



universität
wien

Angewandte Ökonometrie 2

Sommersemester 2017

Timur Sudak A01277687

Empirisches Projekt

1. Aufgabe

Zuerst wählen wir die Variablen für unser Regressionsmodell. Wir logarithmieren unsere abhängige Variable Zigarettenkonsum pro Kopf, um dann die Koeffizienten des Regressionsmodells als Elastizität interpretieren zu können. Als unabhängige Variablen nehmen wir verfügbares Einkommen pro Kopf und den Preis pro Zigarettenpackung. Wir inkludieren Pop und Pop16 in unserem Modell nicht, weil das Einkommen und der Konsum schon pro Kopf angegeben sind. Es gibt auch keinen Grund, die Bevölkerungsanzahl im Modell hinzuzufügen, weil die Pro-Kopf-Nachfrage nicht von der Größe der Population abhängt.

Die Regression sieht aus wie folgt:

$$\ln_cons = \beta_0 + \beta_1 * \ln_income + \beta_2 * \ln_price + \varepsilon_{it}, \varepsilon_{it} = \mu_i + e_{it}$$

Wir wissen, dass der Preis durch Angebot und Nachfrage bestimmt ist. In unserem Modell haben wir den Preis, der durch die abhängige Variable Zigarettenkonsum bestimmt ist. Das heißt, dass unser Preis endogen ist, weil der Zigarettenkonsum den Preis beeinflusst.

Aber jetzt prüfen wir die Endogenität statistisch. Wir führen einen Hausman-Test durch. Wir verwerfen H_0 bei dem Test, was heißt, dass der Preis endogen ist, deshalb müssen wir

Tests of endogeneity

Ho: variables are exogenous

Robust score chi2(1) = 13.611 (p = 0.0002)
Robust regression F(1,1376) = 14.4743 (p = 0.0001)

hausman fe re

	Coefficients		(b-B) Difference	sqrt(diag(V_b-V_B)) S.E.
	(b) fe	(B) re		
ln_price	-.5396432	-.543814	.0041709	.0135497
ln_income	.3947206	.3982387	-.0035181	.0114055

b = consistent under Ho and Ha; obtained from xtivreg
B = inconsistent under Ha, efficient under Ho; obtained from xtivreg

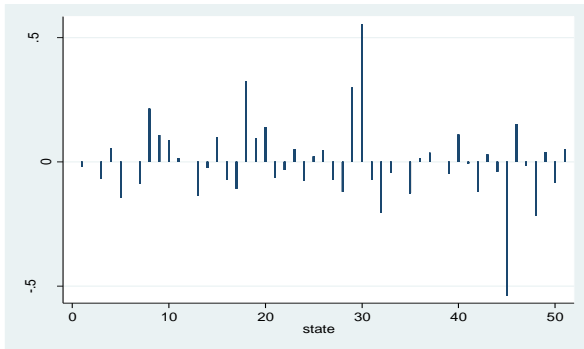
Test: Ho: difference in coefficients not systematic

chi2(2) = (b-B)'[(V_b-V_B)^(-1)](b-B)
= 0.10
Prob>chi2 = 0.9534

die Variable instrumentieren. Wir wollen jetzt eine Instrument Variable finden, die mit dem Zigarettenkonsum unkorreliert und mit dem Preis hochkorreliert ist. Die einzige Variable, die solche Eigenschaften hat, ist der Konsumentenpreisindex.

Als nächstes machen wir den Hausman-Test und testen die Annahme strikter Exogenität. Mit dem Test finden wir heraus, um welches Setting es sich bei unserem Modell handelt: Feste Effekte oder Zufällige Effekte. Der Hausman-Test zeigt uns, ob die Schätzer stark voneinander abweichen. Wir haben herausgefunden, dass die Schätzer nicht so stark voneinander abweichen und deshalb wird die Nullhypothese nicht verworfen. Das heißt, dass der GLS-Schätzer konsistent und effizient in unserem Modell ist. Der Between-Schätzer wäre ineffizient, weil er die Information von der Within Variation ignoriert. Deshalb nehmen wir GLS Schätzer, weil er sowohl Within als auch Between Variation berücksichtigt.

Jetzt wollen wir mit LM-Test(Breusch-Pagan Test) testen, ob die Varianz von individuenspezifischen Effekte σ^2_{μ} gleich Null ist. In unserem Modell verwerfen wir H_0 , es heißt, dass σ^2_{μ} bei uns ungleich Null ist.



```
. xttest0
Breusch and Pagan Lagrangian multiplier test for random effects

ln_cons[state,t] = Xb + u[state] + e[state,t]

Estimated results:
+-----+-----+
|               | Var      | sd = sqrt(Var) |
+-----+-----+
| ln_cons       | .0504448 | .2245991       |
| e             | .0149784 | .1223864       |
| u             | .0316827 | .1779964       |
+-----+-----+

Test:  Var(u) = 0
      chibar2(01) = 8933.44
      Prob > chibar2 = 0.0000
```

Dann schätzen wir μ . In unserem Modell ist der Erwartungswert von μ_i gleich Null. Die Varianz ist gleich σ^2_{μ} . Und wir haben mit dem LM-Test herausgefunden, dass σ^2_{μ} ungleich Null ist. Wir lesen aus dem Graphen

heraus, dass μ_i um dem Erwartungswert null streuen. Jetzt ist es schon offensichtlich, dass der Gepoolte-Schätzer für dieses Modell nicht geeignet ist, weil unsere μ_i verschieden sind. Es heißt, dass die Einstellung zum Rauchen in verschiedenen Städte unterschiedlich ist.

```
correlate e_fe L.e_fe L2.e_fe L3.e_fe L4.e_fe
obs=1,196)
```

	e_fe	L.e_fe	L2.e_fe	L3.e_fe	L4.e_fe
e_fe	1.0000				
L1.	0.8616	1.0000			
L2.	0.7611	0.8612	1.0000		
L3.	0.6597	0.7626	0.8621	1.0000	
L4.	0.5739	0.6589	0.7607	0.8675	1.0000

korrelieren.

2.Aufgabe

Um zu bestimmen, ob Zeit-spezifische Effekte vorliegen fügen wir jetzt die Zeit Variable in unsere Regression für jedes einzelne Jahr hinzu und führen den Test auf Signifikanz durch. Wir finden heraus, dass alle zeitspezifischen Variablen hoch signifikant sind. Aus dem Modell kann man einen Zeittrend ablesen, den man so interpretieren kann, dass der

Zigarettenkonsum mit der Zeit abnimmt. Wir sehen auch, dass alle zeitspezifischen Variablen zusammen ungleich null sind. Aber es stellt sich heraus, dass die Preisvariable in diesem

Weiter haben wir herausgestellt, dass die Fehlerterme e_{it} autokorreliert sind. Deshalb fügen wir zu unserer Regression VCE(Cluster) hinzu.

Wir nehmen diese Option, um die Korrelation innerhalb der Gruppen zuzulassen. Die Fehlerterme dürfen innerhalb der Staaten

```
corr(u_i, X) = 0 (assumed)
Wald chi2(30) = 400262.21
Prob > chi2 = 0.0000
```

(Std. Err. adjusted for 47 clusters in state)						
ln_cons	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
ln_price	-.0783108	.2848546	-0.27	0.783	-.6366155	.4799939
ln_income	.665086	.1221851	5.44	0.000	.4256076	.9045644
year						
64	-.0675382	.0057031	-11.84	0.000	-.0787161	-.0563603
65	-.099883	.0065374	-11.67	0.000	-.116655	-.0831111
66	-.1429481	.0106987	-13.36	0.000	-.1639173	-.121979
67	-.171212	.0059014	-17.29	0.000	-.1906185	-.1518055
68	-.2228521	.013708	-16.26	0.000	-.2497192	-.195985
69	-.2733651	.0168791	-16.20	0.000	-.3064475	-.2402827
70	-.3395247	.030178	-11.25	0.000	-.3986725	-.280377
71	-.3566209	.0384553	-9.27	0.000	-.4319918	-.2812499
72	-.3732122	.04	-9.33	0.000	-.4516107	-.2948138
73	-.441278	.0325966	-13.54	0.000	-.5051662	-.3773897
74	-.4610683	.0328343	-14.04	0.000	-.5254223	-.3967142
75	-.4975944	.0414461	-12.01	0.000	-.5788272	-.4163615
76	-.5203479	.0518857	-10.03	0.000	-.622042	-.4186537
77	-.5769307	.0514756	-11.21	0.000	-.677821	-.4760404
78	-.6395291	.0680279	-9.40	0.000	-.7728614	-.5061969
79	-.716958	.068361	-10.49	0.000	-.8509431	-.582973
80	-.7641709	.0740549	-10.32	0.000	-.9093158	-.6190259
81	-.8238161	.0756944	-10.88	0.000	-.9721745	-.6754578
82	-.8668707	.0978023	-8.86	0.000	-.1.05856	-.6751817
83	-.9293883	.1337252	-6.95	0.000	-1.191485	-.6672917
84	-1.018353	.159866	-6.37	0.000	-1.331684	-.7050213
85	-1.059314	.1717224	-6.17	0.000	-1.395884	-.7227445
86	-1.106868	.1846366	-5.99	0.000	-1.468749	-.7449865
87	-1.162111	.1966527	-5.91	0.000	-1.547543	-.7766787
88	-1.236186	.2108703	-5.86	0.000	-1.649484	-.8228874
89	-1.308352	.2278362	-5.74	0.000	-1.754903	-.8618013
90	-1.390891	.2491501	-5.58	0.000	-1.879216	-.9025652
91	-1.431058	.2609047	-5.48	0.000	-1.942422	-.9196946
92	-1.479733	.2887072	-5.13	0.000	-2.045589	-.9138771

Modell insignifikant ist. Die wahrscheinlichste Ursache ist, dass

die Variation vom Zigarettenkonsum von einzelnen Jahren sehr

gut aufgefangen wird. Sie erklären die Variation vom Zigarettenkonsum so gut, dass der Preis

insignifikant wird. Der Preis hat einen sehr kleinen Koeffizienten. Das heißt, dass der

Zigarettenkonsum sich sehr schwach bei der Preissteigung verändert, was nicht unseren

`chi2(29) = 5303.72`

`Prob > chi2 = 0.0000`

```
. xtivreg ln_cons ln_income (ln_price=ln_CPI) year, re vce(cluster state)

32SLS random-effects IV regression      Number of obs   =    1,380
Group variable: state                   Number of groups  =     46

R-sq:
    within = 0.5991
    between = 0.1861
    overall = 0.3078

Obs per group:
    min = 30
    avg = 30.0
    max = 30

Wald chi2(3) = 509.46
Prob > chi2 = 0.0000

corr(u_i, X) = 0 (assumed)

(Std. Err. adjusted for 47 clusters in state)
```

	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]
ln_price	-.4104533	.1723985	-2.38	0.017	[-.7483481, -.0725584]
ln_income	.7453917	.0709191	10.51	0.000	[.6063929, .8843905]
year	-.0351542	.0154741	-2.27	0.023	[-.0654828, -.0048256]
_cons	2.690722	.2178741	12.35	0.000	[2.263697, 3.117748]
sigma_u	.17875971				
sigma_e	.08251608				
rho	.82434952				(fraction of variance due to u_i)

```
Instrumented: ln_price
Instruments: ln_income year ln_CPI
```

Erwartungen entspricht. Deshalb probieren wir

die Zeit-Spezifische Variablen als gemeinsame

Variable für alle Jahre zusammen in unserem

Modell hinzufügen. Wir kriegen dann eine

signifikante Preisvariable. Unsere zeitspezifische

Variable ist ebenfalls signifikant.

`estimates table RE FD FE POOL, b(*9.4f) se stats(N)`

Variable	RE	FD	FE	POOL
ln_price	-0.4105 0.1724		-0.4316 0.1695	-2.1482 1.1173
ln_income	0.7454 0.0709		0.7429 0.0693	0.5431 0.0471
year	-0.0352 0.0155		-0.0336 0.0151	0.0924 0.0703
ln_price D1.		-1.7521 0.8580		
ln_income D1.		0.0862 0.1487		
year D1.		(omitted)		
_cons	2.6907 0.2179	0.0976 0.0648	2.6777 0.2153	1.6208 1.1353
N	1380	1334	1380	1380

3.Aufgabe.

Durch die Tabelle wird offensichtlich, dass unsere Schätzer sich

voneinander stark unterscheiden. Man kann schlussfolgern, dass

der FD und Pooled-Schätzer stark verzerrt sind. Der Fixed-Effekt

Schätzer hat eine kleinere Verzerrung als der Random-Effekt

Schätzer.

Der Between-Schätzer eliminiert die Variablen, die über die

Individuen hinweg konstant sind. Da wir aber als Instrumentvariable den Konsumentenpreisindex

(konstant in allen Staaten) genommen haben, ist der Between-Schätzer irrelevant für unser Modell.

Der RE-Schätzer ist der relevante Schätzer für unser Modell. Die statistische Validität von RE- Schätzer

haben wir schon in der Aufgabe 1 gezeigt. Die geschätzte Preiselastizität beträgt -0.41. Das heißt, wenn

der Preis um 1 Einheit steigt, dann fällt der Zigarettenkonsum um 0.41 Prozent. Der Wert für die

Schätzung, den wir gerechnet haben, unterscheidet sich nicht so stark von den Werten, die in anderen

Studien in gleichem Zeitraum gerechnet sind. Zum Beispiel in Baltagi¹(1987) beträgt die Preiselastizität

für die Periode zwischen 1956-64 -0.556 und für 1965-71 -0.433. So kann man die ökonomische und

statistische Validität unseres Schätzers begründen.

¹ Baltagi B. H., Goel R. K. Quasi-experimental price elasticities of cigarette demand and the bootlegging effect //American Journal of Agricultural Economics. – 1987. – p. 750-754.