REVIEW

Araştırmada Güney Afrika’da üniversite eğitiminin gelirlere nasıl yansıdığını inceleyeceğiz.

Aslında Güney Afrika’da 2001 ve 2011 yıllarınd ayapılmış detay nüfus sayım ve anketleri mevcut ve erişimimize açık. Çalışmada hiç okula gitmemiş, some primary, some secondary, high education gibi gruplara ayrılmış.

SORUNLAR

1. Gözlem sayımız düşük:

* Anketler kamuoyuna doğrudan açık olmadığı için, binlerce gözlem, bizde yalnızca tek bir değer; ortalama değer olarak mevcut.
* Yıl gözlem sayımız iki, bu da panel datadan faydalanmak için yetersiz
* Toplam gözlem sayımız: 46

THEORY

Bu konudaki genel çalışmalar ülkelerden bağımsız olarak, üniversite eğitiminin maaşlara pozitif yansıdığını gösteriyor.

Biz Güney Afrika’da ne kadar pozitif buna bakmaya çalışacağız. Cinsiyet, renk gibi etmenlerden ne kadar etkilenmiş bunları göremeye çalışıcaz.

Hazandırıyorsa, hangi seviye bir eğitim ne kadar daha fazla kazandırmış bunu göreceğiz.

DATA

**Datayı Güney Afrika İstatistik Kurumunun resmi sitesinden oluşturduk.**

**Sadece 2001 ve 2011 yılına dair eğitimle ilgili bir sayım mevcut.**

**Her bir gözlemim şu şekilde oluşturarak datamı grupladım:**

**Renk, Cinsiyet, Bölge, Eğitim Seviyesi üzerinden identifierlar oluşturdum.**

**Aynı bölgede yaşayıp aynı cinsiyete ve aynı eğitim seviyesine sahip olan kişilerin;**

**Ne kadar bir oranının hangi gelir grubuna dahil olduğunun datasını oluşturdum.**

**Bana Güney Afrika istatistik Kurumunun müsade ettiği bütün mümkün alt gruplandırmalardan faydalanarak maksimum 46 farklı gözleme eriştim.**

**Excel yardımıyla ham datada sayımdan çıkan toplam rakamları, yüzdelere çevirdim. Regresyona bunları soktum.**

## ****Tanım:****

R0 R1 .... R10 R11 Gelir gruplarımızın kısaltmaları

R0 hiçbir geliri olmayanlar, R11 ise en yüksek gelir grubumuz

Model Reasoning

Cross Section Data’yı daha uygun gördüm. Çünkü 46 satır dar datamın her satırını bağımsız değerlendirerek daha fazla değişkeni regresyona sokabilirim.

Time series yapamam, çünkü sadece iki dönem gözlemim var.

Panel data kullanamam, çünkü hem dönem sayım düşük hem de elimizde anket datası gibi bir data mevcut değil.

İki dönemlik bir araştırma olduğu için, hiçbir makro değişkeni (büyüme, kur, enflasyon, işssizlik vs) regresyonumuza katamıyoruz. İkşer makro değerimiziz bütün satırlara kopyalanacağından R, varyans elde edemez ve bize herhangi bir katsayı döndüremez.

Result

**Just Education level**

İlk olarak en üst gelir kesimindeki nüfusun eğitim seviyesi dağılımına baktık.

reg1 = lm(data = dat, r11 ~ educ)

Coefficients:

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

(Intercept) 2.2091 3.9264 0.563 0.5738

educGrade 12/Std 10 3.4109 5.5528 0.614 0.5392

educHigher 6.2680 5.5528 1.129 0.2592

educNo schooling 12.1850 5.5528 2.194 0.0284 \*

educSome primary 0.5188 5.5528 0.093 0.9256

educSome secondary -0.1429 5.5528 -0.026 0.9795

---

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

Residual standard error: 57.71 on 1290 degrees of freedom

Multiple R-squared: 0.005861, Adjusted R-squared: 0.002008

F-statistic: 1.521 on 5 and 1290 DF, p-value: 0.1802

R aşırı düşük çıktı. Yani gelir gruplarının dağılımını eğitimle açıklamak zor görünüyor

Fakat herhangi bir eğitimi olmayan kesimin üst gelir seviyesinde, biraz daha yoğunlaştığına dair bir işaret var. (no schooling anlamlı ve pozitif çıktı)

**Education + Sex + Stae + Year + Color**

Elde edebildiğimiz diğer değişkenleri de regresyona ekleyerek, nüfusun gelir gruplarına dağılımını açıklamaya çalışalım

reg2 = lm(data = dat, r11 ~ educ + sex + state + year + color)

summary(reg2)

Coefficients:

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

(Intercept) 2.5492 7.1523 0.356 0.7216

educGrade 12/Std 10 3.4109 5.5401 0.616 0.5382

educHigher 6.2680 5.5401 1.131 0.2581

educNo schooling 12.1850 5.5401 2.199 0.0280 \*

educSome primary 0.5188 5.5401 0.094 0.9254

educSome secondary -0.1429 5.5401 -0.026 0.9794

sexMale 4.6288 3.1986 1.447 0.1481

stateFree State -10.5090 6.7853 -1.549 0.1217

stateGauteng -10.7027 6.7853 -1.577 0.1150

stateKwaZulu-Natal -11.3000 6.7853 -1.665 0.0961 .

stateLimpopo -10.5290 6.7853 -1.552 0.1210

stateMpumalanga -11.8257 6.7853 -1.743 0.0816 .

stateNorth West -10.3585 6.7853 -1.527 0.1271

stateNorthern Cape -7.8831 6.7853 -1.162 0.2455

stateWestern Cape -12.5024 6.7853 -1.843 0.0656 .

year2001 1.1622 3.9175 0.297 0.7668

year2011 7.7708 3.9175 1.984 0.0475 \*

colorColoured 0.4014 4.5235 0.089 0.9293

colorIndian or Asian 11.4450 4.5235 2.530 0.0115 \*

colorWhite 3.6739 4.5235 0.812 0.4168

---

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

Residual standard error: 57.57 on 1276 degrees of freedom

Multiple R-squared: 0.02114, Adjusted R-squared: 0.00656

F-statistic: 1.45 on 19 and 1276 DF, p-value: 0.09482

Eğitim seviyesi yine (no schooling dışında) anlamsız çıktı.

Bir kaç bölgenin dezavantajlı olduğuna dair %10 seviyesinde anlamlı işaretler var fakat güçlü değil:

KwaZulu-Natal

Mpumalanga

Western Cape

no schooling yine %12 civarında çıktı. Eğitim almamışların %12’si bu kesimde yer alıyor gibiler.

Bunun bir sebebi eğitimsiz bireylerin gelir gruplarına dağılımının daha katı olması da olabilir. Fakat 12 gelir grubundan en zengininde, eğitimsiz nüfusun %12’sinin bulunması yüksek bir oran.

2011 yılında bu gelir grubun oranı genel olarak artmış (en zengin r11)

Bunun sebebi, gelir gruplarımızın al ve üst sınırlarının statik olması olabilir. Nominal gelirler artmış ise dahi, nüfusun 10 sene içerisinde bir miktar yukarı gruplara kaymış gözükmesi normal.

R karemiz hala çok düşük.

Acaba en sadece en üst kesime bakmak sorunlu olabilir mi? Nüfusun küçük bir oranını temsil eden bu ayrıcalıklı grubu biraz genişletip tekrar deneyelim.

**Üst 4 gelir grubu ~ yıl + renk**

Reg4 = lm(data = dat, I(r11+r10+r9+r8) ~ year + color)

summary(reg4)

Call:

lm(formula = I(r11 + r10 + r9 + r8) ~ year + color, data = dat)

Residuals:

Min 1Q Median 3Q Max

-213.0 -103.0 -51.4 -5.8 17725.2

Coefficients:

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

(Intercept) -28.05 43.84 -0.640 0.52248

year2001 13.75 43.84 0.314 0.75392

year2011 108.66 43.84 2.478 0.01332 \*

colorColoured 42.55 50.63 0.840 0.40083

colorIndian or Asian 132.42 50.63 2.616 0.00901 \*\*

colorWhite 76.61 50.63 1.513 0.13044

---

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

Residual standard error: 644.4 on 1290 degrees of freedom

Multiple R-squared: 0.0112, Adjusted R-squared: 0.007369

F-statistic: 2.923 on 5 and 1290 DF, p-value: 0.01248

2011 yılında üst gelir gruplarına girmiş nüfus payında yine artış var

yine hintli ve asyalılar sadece en üst değil üst gelir gruplarında da genel olarak yoğunlar.

Anlamlı çıkabilen değerlerle birlikte eğitimi de katalım. Belkidaha iyi bir R kare yakalarız.

**Üst 4 gelir grubu ~ eğitim + yıl + renk**

reg4 = lm(data = dat, I(r11+r10) ~ educ + year + color)

summary(reg4)

Coefficients:

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

(Intercept) -8.305 5.895 -1.409 0.159083

educGrade 12/Std 10 4.725 6.157 0.767 0.442973

educHigher 10.567 6.157 1.716 0.086338 .

educNo schooling 14.707 6.157 2.389 0.017050 \*

educSome primary 4.546 6.157 0.738 0.460439

educSome secondary -1.490 6.157 -0.242 0.808842

year2001 2.975 4.353 0.683 0.494462

year2011 13.176 4.353 3.027 0.002523 \*\*

colorColoured 2.885 5.027 0.574 0.566184

colorIndian or Asian 18.278 5.027 3.636 0.000288 \*\*\*

colorWhite 11.311 5.027 2.250 0.024615 \*

---

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

Residual standard error: 63.98 on 1285 degrees of freedom

Multiple R-squared: 0.02763, Adjusted R-squared: 0.02007

F-statistic: 3.652 on 10 and 1285 DF, p-value: 8.02e-05

no\_schooling ısrarla anlamlı ve pozitif çıkıyor. Demekki ülke'nin en zengin iki gelir

kategorisinde de eğitimsiz bir nüfusun ağırlığı mevcut (%14 oranında bu gelir grubundalar)

Dönem etkisini yine gördük.

Hintli ve asyalılar da %18 oranında en yüksek iki gelir grubundalar.

Ve aradığımız şeyle ilgili bir işaret nihayet karşımıza çıktı.

Universite mezunları %10 seviyesinde anlamlı olacak şekilde, bu iki gelir grubunda nispeten yoğunlaşmışlar (yaklaşık %11’lik kısımları en zengin iki grupta)

Eğitimin zenginliğe ve refaha eriştirdiğini tam olarak bulamadık. Peki yoksulluktan uzaklaştırıyor mu?

**Alt 5 gelir grubu ~ eğitim + yıl + renk**

reg5 = lm(data = dat, I(r0+r1+r2+r3+r4) ~ educ + year + color)

summary(reg5)

Coefficients:

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

(Intercept) -1.535e+06 2.548e+06 -0.602 0.5470

educGrade 12/Std 10 2.802e+02 2.662e+06 0.000 0.9999

educHigher 5.248e+03 2.662e+06 0.002 0.9984

educNo schooling 4.610e+06 2.662e+06 1.732 0.0835 .

educSome primary -8.475e+01 2.662e+06 0.000 1.0000

educSome secondary -4.942e+01 2.662e+06 0.000 1.0000

year2001 2.302e+06 1.882e+06 1.223 0.2214

year2011 -2.756e+03 1.882e+06 -0.001 0.9988

colorColoured -3.680e+02 2.173e+06 0.000 0.9999

colorIndian or Asian 3.073e+06 2.173e+06 1.414 0.1575

colorWhite 3.439e+03 2.173e+06 0.002 0.9987

---

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

Residual standard error: 27660000 on 1285 degrees of freedom

Multiple R-squared: 0.007717, Adjusted R-squared: -4.868e-06

F-statistic: 0.9994 on 10 and 1285 DF, p-value: 0.4417

en alt 5 gelir grubunda eğitimsiz kesimin daha katı dağılımına ilişkin işaret görüyoruz ama R karemiz o kadar küçük ki anlamlı bir regresyon genel olarak elimizde yok.

Eğitimsizlik yoksulluğu derinleştiriyor olabilse de böyle bir ihtimalden söz edebilsek de, yüksek bir gelir elde edebilmek için üniversite eğitimini yeterli sayacak hiç bir işaret bulamadık.

**üst 4 gelir grubu ~ eğitim + yıl +cinsiyet + bölge + renk**

Belki gelir grubumuzun alt üst sınırlarını biraz daha genişletirsek yüksek öğretimin katkısına bir işaret görebiliriz.

Açıklayıcı olabilecek her şeyi regresyonumuza kattık.

reg6 = lm(data = dat, I(r11+r10+r9+r8) ~ educ + sex + state + year + color)

summary(reg6)

Coefficients:

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

(Intercept) -99.81 79.78 -1.251 0.211135

educGrade 12/Std 10 42.58 61.80 0.689 0.490916

educHigher 211.61 61.80 3.424 0.000636 \*\*\*

educNo schooling 33.86 61.80 0.548 0.583889

educSome primary 21.24 61.80 0.344 0.731131

educSome secondary -14.90 61.80 -0.241 0.809505

sexMale 15.80 35.68 0.443 0.657897

stateFree State -15.92 75.69 -0.210 0.833446

stateGauteng 41.44 75.69 0.548 0.584107

stateKwaZulu-Natal 96.80 75.69 1.279 0.201172

stateLimpopo 38.30 75.69 0.506 0.612966

stateMpumalanga -26.27 75.69 -0.347 0.728635

stateNorth West 14.68 75.69 0.194 0.846254

stateNorthern Cape 15.94 75.69 0.211 0.833204

stateWestern Cape -31.76 75.69 -0.420 0.674870

year2001 13.75 43.70 0.315 0.753131

year2011 108.66 43.70 2.487 0.013025 \*

colorColoured 42.55 50.46 0.843 0.399285

colorIndian or Asian 132.42 50.46 2.624 0.008789 \*\*

colorWhite 76.61 50.46 1.518 0.129178

---

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

Residual standard error: 642.2 on 1276 degrees of freedom

Multiple R-squared: 0.02835, Adjusted R-squared: 0.01388

F-statistic: 1.96 on 19 and 1276 DF, p-value: 0.008065

R malesef hala çok düşük. Ülkemizde nihayet gelir dağılımında üniversite mezunu olmak bir şey ifade etti. Hem de binde bir seviyesinde bile.

Demek ki evet gelirlerde bir artış yaratması muhtemelmiş.

2011de üst 4 gelir grubu mevcudu yine artmış

Indian Asian yine bu gruplarda yoğun gözüküyor.

Çıkan coefficentları düzgün yorumlamak zor çünkü interceptimiz anlamsız ve -100 çıktı.

Cinsiyet bir şey ifade etmedi. Gelirin aile içinde bölündüğünü düşünürsek, bu da normal bir durum.

Yaşanılan bölgeler ile gelir arasında da bir ilişki bulamadık

Son olarak anlamlı iki değişkenimizden (eğitim ve renk) birleşik değişken oluşturup, yıl etkisini de ayrı olarak hesaplayık regresyon kuralım. Belki farklı eğtim seviyeleri farklı etnisitelerde farklı sonuçlar doğuruyordur.

**üst 4 gelir grubu ~ eğitim \* renk + yıl**

reg7 = lm(data = dat, I(r11+r10+r9+r8) ~ educ\*color + year)

summary(reg7)

Coefficients:

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

(Intercept) -39.8611 90.5752 -0.440 0.6599

educGrade 12/Std 10 5.1537 123.0675 0.042 0.9666

educHigher 64.2566 123.0675 0.522 0.6017

educNo schooling 0.4435 123.0675 0.004 0.9971

educSome primary -0.2267 123.0675 -0.002 0.9985

educSome secondary 1.2598 123.0675 0.010 0.9918

colorColoured 6.4616 123.0675 0.053 0.9581

colorIndian or Asian 44.4506 123.0675 0.361 0.7180

colorWhite 51.6625 123.0675 0.420 0.6747

year2001 13.7469 43.5109 0.316 0.7521

year2011 108.6604 43.5109 2.497 0.0126 \*

educGrade 12/Std 10:colorColoured 2.5023 174.0438 0.014 0.9885

educHigher:colorColoured 202.9854 174.0438 1.166 0.2437

educNo schooling:colorColoured 7.5628 174.0438 0.043 0.9653

educSome primary:colorColoured -2.0112 174.0438 -0.012 0.9908

educSome secondary:colorColoured 5.4685 174.0438 0.031 0.9749

educGrade 12/Std 10:colorIndian or Asian 5.0955 174.0438 0.029 0.9766

educHigher:colorIndian or Asian 444.7728 174.0438 2.556 0.0107 \*

educNo schooling:colorIndian or Asian 115.6315 174.0438 0.664 0.5066

educSome primary:colorIndian or Asian -14.8630 174.0438 -0.085 0.9320

educSome secondary:colorIndian or Asian -22.8482 174.0438 -0.131 0.8956

educGrade 12/Std 10:colorWhite 142.1199 174.0438 0.817 0.4143

educHigher:colorWhite -58.3521 174.0438 -0.335 0.7375

educNo schooling:colorWhite 10.4598 174.0438 0.060 0.9521

educSome primary:colorWhite 102.7425 174.0438 0.590 0.5551

educSome secondary:colorWhite -47.2624 174.0438 -0.272 0.7860

---

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

Residual standard error: 639.5 on 1270 degrees of freedom

Multiple R-squared: 0.04122, Adjusted R-squared: 0.02235

F-statistic: 2.184 on 25 and 1270 DF, p-value: 0.0006755

Yıl etkisinden arındılırmış şekilde bakıyoruz.

Önceki regresyonda üni eğitiminin olumlu etkisini görmüştük.

Bu sefer bu yoğunlaşmanın temel aktörü biraz öne çıktı.

Sadece hint ve asya kökenlilerin üniversite mezunları, bu üst 4 grupta yoğunlaşmış gözüküyor.

Yani Güney afrikada etnik kökeniniz, aldığınız eğitimin, gelecekteki katkısında gayet belirleyiciymiş.

**Sonuç**

Açıklayıcı değişkenlerimizin azlığı ve doğru değişkenleri yakalayamamamız, güçlü bir regresyon kurmamıza olanan sağlamadı.

Datamız aslında oldukça geniş, bu yüzden düşük R değerleri görmemiz normal.

Güney afrikada eğitim seviyesi ile gelir arasında ikna edici bir ilişki yakalayamadık.

Yüksek öğretim, bazı etnik gruplarda anlam teşkil etse de, refah için genel bir sebep değil.

Güney afrikada gelir dağılımında, etnik kökenin anlamlı olduğuna dair işaretler gördük.

Yakaladığımız dar ilişkinin (hint kökenli mezunların üst gelir gruplarına giriyor olması) da bazı sorunları hala içinde barındıyor:

O da ilişkinin yönü.

Belki de bu etnik grup daha üst gelirlere sahip olduğundan (ki regresyonlarımız ona da işaret ediyordu) çocuklarını okutabilme fırsatı buldular.

Böyle bir sorunu çözmek içinse malesef hiçbir instrumental variable bulamadık.

Özetle, ülkemizde, eğitim ile gelir ya da maaşlar arasında iyi bir ilişki olmadığı kanaatindeyiz.