Spatial Economics – Assignment 2

Max Heinze (h11742049@s.wu.ac.at)   Jaime Miravet (h12235992@s.wu.ac.at)
Jan Trimmel (h11809096@s.wu.ac.at)

April 12, 2024

Contents

> *The executable code that was used in compiling the assignment is available on GitHub at*
> *https://github.com/maxmheinze/spatial.*

Task A

Text

Task B

Nice Maps

We create four maps:

Distance from Jesuit Missions by Municipality

distmiss
	0.1 to 56.8
	56.8 to 113.5
	113.5 to 170.2
	170.2 to 226.9
	226.9 to 283.6
	283.6 to 340.3
	340.3 to 396.9
	396.9 to 453.6
	453.6 to 510.3
	510.3 to 567.0
	Missing

Distance from Franciscan Missions by Municipality

distfran
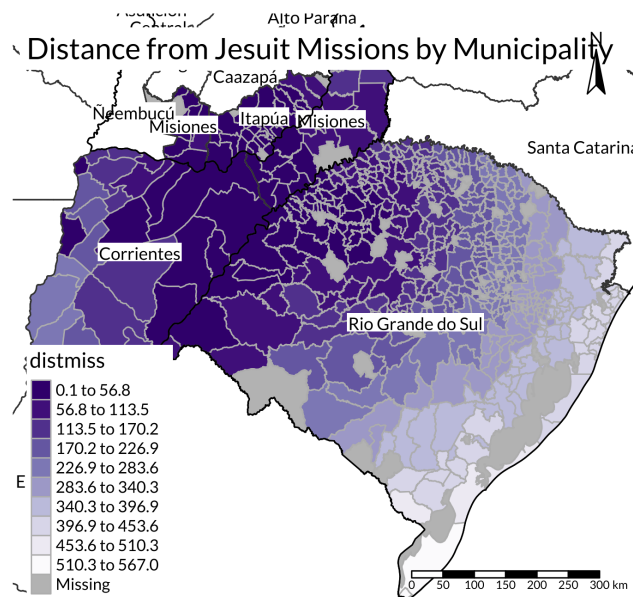|	
	17.9 to 99.2
	99.2 to 180.5
	180.5 to 261.8
	261.8 to 343.0
	343.0 to 424.3
	424.3 to 505.6
	505.6 to 586.9
	586.9 to 668.2
	668.2 to 749.5
	749.5 to 830.8
	Missing

Replication of Table 2

Below are code and results of our replication of Table 2 from Valencia Caicedo ([2019](#)). Since the original author used Stata's robust standard errors, a notorious problem for replication in R, we specifically report Stata-style standard errors in the table below using the `starprep` function from the `estimatr` package. Using these standard errors, we can reproduce both coefficients and standard errors exactly.

```
# Replicate Results --------------------------------------------------------
col1 <- lm(illiteracy ~ distmiss + lati + longi + corr + ita + mis + mis1, data = litr)
col2 <- lm(illiteracy ~ distmiss + lati + longi + area + tempe + alti + preci + rugg +
    river + coast + corr + ita + mis + mis1, data = litr)


litr_bra <- subset(litr, country == "BRA")
col3 <- lm(illiteracy ~ distmiss + lati + longi + as.factor(mesorregi), data = litr_bra)
col4 <- lm(illiteracy ~ distmiss + lati + longi + area + tempe + alti + preci + rugg +
    river + coast + as.factor(mesorregi), data = litr_bra)


litr_arg <- subset(litr, country == "Argentina")
col5 <- lm(illiteracy ~ distmiss + lati + longi + corr, data = litr_arg)
col6 <- lm(illiteracy ~ distmiss + lati + longi + area + tempe + alti + preci + rugg +
```

```
    river + coast + corr, data = litr_arg)

litr_pry <- subset(litr, country == "Paraguay")
col7 <- lm(illiteracy ~ distmiss + ita, data = litr_pry)
col8 <- lm(illiteracy ~ distmiss + area + tempe + alti + preci + rugg + river + coast +
    ita, data = litr_pry)
```

	Dependent variable:			
	illiteracy			
	(1)	(2)	(3)	(4)
distmiss	0.011***	0.011**	0.016*	0.030***
	(0.004)	(0.005)	(0.009)	(0.010)
lati	0.556**	0.070	0.408	4.575**
	(0.238)	(0.781)	(0.553)	(1.807)
longi	−1.108***	−1.007*	−1.022	−5.694***
	(0.257)	(0.556)	(0.689)	(1.811)
area		0.0001		−0.0002
		(0.0002)		(0.0004)
tempe		0.059		−0.062
		(0.077)		(0.124)
alti		0.006		0.001
		(0.004)		(0.005)
preci		−0.003		0.001
		(0.002)		(0.003)
rugg		−0.00000		−0.00000
		(0.00000)		(0.00000)
river		1.470**		1.723*
		(0.712)		(0.893)
coast		0.209		−4.976**
		(0.894)		(2.147)
corr	−5.341***	−6.032***		
	(1.286)	(1.583)		
ita	−3.187***	−2.409***		
	(0.728)	(0.833)		
mis	−4.324***	−4.734***		
	(1.122)	(1.488)		
mis1	−3.279***	−2.299**		
	(0.860)	(0.974)		
as.factor(mesorregi)4302			−2.720***	−2.543**
			(0.863)	(1.037)
as.factor(mesorregi)4303			−0.483	−0.383
			(1.133)	(1.228)
as.factor(mesorregi)4304			−0.771	0.196
			(1.150)	(1.403)
as.factor(mesorregi)4305			−3.023**	−1.290
			(1.312)	(1.541)
as.factor(mesorregi)4306			−1.724	−3.421
			(2.015)	(2.184)
as.factor(mesorregi)4307			−0.437	0.327
			(2.553)	(2.634)
Constant	−35.328***	−53.741*	−35.274	−143.869**
	(11.797)	(32.497)	(38.125)	(66.532)
Observations	549	548	467	467
R^2	0.042	0.073	0.094	0.135
Adjusted R^2	0.029	0.049	0.076	0.104
Residual Std. Error	3.948 (df = 541)	3.912 (df = 533)	4.040 (df = 457)	3.978 (df = 450)

Note:	*p<0.1; **p<0.05; ***p<0.01

	Dependent variable:			
	illiteracy			
	(1)	(2)	(3)	(4)
distmiss	0.016**	0.067***	0.005	0.014
	(0.007)	(0.022)	(0.012)	(0.027)
lati	0.084	−9.338**		
	(0.758)	(3.831)		
longi	1.095	7.186***		
	(0.803)	(2.676)		
area		−0.0001		0.0004
		(0.0003)		(0.001)
tempe		0.968***		0.360
		(0.237)		(0.220)
alti		0.065***		0.016
		(0.011)		(0.014)
preci		−0.017**		0.0001
		(0.008)		(0.005)
rugg		−0.00005***		0.0001
		(0.00002)		(0.00005)
river		9.795***		0.983
		(2.173)		(5.492)
coast		1.889		0.826
		(3.522)		(4.557)
corr	3.771**	−3.043		
	(1.850)	(3.644)		
ita			−0.231	0.829
			(0.832)	(2.316)
Constant	69.263*	−41.058	8.673***	−80.723*
	(36.792)	(54.628)	(0.694)	(43.888)
Observations	42	42	40	39
R^2	0.165	0.669	0.004	0.251
Adjusted R^2	0.075	0.547	−0.050	0.019
Residual Std. Error	2.924 (df = 37)	2.045 (df = 30)	2.150 (df = 37)	2.101 (df = 29)
Note:				*p<0.1; **p<0.05; ***p<0.01

Next, we try to reproduce the Conley standard errors. We try two different approaches, first using the `conleyreg` package and then using the `fixest` package. Valencia Caicedo (2019) specifies a cutoff distance of 0.1 degrees. Both packages we use only allow us to specify the cutoff distance in kilometers, so for the sake of simplicity, we use the distance that 0.1 degrees equal at the equator, which is 6 nautical miles, or approximately 11.112 kilometers. We first print the results from the `conleyreg` package and then from the `fixest` package.

```
lit1 <- litr %>%
    drop_na(lati, longi) %>%
    mutate(lat = lati, lon = longi)

col1c <- conleyreg(illiteracy ~ distmiss + lati + longi + corr + ita + mis + mis1,
    data = lit1, dist_cutoff = 11.112, lat = "lat", lon = "lon")
col2c <- conleyreg(illiteracy ~ distmiss + lati + longi + area + tempe + alti + preci +
    rugg + river + coast + corr + ita + mis + mis1, data = lit1, dist_cutoff = 11.112,
    lat = "lat", lon = "lon")

lit1_bra <- subset(lit1, country == "BRA")
col3c <- conleyreg(illiteracy ~ distmiss + lati + longi + as.factor(mesorregi), data =
lit1_bra,
    dist_cutoff = 11.112, lat = "lat", lon = "lon")
col4c <- conleyreg(illiteracy ~ distmiss + lati + longi + area + tempe + alti + preci +
    rugg + river + coast + as.factor(mesorregi), data = lit1_bra, dist_cutoff = 11.112,
    lat = "lat", lon = "lon")

lit1_arg <- subset(lit1, country == "Argentina")
col5c <- conleyreg(illiteracy ~ distmiss + lati + longi + corr, data = lit1_arg,
    dist_cutoff = 11.112, lat = "lat", lon = "lon")
col6c <- conleyreg(illiteracy ~ distmiss + lati + longi + area + tempe + alti + preci +
```

```
    rugg + river + coast + corr, data = lit1_arg, dist_cutoff = 11.112, lat = "lat",
    lon = "lon")

lit1_pry <- subset(lit1, country == "Paraguay")
col7c <- conleyreg(illiteracy ~ distmiss + ita, data = lit1_pry, dist_cutoff = 11.112,
    lat = "lat", lon = "lon")
col8c <- conleyreg(illiteracy ~ distmiss + area + tempe + alti + preci + rugg + river +
    coast + ita, data = lit1_pry, dist_cutoff = 11.112, lat = "lat", lon = "lon")
```

	Dependent variable:			
	(1)	(2)	(3)	(4)
distmiss	0.011***	0.011**	0.016*	0.030***
	(0.004)	(0.005)	(0.009)	(0.010)
lati	0.556**	0.070	0.408	4.575**
	(0.247)	(0.773)	(0.570)	(1.792)
longi	−1.108***	−1.007*	−1.022	−5.694***
	(0.266)	(0.549)	(0.696)	(1.780)
area		0.0001		−0.0002
		(0.0002)		(0.0003)
tempe		0.059		−0.062
		(0.080)		(0.130)
alti		0.006		0.001
		(0.004)		(0.006)
preci		−0.003		0.001
		(0.002)		(0.003)
rugg		−0.00000		−0.00000
		(0.00000)		(0.00000)
river		1.470**		1.723*
		(0.731)		(0.904)
coast		0.209		−4.976**
		(0.885)		(2.125)
corr	−5.341***	−6.032***		
	(1.325)	(1.612)		
ita	−3.187***	−2.409***		
	(0.758)	(0.848)		
mis	−4.324***	−4.734***		
	(1.159)	(1.519)		
mis1	−3.279***	−2.299**		
	(0.876)	(0.980)		
as.factor(mesorregi)4302			−2.720***	−2.543**
			(0.896)	(1.067)
as.factor(mesorregi)4303			−0.483	−0.383
			(1.150)	(1.250)
as.factor(mesorregi)4304			−0.771	0.196
			(1.215)	(1.448)
as.factor(mesorregi)4305			−3.023**	−1.290
			(1.393)	(1.659)
as.factor(mesorregi)4306			−1.724	−3.421
			(2.027)	(2.190)
as.factor(mesorregi)4307			−0.437	0.327
			(2.576)	(2.654)
Constant	−35.328***	−53.741	−35.274	−143.869**
	(12.074)	(33.356)	(38.313)	(66.608)

Note:	$^*p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

	Dependent variable:			
	(5)	(6)	(7)	(8)
distmiss	0.016**	0.067***	0.005	0.014
	(0.007)	(0.019)	(0.011)	(0.023)
lati	0.084	−9.338***		
	(0.711)	(3.238)		
longi	1.095	7.186***		
	(0.754)	(2.262)		
area		−0.0001		0.0004
		(0.0002)		(0.001)
tempe		0.968***		0.360*
		(0.200)		(0.190)
alti		0.065***		0.016
		(0.010)		(0.012)
preci		−0.017**		0.0001
		(0.007)		(0.004)
rugg		−0.00005***		0.0001
		(0.00002)		(0.00004)
river		9.795***		0.983
		(1.837)		(4.738)
coast		1.889		0.826
		(2.976)		(3.941)
corr	3.771**	−3.043		
	(1.736)	(3.080)		
ita			−0.231	0.829
			(0.794)	(1.998)
Constant	69.263*	−41.058	8.673***	−80.723**
	(34.532)	(46.169)	(0.666)	(37.732)

Note:	*p<0.1; **p<0.05; ***p<0.01

```r
col1cf <- feols(illiteracy ~ distmiss + lati + longi + corr + ita + mis + mis1, data = litr,
    vcov_conley(lat = "lati", lon = "longi", cutoff = 11.112, distance = "spherical"))
col2cf <- feols(illiteracy ~ distmiss + lati + longi + area + tempe + alti + preci +
    rugg + river + coast + corr + ita + mis + mis1, data = litr, vcov_conley(lat = "lati",
    lon = "longi", cutoff = 11.112, distance = "spherical"))

col3cf <- feols(illiteracy ~ distmiss + lati + longi + as.factor(mesorregi), data =
litr_bra,
    vcov_conley(lat = "lati", lon = "longi", cutoff = 11.112, distance = "spherical"))
col4cf <- feols(illiteracy ~ distmiss + lati + longi + area + tempe + alti + preci +
    rugg + river + coast + as.factor(mesorregi), data = litr_bra, vcov_conley(lat = "lati",
    lon = "longi", cutoff = 11.112, distance = "spherical"))

col5cf <- feols(illiteracy ~ distmiss + lati + longi + corr, data = litr_arg,
vcov_conley(lat = "lati",
    lon = "longi", cutoff = 11.112, distance = "spherical"))
col6cf <- feols(illiteracy ~ distmiss + lati + longi + area + tempe + alti + preci +
    rugg + river + coast + corr, data = litr_arg, vcov_conley(lat = "lati", lon = "longi",
    cutoff = 11.112, distance = "spherical"))

col7cf <- feols(illiteracy ~ distmiss + ita, data = litr_pry, vcov_conley(lat = "lati",
    lon = "longi", cutoff = 11.112, distance = "spherical"))
col8cf <- feols(illiteracy ~ distmiss + area + tempe + alti + preci + rugg + river +
    coast + ita, data = litr_pry, vcov_conley(lat = "lati", lon = "longi", cutoff = 11.112,
    distance = "spherical"))
```

| Dependent Variable: | | | | illiteracy | | | | |
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Variables								
Constant	-35.33***	-53.74	-35.27	-143.9*	69.26*	-41.06	8.673***	-80.72*
	(12.96)	(38.28)	(40.26)	(73.54)	(36.35)	(53.97)	(0.6825)	(43.22)
distmiss	0.0105***	0.0112*	0.0164*	0.0297**	0.0157**	0.0669***	0.0045	0.0138
	(0.0040)	(0.0057)	(0.0092)	(0.0115)	(0.0073)	(0.0217)	(0.0112)	(0.0261)
lati	0.5561*	0.0698	0.4078	4.575**	0.0837	-9.338**		
	(0.2841)	(0.7983)	(0.6512)	(1.898)	(0.7486)	(3.785)		
longi	-1.108***	-1.007*	-1.022	-5.694***	1.095	7.186**		
	(0.2957)	(0.5727)	(0.7468)	(1.857)	(0.7939)	(2.644)		
corr	-5.341***	-6.032***			3.771**	-3.043		
	(1.462)	(1.776)			(1.828)	(3.600)		
ita	-3.187***	-2.409**					-0.2311	0.8290
	(0.8686)	(0.9485)					(0.8123)	(2.294)
mis	-4.324***	-4.734***						
	(1.289)	(1.695)						
mis1	-3.279***	-2.299**						
	(0.9438)	(1.042)						
area		0.0001		-0.0002		-8.9×10^{-5}		0.0004
		(0.0002)		(0.0004)		(0.0003)		(0.0007)
tempe		0.0587		-0.0625		0.9675***		0.3598
		(0.0937)		(0.1558)		(0.2343)		(0.2205)
alti		0.0057		0.0006		0.0654***		0.0160
		(0.0042)		(0.0067)		(0.0112)		(0.0137)
preci		-0.0026		0.0010		-0.0171**		0.0001
		(0.0027)		(0.0037)		(0.0083)		(0.0051)
rugg		-3.56×10^{-6}		-3.09×10^{-6}		-4.8×10^{-5}**		6.93×10^{-5}
		(4.8×10^{-6})		(4.71×10^{-6})		(1.82×10^{-5})		(4.85×10^{-5})
river		1.470*		1.723*		9.795***		0.9834
		(0.8202)		(1.011)		(2.147)		(5.445)
coast		0.2086		-4.976**		1.889		0.8264
		(0.9177)		(2.244)		(3.480)		(4.543)
as.factor(mesorregi)4302			-2.720**	-2.543**				
			(1.065)	(1.256)				
as.factor(mesorregi)4303			-0.4829	-0.3830				
			(1.232)	(1.382)				
as.factor(mesorregi)4304			-0.7711	0.1960				
			(1.426)	(1.686)				
as.factor(mesorregi)4305			-3.023*	-1.290				
			(1.648)	(1.987)				
as.factor(mesorregi)4306			-1.724	-3.421				
			(2.130)	(2.350)				
as.factor(mesorregi)4307			-0.4368	0.3267				
			(2.751)	(2.881)				
Fit statistics								
Observations	549	548	467	467	42	42	40	39
R^2	0.04178	0.07299	0.09402	0.13481	0.16514	0.66887	0.00359	0.25135
Adjusted R^2	0.02938	0.04864	0.07618	0.10405	0.07488	0.54745	-0.05027	0.01901

Conley (11.112km) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Evidently, we could not reproduce the exact Conley standard errors reported in Valencia Caicedo (2019). And the two packages yielded different errors even though we specified the same cutoff (11.112 kilometers) and the same method of distance calculation (spherical).

Ways to Use Distance

In his specification, Felipe Valencia Caicedo (2019) uses the following specification (Eq. 1 in the paper):

$$Y_{2000,ij} = \alpha + \beta d(M_{ij}) + \gamma GEO_{ij} + \mu_j + \varepsilon_{ij},$$

where $d(\cdot)$ is described as either being a dummy for the presence of a mission or the plain (i.e., untransformed) distance, although we do not fine examples of the former in any of the printed tables in the paper, so we will just be calling it the distance. The decaying impact of a mission can be plotted against distance as follows:
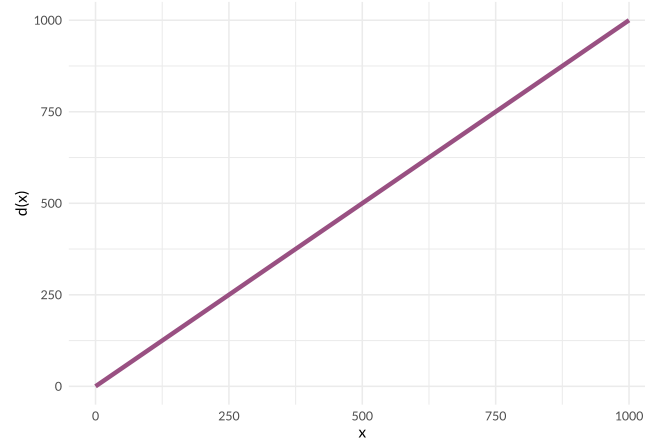
If we view $d(\cdot)$ as *impact* exerted by a mission, it might be more intuitive to plot the negative of it, so that a higher value corresponds to more impact.
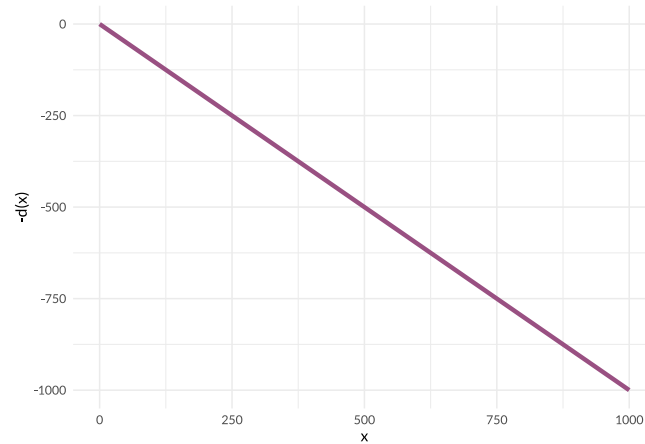
If we draw a map and plot distance from a mission using this function, we get the same map as above.
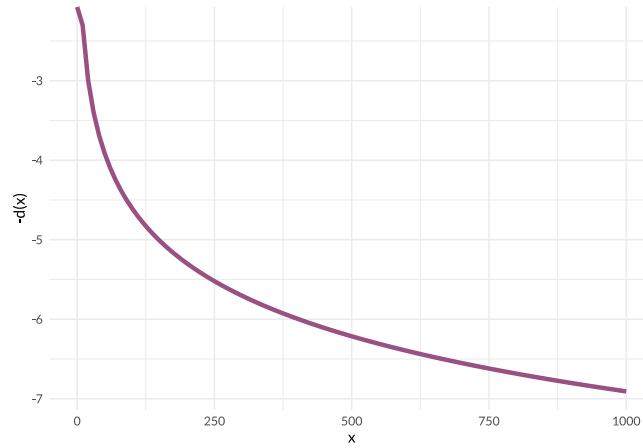

Distance from Jesuit Missions, Linear
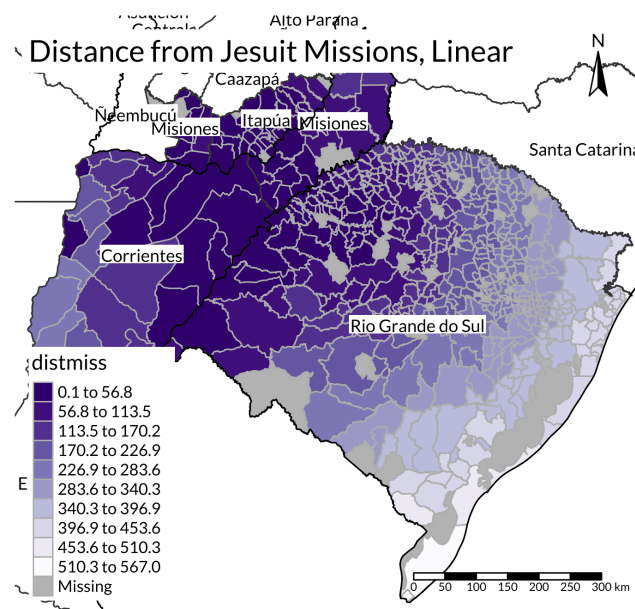
Letting d be

$$d : \mathbb{R} \to \mathbb{R}, \quad x \mapsto x$$

means that we attribute the same change in impact to moving from 0km from the treatment position to 100km from it, as to moving from 900km distance to 1000km. This is quite an assumption. Alternatively, we could use the **logarithm** of the distance,

$$d : \mathbb{R} \to \mathbb{R}, \quad x \mapsto \log(x),$$
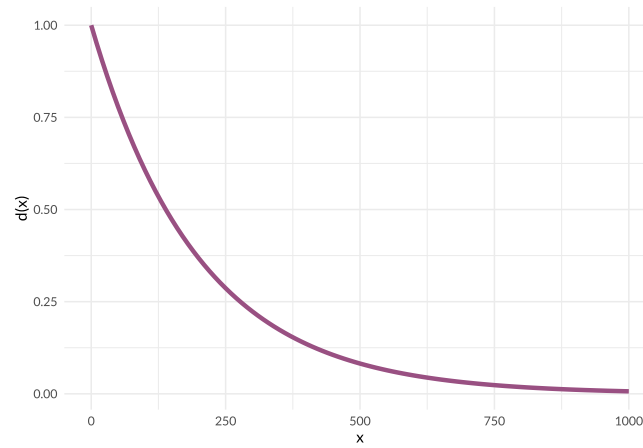
like this:

This gives the following map.
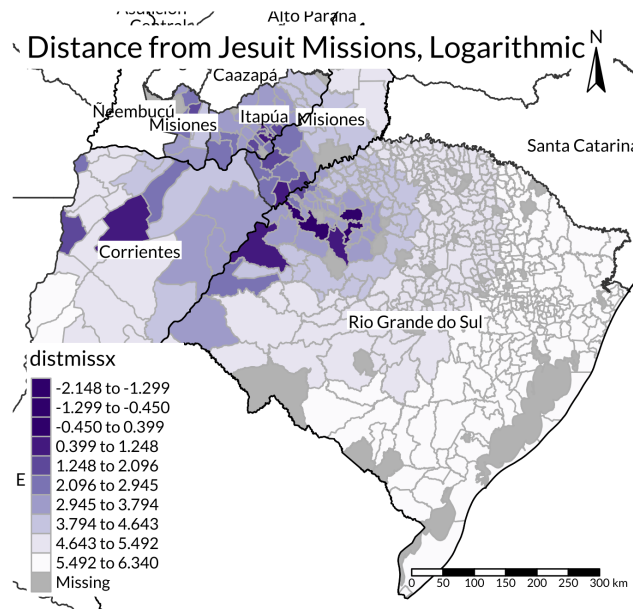
We can see that unlike in the linear case, the differences between distances at a greater length from the locations of treatment are weighted down, whereas distance differences closer to the mission are exaggerated. It could make sense to use such a specification instead of a linear one if we suspect impacts of a mission to be roughly comparable between places that are 400 or 500 kilometers from the closest mission, or if we are concerned that portions of Rio Grande do Sul are included in the analysis that are much farther from where Jesuits went than portions of Argentinian or Paraguayan states that were not included.
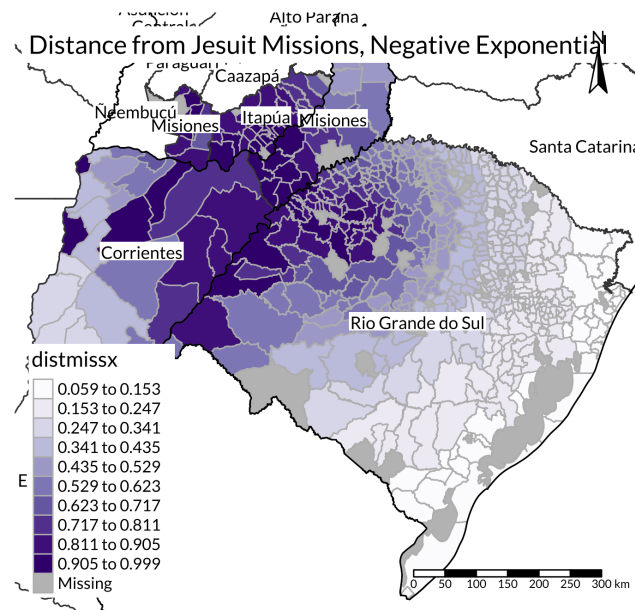
Alternatively, we could consider a **negative exponential** specification,

$$d : \mathbb{R} \to \mathbb{R}, \quad x \mapsto \exp(-\lambda x),$$

which is plotted below for $\lambda = \frac{1}{500}$:
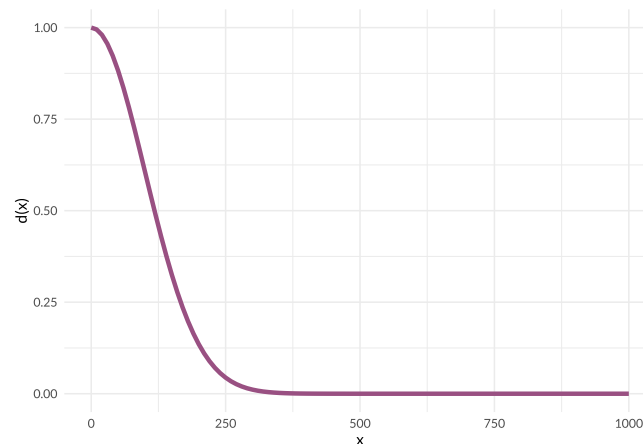
Applying to our data, it looks like this:

Distance from Jesuit Missions, Negative Exponential

By adjusting λ, we can make the decay more or less "aggressive." With our present specification, it is "between" the linear and logarithmic cases.
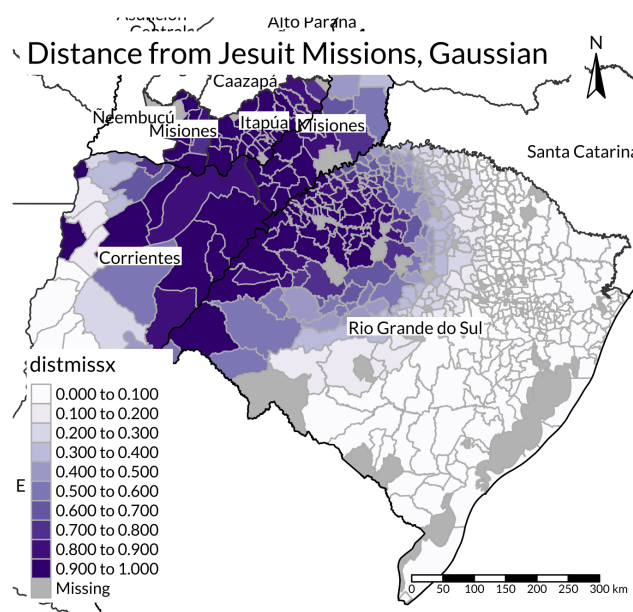
Lastly, we could consider using a **Gaussian** decay function, like this:

$$d : \mathbb{R} \to \mathbb{R}, \quad x \mapsto \exp\left(-\frac{x^2}{2\sigma^2}\right)$$

Letting $\sigma = 100$ gives the following:

We can see that by using the Gaussian decay function, we can model something that is closer to a "cutoff." Applying this function to the data gives the following map:

We can clearly see that different functions yield vastly different distance measures across municipalities. So it is quite sensible to suspect that estimation results are subject to the choice of distance measure. To check this, we run the regression from column (2) of Table 2 again – but each time with a different distance measure.

```r
litrd <- litr %>%
    mutate(distmiss_log = log(distmiss), distmiss_exp = exp(-distmiss/200), distmiss_gau =
    exp(-(distmiss^2)/(2 *
        100^2)))

col2_reg <- lm(illiteracy ~ distmiss + lati + longi + area + tempe + alti + preci +
    rugg + river + coast + corr + ita + mis + mis1, data = litrd)
col2_log <- lm(illiteracy ~ distmiss_log + lati + longi + area + tempe + alti + preci +
    rugg + river + coast + corr + ita + mis + mis1, data = litrd)
col2_exp <- lm(illiteracy ~ distmiss_exp + lati + longi + area + tempe + alti + preci +
    rugg + river + coast + corr + ita + mis + mis1, data = litrd)
col2_gau <- lm(illiteracy ~ distmiss_gau + lati + longi + area + tempe + alti + preci +
    rugg + river + coast + corr + ita + mis + mis1, data = litrd)
```

	Dependent variable:			
	illiteracy			
	(1)	(2)	(3)	(4)
distmiss	0.011**			
	(0.005)			
distmiss_log		0.078		
		(0.289)		
distmiss_exp			−3.293	
			(2.082)	
distmiss_gau				−1.631
				(1.246)
lati	0.070	−0.065	−0.216	−0.182
	(0.781)	(0.795)	(0.805)	(0.796)
longi	−1.007*	−0.360	−0.580	−0.454
	(0.556)	(0.536)	(0.535)	(0.539)
area	0.0001	0.0001	0.0001	0.0001
	(0.0002)	(0.0002)	(0.0002)	(0.0002)
tempe	0.059	−0.028	0.014	−0.009
	(0.077)	(0.071)	(0.075)	(0.070)
alti	0.006	0.002	0.003	0.002
	(0.004)	(0.003)	(0.003)	(0.003)
preci	−0.003	−0.003	−0.002	−0.002
	(0.002)	(0.002)	(0.002)	(0.002)
rugg	−0.00000	−0.00000	−0.00000	−0.00000
	(0.00000)	(0.00000)	(0.00000)	(0.00000)
river	1.470**	1.305*	1.292*	1.330*
	(0.712)	(0.717)	(0.710)	(0.713)
coast	0.209	0.564	0.578	0.652
	(0.894)	(0.874)	(0.890)	(0.892)
corr	−6.032***	−3.740***	−4.638***	−4.191***
	(1.583)	(1.419)	(1.451)	(1.451)
ita	−2.409***	−1.922**	−1.776**	−1.823**
	(0.833)	(0.888)	(0.841)	(0.832)
mis	−4.734***	−3.296**	−3.489**	−3.389**
	(1.488)	(1.398)	(1.364)	(1.345)
mis1	−2.299**	−2.212**	−1.970**	−2.072**
	(0.974)	(1.042)	(0.998)	(0.985)
Constant	−53.741*	−3.482	−28.162	−16.605
	(32.497)	(28.517)	(29.882)	(26.957)
Observations	548	548	548	548
R^2	0.073	0.063	0.067	0.066
Adjusted R^2	0.049	0.038	0.043	0.041
Residual Std. Error (df = 533)	3.912	3.934	3.924	3.927

Note: *p<0.1; **p<0.05; ***p<0.01

Of course, coefficient magnitudes cannot be directly compared, since we wildly scaled and transformed the distance measure. Even the sign changes do not necessarily need to worry us, since we transformed distance measures in a way that they sometimes shrank and sometimes increased with increasing distance. However, we see that using the different distance measures, effect sizes are no longer statistically different from zero.

Persistence and Space, or: Where's Waldo?

Because any text about spatial autocorrelation would be incomplete without reciting Waldo[1] Tobler's First Law of Geography, here it is in all its beauty: "Everything is related to everything else, but near things are more related than distant things." Applied to the present context, this has important implications.

First, the presence of a mission in some given municipality is likely autocorrelated with whether adjacent municipalities were subject to missionary activity too. If the first missionaries went to some place in what we today call Misiones, it is likely that subsequent missionaries built on locally specific knowledge of earlier missions and settled near them. Our distance plots, especially the logarithmic one, are a good indication of that autocorrelation, since we can see that all missions were more or less in the same region.

Second, we can assume literacy rates to be spatially autocorrelated as well. Again, we can see this from the map we originally produced where we plotted literacy rates across the regions analyzed by Valencia Caicedo (2019). We can see that high literacy rates concentrate in one region around the Argentinian state of Misiones and the northwestern part of Rio Grande do Sul, and in one region in the eastern part of Rio Grande do Sul. And if we take a glance at the map below, we can also see that these regions are clusters of higher population density.

Kelly (2021) notes that this covert relation leads to standard errors usually being too low since the presence of spatial clusters in two variables means that whatever correlation exists between the variables in one locality is likely to also be present in adjacent localities. He also notes that different types of available standard error corrections lead to widely different results. Conley standard errors, which Valencia Caicedo (2019) uses as a robustness check, tend to be the most conservative of the corrections. Having standard errors that are too small, of course, does not bias the estimate, but will cause false positive results.

The same can be thought of if there are potential confounders. If there is a unobserved spatial variable that influences the outcome and that is correlated with the covariates that are included in the model, the coefficient estimates will be biased (Dupont et al., 2023). Assuming the variables considered in the model in Valencia Caicedo (2019) behave like most other socioeconomic variables in space, we can assume that any spatial effect will be correlated with the model covariates, leading to the possibility of biased coefficients.

---

[1]Here he is.

Task C

The perils of ignoring peer effects

Task D

The image is a screenshot, and those are conventionally stored in PNG format. The photo *contained* in the screenshot is a photograph, and it is difficult to guess which format it was originally saved in. Let's say it's JPEG. Then, someone inserted the image into the assignment PDF, meaning it is technically not stored as a PNG anymore. What all those ways of storing the image have in common is that they are **raster formats**, as they consist of individual pixels. And even if we print the document containing the image, it gets printed as dots, which are not exactly pixels, but certainly form a raster rather than a vector.

References

Dupont, E., Marques, I., & Kneib, T. (2023). *Demystifying Spatial Confounding* (Working Paper). arXiv. https://doi.org/10.48550/arXiv.2309.16861

Kelly, M. (2021). *Persistence, randomization, and spatial noise* (Working Paper). Dublin: University College Dublin, UCD School of Economics. https://www.econstor.eu/handle/10419/246498

Valencia Caicedo, F. (2019). The Mission: Human Capital Transmission, Economic Persistence, and Culture in South America. *Q. J. Econ.*, *134*(1), 507–556. https://doi.org/10.1093/qje/qjy024