

empirical_project1_memo

Lindsey Greenhill

3/7/2022

looking at variables I care about

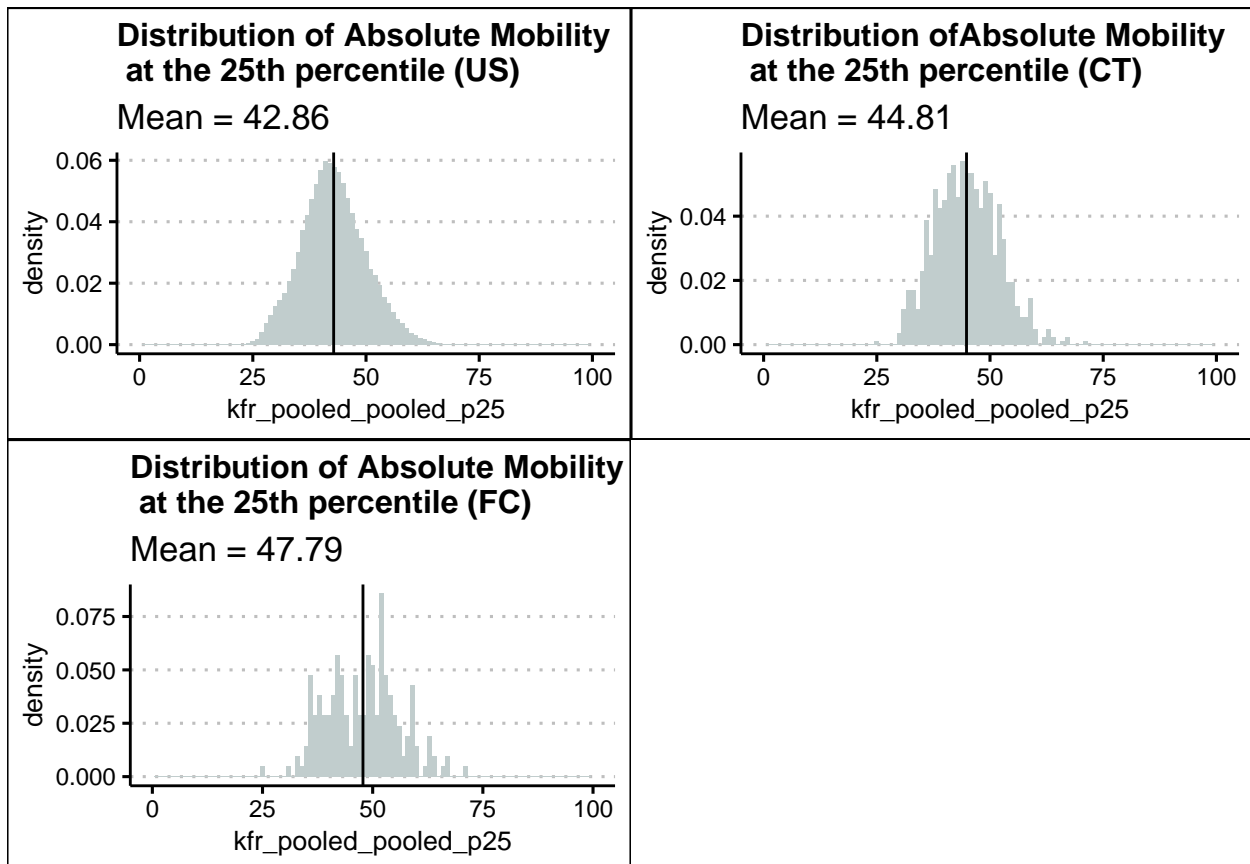
```
## Warning: Removed 1192 rows containing non-finite values (stat_bin).
```

```
## Warning: Removed 2 rows containing missing values (geom_bar).
```

```
## Warning: Removed 2 rows containing non-finite values (stat_bin).
```

```
## Warning: Removed 2 rows containing missing values (geom_bar).
```

```
## Warning: Removed 2 rows containing missing values (geom_bar).
```



Standard Deviations: US, CT, FC

FC has largest sd	
place	sd_kfr_pooled_pooled_p25
Fairfield County	8.15
CT	7.00
USA	7.12

looking at education now

here we see the national pattern

```
school_national <- atlas %>%
  ggplot(aes(x = gsmn_math_g3_2013, y = kfr_pooled_pooled_p25)) +
  stat_binmean(n = 20, geom = "point") +
  stat_smooth(method = "lm", se = FALSE) +
  labs(title = "Upward Mobility vs. Third Grade Standardized Tests",
        subtitle = "Mobility is higher in areas with higher scores") +
  theme_clean()

ct_acs <- get_acs(geography = "tract",
                  geometry = TRUE,
                  state = "CT",
                  variables = "B19013_001")
```

Getting data from the 2015-2019 5-year ACS

Downloading feature geometry from the Census website. To cache shapefiles for use in future sessions

|

```
ct_joined <- ct_acs %>%
  mutate(tract = as.double(str_sub(GEOID, start = -6))) %>%
  left_join(ct, by = "tract") %>%
  select(GEOID, county, NAME, tract, gsmn_math_g3_2013, kfr_pooled_pooled_p25) %>%
  filter(county == 1) %>%
  mutate(dot_color = as_factor(case_when(tract == 70300 ~ "Westside, Bridgeport",
                                          tract == 35400 ~ "New Canaan",
                                          TRUE ~ "other")))
```

now going to zoom in to fairfield county

```
school_fc <- ct_joined %>%
  ggplot(aes(x = gsmn_math_g3_2013, y = kfr_pooled_pooled_p25)) +
  geom_point(aes(col = dot_color)) +
  stat_smooth(method = "lm", se = FALSE) +
  scale_x_continuous(limits = c(.75, 7)) +
  scale_color_manual(values = c("black", "red", "forestgreen")) +
  annotate("text", x = 2.3, y = 25, label = "Westside, Bridgeport",
          col = "red") +
```

```

  annotate("text", x = 6.5, y = 66, label = "New Canaan",
          col = "forestgreen") +
  guides(col = FALSE) +
  labs(title = "Absolute mobility p25 vs. 3rd Grade Test Scores") +
  theme_clean()

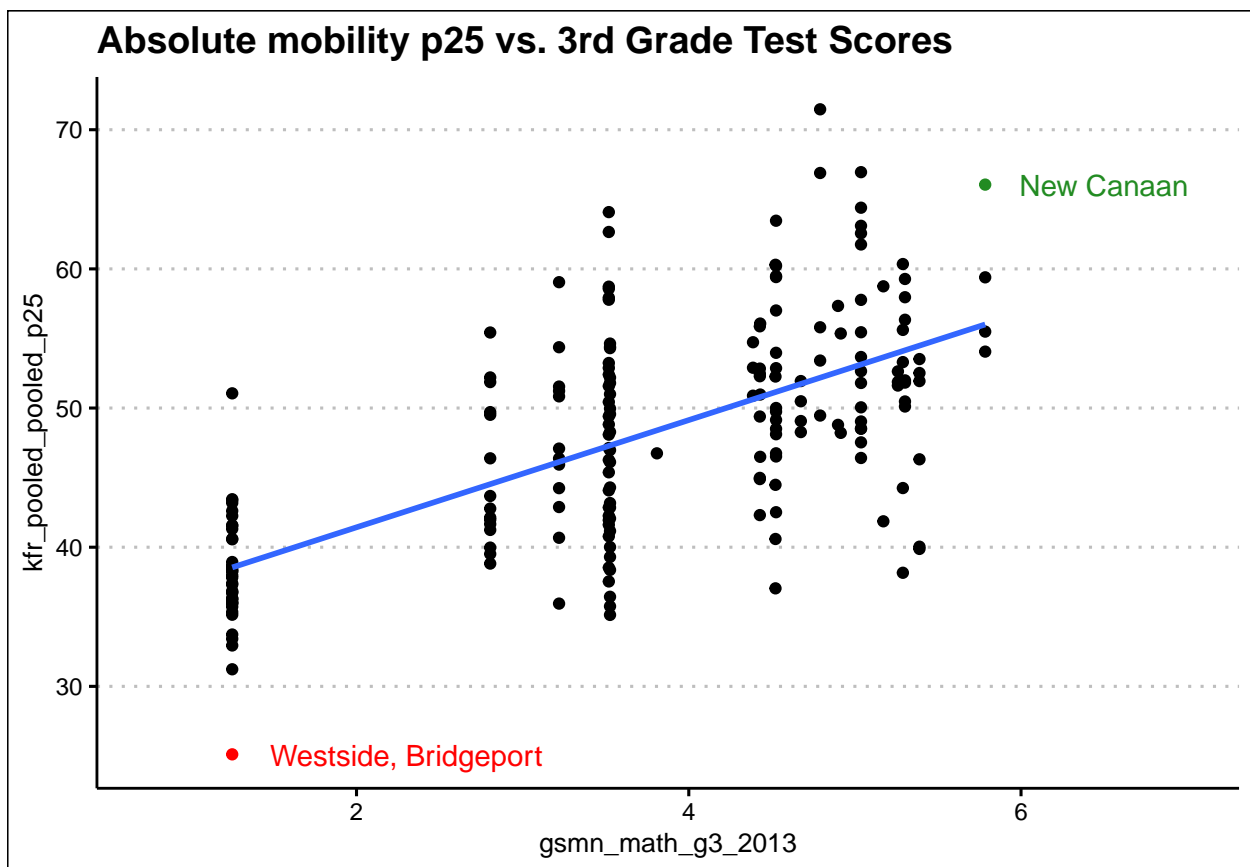
```

school_fc

```
## 'geom_smooth()' using formula 'y ~ x'
```

```
## Warning: Removed 6 rows containing non-finite values (stat_smooth).
```

```
## Warning: Removed 6 rows containing missing values (geom_point).
```



```

fc_limited <- fc %>%
  select(czname, tract, gsmn_math_g3_2013, kfr_pooled_pooled_p25) %>%
  arrange(gsmn_math_g3_2013) %>%
  head(10)

```

```
# lowest scores are tracts
```

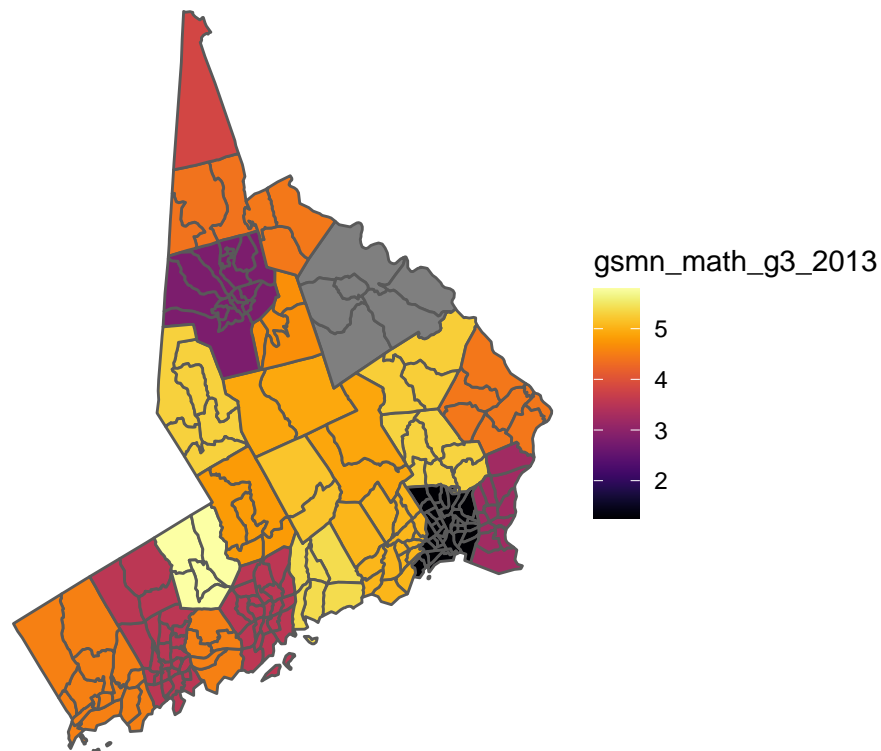
```

pal <- wes_palette("GrandBudapest1", 100, type = "continuous")

ct_joined %>%
  filter(county == 1) %>%
  ggplot(aes(fill = gsmn_math_g3_2013)) +
  geom_sf() +
  scale_fill_viridis(option = "B") +
  labs(title = "Fairfield County 3rd Grade Test Scores") +
  theme_void()

```

Fairfield County 3rd Grade Test Scores



bridgeport and Danbury have the worst it looks like

regressions

```

mod_1 <- lm(kfr_pooled_pooled_p25 ~ gsmn_math_g3_2013, data = fc)

stargazer(mod_1,
  type = "latex")

```

% Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu
 % Date and time: Wed, Mar 09, 2022 - 21:59:27

	<i>Dependent variable:</i>
	kfr_pooled_pooled_p25
gsmn_math_g3_2013	3.852*** (0.318)
Constant	33.726*** (1.238)
Observations	204
R ²	0.421
Adjusted R ²	0.418
Residual Std. Error	6.258 (df = 202)
F Statistic	146.887*** (df = 1; 202)
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01