

# Προηγμένες μέθοδοι Δεδομένων Πάνελ

Wooldridge, Κεφ 14

# Εναλλακτικές μέθοδοι (από τον διαφορισμό) για εξάλειψη σταθερής επίδρασης

- Μετασχηματισμός σταθερών επιδράσεων
- Μετασχηματισμός τυχαίων επιδράσεων

# Μετασχηματισμός σταθερών επιδράσεων

$$y_{it} = \beta_1 x_{it} + a_i + u_{it}, t = 1, 2, \dots, T.$$

Έχει διαχρονικό μέσο:

$$\bar{y}_i = \beta_1 \bar{x}_i + a_i + \bar{u}_i,$$

Οπότε αφαιρώντας:

$$\ddot{y}_{it} = \beta_1 \ddot{x}_{it} + \ddot{u}_{it}, t = 1, 2, \dots, T,$$

Και γενικότερα:

$$\ddot{y}_{it} = \beta_1 \ddot{x}_{it1} + \beta_2 \ddot{x}_{it2} + \dots + \beta_k \ddot{x}_{itk} + \ddot{u}_{it}, t = 1, 2, \dots, T,$$

Εκτίμηση σταθερών επιδράσεων / “εντός”

# Υποθέσεις

- $E(u_{it}|x_{jst})$  για όλα τα  $j,s,k$
- Επιτρέπεται  $\text{Corr}(a_i, x_{isk}) \sim 0$
- $x_{jst}$  δεν μπορεί να είναι χρονικά σταθερό
- Ομοσκεδαστικότητα
- Έλλειψη αυτοσυσχέτισης

# Ιδιότητες

- Συνηθισμένες OLS
- Βαθμοί ελευθερίας:  $N(T-1)-k$  λόγω αφαίρεσης διαχρονικού μέσου
- $R^2$  μετράει εξηγησιμότητα διαχρονικών διακυμάνσεων εντός στρώματος

# Επίδραση επαγγελματικής εκπαίδευσης σε ελαττώματα σε προϊόντα

Εξαρτημένη μεταβλητή: $\log(scrap)$	
Ανεξάρτητες μεταβλητές	
<i>d88</i>	-.080 (.109)
<i>d89</i>	-.247 (.133)
<i>grant</i>	-.252 (.151)
<i>grant</i> <sub>-1</sub>	-.422 (.210)
Παρατηρήσεις	162
Βαθμοί ελευθερίας	104
<i>R</i> -τετράγωνο	.201

54 εταιρείες

1987: 0 επιδοτήσεις

1988: 19 επιδοτήσεις

1999: 10 επιδοτήσεις

Εκτιμάται εντός

Παρουσιάζεται κλασσικά

Αν συμπεριλάβουμε τις μεταβλητές:  $\log(\text{sales})$ ,  $\log(\text{employ})$

3 εταιρείες δεν έχουν τέτοια δεδομένα

5 παρατηρήσεις χάνονται γιατί λείπουν τα δεδομένα για κάποιες εταιρείες σε κάποια χρόνια

⇒  $n=148$ , “μη ισορροπημένο πάνελ”

Επίδραση μεγαλώνει:

$\beta_{\text{grant}} =$                        $-.297$ ,  $t=-1,89$

$\beta_{\text{grant}-1} =$                        $-.536$ ,  $t=-2,39$

Εταιρείες με 1 παρατήρηση δεν έχουν επίδραση σε αποτελέσματα  
Πιθανότητα να μην υπάρχει η παρατήρηση μπορεί να συσχετίζεται με επίδραση  $a_i$

# Απόδοση στην μόρφωση διαχρονικά

545 άνδρες

1980-7

$$\log(\text{educ\_it}) = \beta_0 \text{educ}^2 + \beta_1 \text{dmarrried} + \beta_2 \text{dsynd} + \beta_3 \text{d81} + \dots + \beta_9 \text{d87} + \beta_{10} \text{d81} * \text{educ} + \dots + \beta_{16} \text{d87} * \text{educ}$$

$\beta_{16} = 0.03$  (t-stat = 2.48) => Απόδοση 3% μεγαλύτερη από το 1980

Απόδοση 1980 είναι μέρος του  $a_i$

Ιδιαίτερος τρόπος χρήσης στατικής πληροφορίας για το στρώμα

Εκτίμηση σταθερών επιδράσεων [παραληφθείσες μεταβλητές]:

$$\hat{a}_i = \bar{y}_i - \hat{\beta}_1 \bar{x}_{i1} - \dots - \hat{\beta}_k \bar{x}_{ik}, \quad i = 1, \dots, N,$$

Αμερόληπτη, συνεπής  $T$ , όχι  $N$

Ίδια αποτελέσματα με παλινδρόμηση ψευδομεταβλητών

- $d_i$  για κάθε  $i$
- Δύσκολο αν  $N$  μεγάλο, ανακριβές αν  $T$  μικρό
- $R^2$  τώρα συμπεριλαμβάνει ερμηνευτικότητα ψευδομεταβλητών



# Σύγκριση εκτιμήσεων από μετασχηματισμούς σταθερών επιδράσεων και διαφορισμού

T=1: Δ Μη εφαρμόσιμος, ΣΕ όλες μεταβλητές 0

T=2: Ίδια αποτελέσματα

T>=3: Αμερόληπτοι και συνεπής (για N)

Αποτελεσματικότητα:

Εξαρτάται από την αυτοσυσέτιση του  $u_{it}$  [μη παρατηρήσιμο] σε σχέση με του  $\Delta u_{it}$  [παρατηρήσιμο]

- Καλό να χρησιμοποιούνται και οι δύο μετασχηματισμοί

Αν  $T \gg 0$  χάρη σε ασυμπτωτική θεωρία χρονοσειρών:

- Δ προτιμότερος αν κάποιες μεταβλητές είναι I(1)
- ΣΕ προτιμότερες αν δεν ισχύει η ΑΕ

# Μετασχηματισμός τυχαίων επιδράσεων

Αν  $\text{Cov}(x_{itj}, a_i) = 0, t = 1, 2, \dots, T. j = 1, 2, \dots, k.$

δεν απομακρύνουμε απαραίτητη σταθερή επίδραση στο υπόδειγμα:

$$y_{it} = \beta_0 + \beta_1 x_{it1} + \dots + \beta_k x_{itk} + a_i + u_{it},$$

ή

$$y_{it} = \beta_0 + \beta_1 x_{it1} + \dots + \beta_k x_{itk} + v_{it}. \quad v_{it} = a_i + u_{it}.$$

Γιατί χάνουμε βαθμούς ελευθερίας (αποτελεσματικότητα)

Τότε όμως έχουμε αυτοσυσχέτιση:

$$\text{Corr}(v_{it}, v_{is}) = \sigma_a^2 / (\sigma_a^2 + \sigma_u^2), t \neq s,$$

Αφαιρόντας την αυτοσυσχέτιση με διαδικασία αντίστοιχη με ημιδιαφορισμό:

$$y_{it} - \lambda \bar{y}_i = \beta_0(1 - \lambda) + \beta_1(x_{it1} - \lambda \bar{x}_{i1}) + \dots \\ + \beta_k(x_{itk} - \lambda \bar{x}_{ik}) + (v_{it} - \lambda \bar{v}_i),$$

$$\lambda = 1 - [\sigma_u^2 / (\sigma_u^2 + T\sigma_a^2)]^{1/2},$$

Χρησιμοποιείται εκτίμηση για το  $\lambda$

Επιτρέπονται μεταβλητές σταθερές στον χρόνο [αρκεί να είναι **ασυσχέτιστες** με ΣΕ]

Παρατηρόντας ότι:

$$v_{it} - \lambda \bar{v}_i = (1 - \lambda)a_i + u_{it} - \lambda \bar{u}_i$$

Αν υπάρχει συσχέτιση μεταξύ  $a_i$  και  $x_i$

Για  $\lambda \rightarrow 1$  ασήμαντη [ΣΕ]

Για  $\lambda \rightarrow 0$  σημαντική [ομαδοποιημένα στοιχεία]

Σύγκριση αποτελεσμάτων από μετασχηματισμούς ΤΕ και ΣΕ  
δίνει έλεγχο για έλλειψη συσχέτισης

# Απόδοση στην μόρφωση διαχρονικά (ξανά)

545 άνδρες

## ΣΥΓΚΡΙΣΗ ΕΚΤΙΜΗΣΕΩΝ

1980-7

$\lambda^{\wedge}=0.643$

Εξαρτημένη μεταβλητή: $\log(wage)$			
Ανεξάρτητες μεταβλητές	Ομαδοποιημένα ελάχιστα τετράγωνα	Τυχαίες επιδράσεις	Σταθερές επιδράσεις
<i>educ</i>	.091 (.005)	.092 (.011)	————
<i>black</i>	-.139 (.024)	-.139 (.048)	————
<i>hispan</i>	.016 (.021)	.022 (.043)	————
<i>exper</i>	.067 (.014)	.106 (.015)	————
<i>exper<sup>2</sup></i>	-.0024 (.0008)	-.0047 (.0007)	-.0052 (.0007)
<i>married</i>	.108 (.016)	.064 (.017)	.047 (.018)
<i>union</i>	.182 (.017)	.106 (.018)	.080 (.019)

Συμπεριλαμβάνονται και ψευδομεταβλητές έτους

# Άλλες χρήσεις μετασχηματισμών πάνελ

Όταν σε συστάδες [όχι βασισμένες στον χρόνο] του δείγματος αναμένονται απαραίτητες επιδράσεις:

π.χ. Εισόδημα κόρης σε οικογένειες με τουλάχιστον δύο κόρες

$$\log(\text{incneeds}_{fs}) = \beta_0 + \delta_0 \text{sister2}_s + \beta_1 \text{teenbrth}_{fs} + \beta_2 \text{age}_{fs} + \text{άλλοι παράγοντες} + a_f + u_{fs},$$

n=129, 1982,  $\beta^1 = -.33$ , πολύ σ.σ.

$$\Delta \log(\text{incneeds}) = \delta_0 + \beta_1 \Delta \text{teenbrth} + \beta_2 \Delta \text{age} + \dots + \Delta u.$$

n=129, 1982,  $\beta^1 = -.08$ , όχι σ.σ.

Π.χ.2:

149 ομοζυγωτικά δίδυμα

age, gender, race είναι ίδια οπότε παλινδρόμηση:

$\Delta \log(\text{earn}) = \beta_0 \Delta \text{educ}$

$\beta_0 = 0.092, t = 3.83$

Μπορούν να χρησιμοποιηθούν και μετασχηματισμοί Τ.Ε.