11/17/23

- Regressions with area added as a control
- More data sources found for different areas

Simple OLS on outcomes (itm 823: purchase transit)

Explanatory variable: num_stations, the number of stations that exist within the observation's (household's) district

```
. regress itm823 num_stations area itm190 year b3_1 b4_1 b5_1 b11_1 b13_1 b14_1 b16_1
```

Source	SS	df	MS			
Model	8.1264e+11	11	7.3877e+10			
Residual	9.4703e+12	29,600	319944124			
Total	1.0283e+13	29,611	347269229			

Number of obs = 29,612
F(11, 29600) = 230.90
Prob > F = 0.0000
R-squared = 0.0790
Adj R-squared = 0.0787
Root MSE = 17887

itm823	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
num_stations	170.642	44.42827	3.84	0.000	83.5606	257.7233
area	-.2695097	29.24194	-0.01	0.993	-57.585	57.04598
itm190	.0064526	.0001763	36.60	0.000	.0061071	.0067982
year	118.3346	28.30216	4.18	0.000	62.86112	173.8081
b3_1	2124.495	266.6415	7.97	0.000	1601.866	2647.124
b4_1	151.9006	10.68371	14.22	0.000	130.96	172.8411
b5_1	196.2896	62.21161	3.16	0.002	74.35213	318.2271
b11_1	-991.9594	99.72492	-9.95	0.000	-1187.425	-796.4942
b13_1	-47.83084	616.2193	-0.08	0.938	-1255.648	1159.986
b14_1	2016.986	485.2837	4.16	0.000	1065.809	2968.164
b16_1	-9.829504	3.176544	-3.09	0.002	-16.05567	-3.603337
_cons	-231547.9	56372.55	-4.11	0.000	-342040.6	-121055.2

Motor Premiums

```
. regress itm825 num_stations area itm190 year b3_1 b4_1 b5_1 b11_1 b13_1 b14_1 b16_1
```

Source	SS	df	MS		Number of obs	=	17,495
					F(11, 17483)	=	175.04
Model	2.2936e+11	11	2.0851e+10		Prob > F	=	0.0000
Residual	2.0826e+12	17,483	119119296		R-squared	=	0.0992
					Adj R-squared	=	0.0986
Total	2.3119e+12	17,494	132155171		Root MSE	=	10914

itm825	Coef.	Std. Err.	t	P>|t|	[95% Conf. Interval]	
num_stations	-96.23014	31.01581	-3.10	0.002	-157.0242	-35.43607
area	11.67036	22.94985	0.51	0.611	-33.31364	56.65436
itm190	.0038986	.0001304	29.89	0.000	.0036429	.0041543
year	-468.5102	24.23289	-19.33	0.000	-516.0091	-421.0113
b3_1	354.391	208.4132	1.70	0.089	-54.11954	762.9016
b4_1	-9.033233	8.934498	-1.01	0.312	-26.54574	8.479273
b5_1	240.7397	51.73062	4.65	0.000	139.3425	342.1368
b11_1	-1178.432	77.43157	-15.22	0.000	-1330.206	-1026.659
b13_1	4066.13	513.5917	7.92	0.000	3059.439	5072.821
b14_1	456.2581	426.1222	1.07	0.284	-378.9838	1291.5
b16_1	-3.785686	2.6101	-1.45	0.147	-8.901742	1.330369
_cons	941479.3	48356.19	19.47	0.000	846696.3	1036262

Water charge

```
. regress itm765 num_stations area itm190 year b3_1 b4_1 b5_1 b11_1 b13_1 b14_1 b16_1
```

Source	SS	df	MS
Model	2.7606e+10	11	2.5097e+09
Residual	2.1123e+11	29,648	7124681.67
Total	2.3884e+11	29,659	8052829.23

Number of obs	=	29,660
F(11, 29648)	=	352.25
Prob > F	=	0.0000
R-squared	=	0.1156
Adj R-squared	=	0.1153
Root MSE	=	2669.2

itm765	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
num_stations	20.55418	6.622482	3.10	0.002	7.573821	33.53453
area	.0899345	4.364142	0.02	0.984	-8.463976	8.643845
itm190	.00083	.0000263	31.58	0.000	.0007785	.0008815
year	136.4895	4.217663	32.36	0.000	128.2227	144.7563
b3_1	-89.0607	39.72147	-2.24	0.025	-166.9165	-11.20487
b4_1	3.646307	1.594119	2.29	0.022	.5217632	6.770851
b5_1	-100.388	9.287452	-10.81	0.000	-118.5918	-82.18415
b11_1	-195.5313	14.91261	-13.11	0.000	-224.7606	-166.3019
b13_1	250.173	92.08389	2.72	0.007	69.68449	430.6615
b14_1	654.8965	72.44419	9.04	0.000	512.9027	796.8903
b16_1	-5.099749	.4731649	-10.78	0.000	-6.027173	-4.172324
_cons	-268308.3	8400.743	-31.94	0.000	-284774.1	-251842.4

Petrol Fees and Parking tolls

```
. regress itm828 num_stations area itm190 year b3_1 b4_1 b5_1 b11_1 b13_1 b14_1 b16_1
```

Source	SS	df	MS
Model	2.8301e+12	11	2.5728e+11
Residual	1.4238e+13	19,452	731943500
Total	1.7068e+13	19,463	876939642

Number of obs	= 19,464
F(11, 19452)	= 351.51
Prob > F	= 0.0000
R-squared	= 0.1658
Adj R-squared	= 0.1653
Root MSE	= 27054

itm828	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
num_stations	-159.9984	74.19486	-2.16	0.031	-305.4267	-14.5701
area	261.728	54.20507	4.83	0.000	155.4814	367.9746
itm190	.0133979	.0003147	42.57	0.000	.012781	.0140148
year	965.1164	58.80426	16.41	0.000	849.855	1080.378
b3_1	1277.879	489.3404	2.61	0.009	318.7296	2237.028
b4_1	-60.0614	20.8177	-2.89	0.004	-100.8659	-19.25693
b5_1	1282.974	120.2499	10.67	0.000	1047.274	1518.674
b11_1	-3359.943	178.9272	-18.78	0.000	-3710.656	-3009.231
b13_1	7843.453	1189.96	6.59	0.000	5511.03	10175.88
b14_1	5282.12	986.1429	5.36	0.000	3349.195	7215.045
b16_1	-48.461	5.946258	-8.15	0.000	-60.11617	-36.80582
_cons	-1918348	117288.7	-16.36	0.000	-2148244	-1688452

Rent and Water Expenditure

```
. regress itm760 num_stations area itm190 year b3_1 b4_1 b5_1 b11_1 b13_1 b14_1 b16_1
```

Source	SS	df	MS			
				Number of obs	=	29,850
				F(11, 29838)	=	761.68
Model	9.4081e+13	11	8.5528e+12	Prob > F	=	0.0000
Residual	3.3505e+14	29,838	1.1229e+10	R-squared	=	0.2192
				Adj R-squared	=	0.2189
Total	4.2913e+14	29,849	1.4377e+10	Root MSE	=	1.1e+05

itm760	Coef.	Std. Err.	t	P>|t|	[95% Conf. Interval]	
num_stations	2748.723	262.3825	10.48	0.000	2234.441	3263.004
area	-2235.347	172.5516	-12.95	0.000	-2573.555	-1897.138
itm190	.0533082	.0010412	51.20	0.000	.0512674	.0553491
year	3710.664	166.9942	22.22	0.000	3383.348	4037.98
b3_1	10494.02	1572.18	6.67	0.000	7412.476	13575.56
b4_1	836.6901	63.05313	13.27	0.000	713.1032	960.277
b5_1	6430.164	367.2466	17.51	0.000	5710.344	7149.983
b11_1	-14885.79	588.7085	-25.29	0.000	-16039.69	-13731.9
b13_1	39983.48	3638.603	10.99	0.000	32851.66	47115.3
b14_1	17316.48	2863.18	6.05	0.000	11704.52	22928.44
b16_1	-151.9291	18.70484	-8.12	0.000	-188.5914	-115.2668
_cons	-7209587	332625.1	-21.67	0.000	-7861547	-6557627

More data sources: Mexico

Household survey data from IPUMS-I, geography level: municipality

Over 2100 municipalities, 15 years, quarterly, ~ 126,000 observations

Mexico

2020Q3 ⓘ	2020Q1 ⓘ	2019Q4 ⓘ	2019Q3 ⓘ	2019Q2 ⓘ	2019Q1 ⓘ	2018Q3 ⓘ	2018Q1 ⓘ
2018Q2 ⓘ	2018Q4 ⓘ	2017Q4 ⓘ	2017Q3 ⓘ	2017Q2 ⓘ	2017Q1 ⓘ	2016Q4 ⓘ	2016Q1 ⓘ
2016Q2 ⓘ	2016Q3 ⓘ	2015Q4 ⓘ	2015Q3 ⓘ	2015Q2 ⓘ	2015Q1 ⓘ	2014Q4 ⓘ	2014Q3 ⓘ
2014Q2 ⓘ	2014Q1 ⓘ	2013Q4 ⓘ	2013Q3 ⓘ	2013Q2 ⓘ	2013Q1 ⓘ	2012Q4 ⓘ	2012Q3 ⓘ
2012Q2 ⓘ	2012Q1 ⓘ	2011Q4 ⓘ	2011Q3 ⓘ	2011Q2 ⓘ	2011Q1 ⓘ	2010Q4 ⓘ	2010Q3 ⓘ
2010Q2 ⓘ	2010Q1 ⓘ	2009Q1 ⓘ	2009Q2 ⓘ	2009Q4 ⓘ	2009Q3 ⓘ	2008Q4 ⓘ	2008Q3 ⓘ
2008Q2 ⓘ	2008Q1 ⓘ	2007Q2 ⓘ	2007Q1 ⓘ	2007Q3 ⓘ	2007Q4 ⓘ	2006Q4 ⓘ	2006Q3 ⓘ
2006Q2 ⓘ	2006Q1 ⓘ	2005Q1 ⓘ	2005Q4 ⓘ	2005Q3 ⓘ	2005Q2 ⓘ		

China

EPS China data

- City level, for 70 cities, each month, for 15 years (same amount of time as taiwan) ~ 12,600 observations
- Data on transit, health, finance, energy, environment

Society and People's Livelihood

- China Education Statistics
- China Household Survey Statistics
- China Health Statistics
- China Transportation Statistics
- China Tourism Statistics
- China Culture Statistics
- China Civil Affairs Statistics
- China Poverty Alleviation Statistics

United States

Census tract, block group, and even exact block in some cases

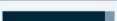However, averaged over 5 years

SELECT DATA ❓

| 6439 SOURCE TABLES | 111 TIME SERIES TABLES | 902 GIS FILES |

| PAGE 1 OF 13 ▶▶ ▶❙ 500 ▾ | VIEW 1 - 500 OF 6,439 |

POPULARITY	TABLE NAME	UNIVERSE	CLASSIFICATIONS	YEAR - DATASET	BREAKDOWNS
➕	B01003. Total Population	Total population		2017_2021_ACS5a	Spatial
➕	B19013. Median Household Income in the Past 12 Months (in 2021 Inflation-Adjusted Dollars)	Households		2017_2021_ACS5a	Spatial
➕	B01003. Total Population	Total population		2016_2020_ACS5a	Spatial
➕	P1. Total Population	Total population		2010_SF1a	Spatial
➕	P1. Total Population	Total population		2020_DHCa	Spatial
➕	B02001. Race	Total population	Race (9)	2017_2021_ACS5a	Spatial