Japonyanın eğitim sistemi çok farklı.

Çok katı, yoğun, ve disiplinli.

Özgürlük, çeşitlilik açısından kötü.

Öğrenciler yüksek oranlarda bir üst seviye eğitimedevam ediyorlar.

Bu da çalışmamızda yaşadığımız farklı seviyelerden mezun kişilerin nüfusa oranları arasında korelasyon sorununu doğuruyor.

4 kategori

Genel lise

Mesleki lise

Yuksek okul

Üniversite ve üstü

Datam Japan statstan.

[www.e-stat.go.jp](http://www.e-stat.go.jp)

Ne anket verisi ne de rapor var maaş ve eğitim seivyesine dair.

Bu yüzden kişi başı gelir ile üniversite mezunlarının nüfusa oranı arasınd ilişki araştırdık

Datam japonyanın bütün ilçeleri bazında.

Konuya dair iki kere sayım yapılmış

2000 ve 2010 yılında.

Toplam 3481 satırlık datam var.

İlçelerin kişi başına düşen gelire kendilerine has etkileri olduğu için,

Panel data kullandım.

Eksik yılı olan ilçeleri temizleyip oanel datayı dengeledim.

Within yöntemiyle ilçenin kendi etkisini yok etmeye çalıştım.

Kısaltmalar

Jhr = temel lise mezunlarının oranı

Shr= meslek lisesi

Jcr = önlisans

Scr = lisans ve üstü mezunlarının oranı

Hepsi ilçe bazında

All education level rates with within model

|  |
| --- |
| Oneway (individual) effect Within Model  Call:  plm(formula = per\_capita\_inc ~ scr + jhr + shr + jcr + non\_educ\_rate,  data = pdat, model = "within")  Balanced Panel: n = 1740, T = 2, N = 3480  Residuals:  Min. 1st Qu. Median 3rd Qu. Max.  -7.9111e+02 -2.4651e+01 1.1369e-13 2.4651e+01 7.9111e+02  Coefficients: (1 dropped because of singularities)  Estimate Std. Error t-value Pr(>|t|)  scr -25.10066 2.56228 -9.7962 < 2.2e-16 \*\*\*  jhr 5.63977 0.90249 6.2491 5.181e-10 \*\*\*  shr 0.49299 0.70656 0.6977 0.4854  jcr -30.42481 3.58640 -8.4834 < 2.2e-16 \*\*\*  ---  Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1  Total Sum of Squares: 26199000  Residual Sum of Squares: 9303600  R-Squared: 0.64489  Adj. R-Squared: 0.28835  F-statistic: 788.156 on 4 and 1736 DF, p-value: < 2.22e-16 |
|  |
| |  | | --- | | > | |

R bize Lisans ve üstü eğitimli kişilerin oranı ile ortalama gelir arasında negatif bir ilişki olduğunu (-25) hem de binde bir hata payıyla iddia ediyor. Bu elbette mantıksız.

Buradaki sorun scr (lisans ve üstü mezun oranı) tek başına artmıyor oluşu. Scr’si yüksek bir ilçenin örneğin jhr’si de yüksek, hem de daha büyük oranda yüksek.

Bu sorunu (multi collinearity) yok etmek için bir yöntem, aralarında korelasyon olan değişkenlerimizi regresyondan temizlemek bir yöntem.

Panel datada sadece uni ve üstü eğitimli kişilerin oranına bakıyorum:

Just Unversity+ graduates rate with within model

Oneway (individual) effect Within Model

Call:

plm(formula = per\_capita\_inc ~ scr, data = pdat, model = "within")

Balanced Panel: n = 1740, T = 2, N = 3480

Residuals:

Min. 1st Qu. Median 3rd Qu. Max.

-831.789 -26.586 0.000 26.586 831.789

Coefficients:

Estimate Std. Error t-value Pr(>|t|)

scr -59.9488 1.2158 -49.309 < 2.2e-16 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Total Sum of Squares: 26199000

Residual Sum of Squares: 10925000

R-Squared: 0.58302

Adj. R-Squared: 0.16579

F-statistic: 2431.42 on 1 and 1739 DF, p-value: < 2.22e-16

Bu sefer de negatif ilişki çıktı. R karemiz çok düşük, yani ortalama geliri tek başına açıklamaya çok yetersiz bir değişken, fakat yüksek mertebede anlamlı çıktı (binde bir)

Bu durum 2000 – 2010 arasında tesadüfi birşey olmamalı. Fakat within modeliyle, ilçelerin kendi ortalamalarından sapmalarına baktık. Yani herhangi bir ilçede o dönem aralığında ortalama gelir arttıysa, üniversite mezunu oranı da azalmış gibi görünüyor.

En azından üniversite mezun oranını bir miktar arttırmak, daha iyi koşullarda bir ilçe yaratmış olmak anlamına gelmiyor. Söz gelimi, insanlar, daha zorlu şartlarda daha fazla okumaya yönelmiş olabilirler.

Ya da iyi iş imkanları, onları okumaktan uzaklaştırmış da olabilir.

University graduates rate and unemployment rate with within model

Oneway (individual) effect Within Model

Call:

plm(formula = scr ~ unempr, data = pdat, model = "within")

Balanced Panel: n = 1740, T = 2, N = 3480

Residuals:

Min. 1st Qu. Median 3rd Qu. Max.

-4.60250 -0.48061 0.00000 0.48061 4.60250

Coefficients:

Estimate Std. Error t-value Pr(>|t|)

unempr 0.615652 0.013161 46.779 < 2.2e-16 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Total Sum of Squares: 4250.2

Residual Sum of Squares: 1882

R-Squared: 0.5572

Adj. R-Squared: 0.11414

F-statistic: 2188.28 on 1 and 1739 DF, p-value: < 2.22e-16

Bu da tezimizi doğrular yönde. Binde bir seviyesinde de anlamlı bir korelasyon bulduk.%1 lik işsizlikte artış, %0.6 lisans ve üstü mezun sayısında artış ile korele. Datamız uzun bir zaman serisi olmadığı için daha detaylı inceleyemiyoruz. R karem çok düşük, elbette üni mezun oranını esas açıklayacak şey işsizlik olmadığından.

University graduates rate and unemployment rate with pooling model

Pooling Model

Call:

plm(formula = scr ~ unemp, data = pdat, model = "pooling")

Balanced Panel: n = 1740, T = 2, N = 3480

Residuals:

Min. 1st Qu. Median 3rd Qu. Max.

-31.09373 -2.86119 -0.95436 1.82482 24.23595

Coefficients:

Estimate Std. Error t-value Pr(>|t|)

(Intercept) 7.9122e+00 7.7469e-02 102.133 < 2.2e-16 \*\*\*

unemp 2.7724e-04 1.2959e-05 21.393 < 2.2e-16 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Total Sum of Squares: 72330

Residual Sum of Squares: 63919

R-Squared: 0.11629

Adj. R-Squared: 0.11603

F-statistic: 457.667 on 1 and 3478 DF, p-value: < 2.22e-16

Aynı pozitif ve anlamlı ilişki, ilçe etkisinden arındırılmamış; iki gözlemimiz olduğu için, özetle trend etkisinden de arındırılmamış, genel ilişkileri için de doğru.

Src oranı ile bir çok başka macro şey korele olabilir.

Bunun yerine biz panel yöntemden vazgeçip pooling ile src gelir genel ilişkisini kontrol edelim.

Just University graduates rate with pooling model

Pooling Model

Call:

plm(formula = per\_capita\_inc ~ scr, data = pdat, model = "pooling")

Balanced Panel: n = 1740, T = 2, N = 3480

Residuals:

Min. 1st Qu. Median 3rd Qu. Max.

-658.621 -151.413 -10.705 140.610 3305.251

Coefficients:

Estimate Std. Error t-value Pr(>|t|)

(Intercept) 720.5754 7.8946 91.274 < 2.2e-16 \*\*\*

scr 53.3782 0.8195 65.135 < 2.2e-16 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Total Sum of Squares: 375030000

Residual Sum of Squares: 168940000

R-Squared: 0.54951

Adj. R-Squared: 0.54939

F-statistic: 4242.56 on 1 and 3478 DF, p-value: < 2.22e-16

Bu sefer pozitif ilişkili çıktı. %1 daha fazla mezunu olan ilçelerde ortalama gelir 53bin daha fazla.

R karem de bir miktar yükseldi.

Doğru yolda gibiyiz.

ek bilgi veriyorum maaşlarda artış fln dediklerimin birimi hep bin yenmiş. Yani 53 daha fazla yerine 53bin yen daha fazla demelisin!

Diğer eğitim seviyelerini deniyorum.

Under gradute level rates with pooling model

Pooling Model

Call:

plm(formula = per\_capita\_inc ~ jcr, data = pdat, model = "pooling")

Balanced Panel: n = 1740, T = 2, N = 3480

Residuals:

Min. 1st Qu. Median 3rd Qu. Max.

-836.422 -180.395 -6.147 161.732 4197.442

Coefficients:

Estimate Std. Error t-value Pr(>|t|)

(Intercept) 516.6866 17.2286 29.990 < 2.2e-16 \*\*\*

jcr 79.2309 2.0019 39.578 < 2.2e-16 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Total Sum of Squares: 375030000

Residual Sum of Squares: 258570000

R-Squared: 0.31052

Adj. R-Squared: 0.31033

F-statistic: 1566.41 on 1 and 3478 DF, p-value: < 2.22e-16

Junior college mezun oranları ile de gelir pozitif ilişkili.

Rakam yüzde bir mezun artışına 79 gelir artışı.

Biliyoruz ki jcr ve scr korele

İkisini birden sokarak deneyelim.

Under graduate + University graduates rate with pooling model

Pooling Model

Call:

plm(formula = per\_capita\_inc ~ jcr + scr, data = pdat, model = "pooling")

Balanced Panel: n = 1740, T = 2, N = 3480

Residuals:

Min. 1st Qu. Median 3rd Qu. Max.

-764.63 -148.53 -11.55 136.15 3099.01

Coefficients:

Estimate Std. Error t-value Pr(>|t|)

(Intercept) 860.8850 15.7274 54.738 < 2.2e-16 \*\*\*

jcr -29.8884 2.9117 -10.265 < 2.2e-16 \*\*\*

scr 66.0438 1.4746 44.787 < 2.2e-16 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Total Sum of Squares: 375030000

Residual Sum of Squares: 163980000

R-Squared: 0.56276

Adj. R-Squared: 0.56251

F-statistic: 2237.62 on 2 and 3477 DF, p-value: < 2.22e-16

R bu sefer gelir ortalamalarında oluşan varyansın sebebini iki mezun grubuna farklı dağıttı.

Aralarında bire bir bi korelasyon olmadığı için, bu dağılımı doğru varsayabiliriz.

Yani lisans ve üstü mezun oranının, önlisans mezun oranlarından daha pozitif etkilediğini.

İlçenin geliri, işsizlik, göç, hangi sektörlerin ağırlıkta olduğu gibi bir çok şeye bağlı olabilir.

Bunları da regresyona katıyorum.

University graduates rate and macro datas with pooling model

Pooling Model

Call:

plm(formula = per\_capita\_inc ~ scr + year + inmr + ourmr + lffr +

unempr + prim\_rate + sec\_rate + ter\_rate, data = pdat, model = "pooling")

Unbalanced Panel: n = 1728, T = 1-2, N = 2862

Residuals:

Min. 1st Qu. Median 3rd Qu. Max.

-880.142 -86.096 -8.551 76.805 3260.317

Coefficients: (1 dropped because of singularities)

Estimate Std. Error t-value Pr(>|t|)

(Intercept) 623.02353 79.25515 7.8610 5.357e-15 \*\*\*

scr 54.79488 1.10505 49.5861 < 2.2e-16 \*\*\*

year2010 -160.86299 9.01581 -17.8423 < 2.2e-16 \*\*\*

inmr 1.87003 4.59302 0.4071 0.683931

ourmr 17.09296 4.78120 3.5750 0.000356 \*\*\*

lffr 3.93191 1.41157 2.7855 0.005380 \*\*

unempr -30.03768 1.84887 -16.2465 < 2.2e-16 \*\*\*

prim\_rate -2.78453 0.51207 -5.4378 5.852e-08 \*\*\*

sec\_rate 4.04927 0.54806 7.3883 1.943e-13 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Total Sum of Squares: 341310000

Residual Sum of Squares: 86904000

R-Squared: 0.74538

Adj. R-Squared: 0.74467

F-statistic: 1043.99 on 8 and 2853 DF, p-value: < 2.22e-16

R 3. Sektör’ün toplamda payını drop etmiş. Normal, çünkü diğer iki sektörden artan paya eşit, bire bir multi collinearity’den elemiş.

Ortalama gelir ile anlamlı ilişkileri olan değerler:

İntercept. Reelde karşılığı olan bir yorumu olmadığı için geçiyorum.

Diğer makro etkilerden arındırdığımız haliyle, üni mezunlarının nüfusa oranında her yüzde bir artış yaklaşık 55bin artış yaratmış ortalama gelirde.

İşsizlik oranı anlamını yitirmiş.

Yıl etkisi varmış. 2000 den 2010 geldiğimizde ortalama gelirler 160bin yen azalmış.

İnmr içeri göç ourmr dışarı göç demek.

Ourmr etkiliymiş. Oranda yüzde birlik artış (nüfusa oranla dışa göç miktarı) ortalama gelirde 17bin yen artış ile ilişkili. Muhtemelen geliri düşük ya da gelirsiz kişiler göç etmişler.

Lffr: kadınların iş gücüne katılım oranı:buradaki yüzde bir artış 3900 yen gelirartışıyla ilişkili çıktı. Elbette çalışan oranında artışın ortalama gelire yansımasını beklerdik.

Unempr: işsizlikte yüzde bir artış, ortalama gelirlerde 30bin yen düşüşle ilişkili. Elbette işsizlikle ücretler ters ilikili olduğundan bu sonucumuz da gerçekçi görünüyor.

Gelire en pozitif katkıyı ikinci sektör: yani sanayi sağlıyor çıktı. Toplam payda oranında %1 lik artış ortalama gelirlere 4bin yen artışla ilişkili.

R karemiz 3bin üzerinde gözlemimiz olmasına rağmen %70lerin üzerinde çıktı. Bu iyi bir rakam.

Sonuç:

Doğrudan ölçemediğimiz maaş – üniversite diploması ilişkisini, dolaylı yoldan, ilçelerin ortalama geliri ve mezun oranları üzerinden inceledik.

Makro verilerin etkisinden arındırılmış halde dahi, üniversite mezunlarının çok olduğu bölgelerde gelirler daha yüksek.

Sadece iki dönemlik gözlem olduğundan panel datayı verimli kullanamadık. Modelimiz auto-regressive olabilir. İlçelerin kendilerinden gelen, gözlemleyemediğimiz bir sebep, yüksek maaşları yaratmış olabilir. Sonuçta r karemiz %75’e çıksa da geri kalan %25’in büyük kısmı burdan geliyor olabilir.

Bulduğumuz data ve çalışmaların yetersizliğinden kaynaklı olarak, güçlü şekilde iddia edemesek de japonya genelinde üniversite mezuniyeti ile daha yüksek gelirler ilişki görünüyor.