	swb	met	age	age2	inc	une	male
-----							
swb	1.0000						
met	-0.0227	1.0000					
age	0.0682*	-0.0407*	1.0000				
age2	0.0686*	-0.0417*	0.9850*	1.0000			
inc	0.1606*	0.0764*	0.0547*	0.0255*	1.0000		
une	-0.0889*	0.0075	-0.1425*	-0.1350*	-0.1374*	1.0000	
male	0.0631*	-0.0192	-0.0321*	-0.0416*	0.2889*	-0.0238	1.0000
hea	0.3006*	0.0418*	-0.1905*	-0.1827*	0.2125*	-0.0357*	0.1202*
kid	0.0514*	-0.0061	-0.2798*	-0.2961*	0.0524*	0.0246*	-0.0592*
col	-0.0061	0.1029*	-0.0514*	-0.0615*	0.2578*	-0.1350*	0.0318*
mar	0.2058*	-0.0111	0.1519*	0.1295*	0.4476*	-0.1123*	0.5064*
nFU	0.1109*	-0.0045	-0.1631*	-0.1911*	0.1962*	-0.0045	0.0803*
whi	0.0568*	-0.1340*	0.1379*	0.1481*	0.2469*	-0.1125*	0.1887*
k	-0.3799*	-0.0095	-0.1768*	-0.1697*	-0.1489*	0.1104*	-0.1115*
-----							
	hea	kid	col	mar	nFU	whi	k
-----							
hea	1.0000						
kid	0.0530*	1.0000					
col	0.1242*	-0.0456*	1.0000				
mar	0.1217*	0.1258*	0.1375*	1.0000			
nFU	0.0545*	0.8620*	-0.0265*	0.3553*	1.0000		
whi	0.0734*	-0.0755*	0.1525*	0.2474*	-0.0181	1.0000	
k	-0.2978*	0.0113	-0.0548*	-0.1872*	-0.0414*	-0.0129	1.0000

### 3.3 Panel Structure of Metro Variable

xttab met

met	Overall		Between		Within
	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

## 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 (0): "Consumption Changes, Not Income Changes, Predict Changes in Subjective Well-Being," *Social Psychological and Personality Science*, 0, 1948550619835215.
- HOOGERBRUGGE, M. AND M. BURGER (????): "Selective Migration and Urban-Rural Differences in Subjective Well-being: Evidence from the United Kingdom'," *Urban Studies*.
- 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.
- 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.