Analisi Multilivello con M*plus*

Seminario 3 – 23 Maggio 2016

Valerio Ghezzi

Dipartimento di Psicologia

Sapienza – Università di Roma

ARGOMENTI

- **Centratura dei predittori nel dettaglio e moderazione cross-livello;**

- **Modelli con variabili osservate con più di un outcome;**

- **La path analysis a 2 livelli.**

CENTRATURA DEI PREDITTORI NEL DETTAGLIO

Il tema della "centratura"

Il tema della centratura dei predittori nella regressione multilivello è un **tema estremamente e complesso e delicato**, che in alcuni casi può determinare una interpretazione dei risultati profondamente diversa a seconda delle scelte che si fanno (Bryk & Raudenbush, 2002; Hox, 2010);

Sebbene esistano diverse strategie di centratura dei predittori (Plewis, 1989), **generalmente nella regressione multilivello i predittori vengono centrati rispetto alla media**, ossia al valore osservato nel predittore ottenuto da un soggetto viene sottratta una media;

Tale media può essere quella del campione generale (*grand mean centering, CGM*) oppure la media del proprio gruppo di appartenenza (*group mean centering* o *centering within context, CWC*).

Grand Mean Centering - CGM

1) CGM **non influenza l'ordine di rango di un soggetto** rispetto al predittore nel campione generale;

2) Il predittore di livello 1 può essere visto come composito che contiene entrambe le componenti (i.e., la variabilità del predittore *within* è correlato con la sua controparte *between*);

3) L'intercetta è interpretabile come **la media dell'outcome per un qualsiasi cluster *j*, meno un aggiustamento che dipende**:

a) dalla slope (effetto fisso);

b) dalla **deviazione tra la media del predittore nel cluster *j* e *la* media generale del campione** (*grand mean*).

Grand Mean Centering - CGM

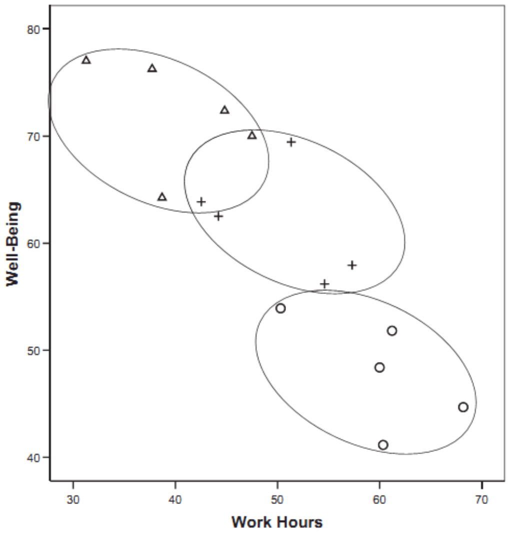
LIVELLO 1 $$WELLBEING_{ij} = \beta_{0j} + \beta_{1j}(HOURS_{ij} - \bar{x}_{HOURS}) + r_{ij}$$

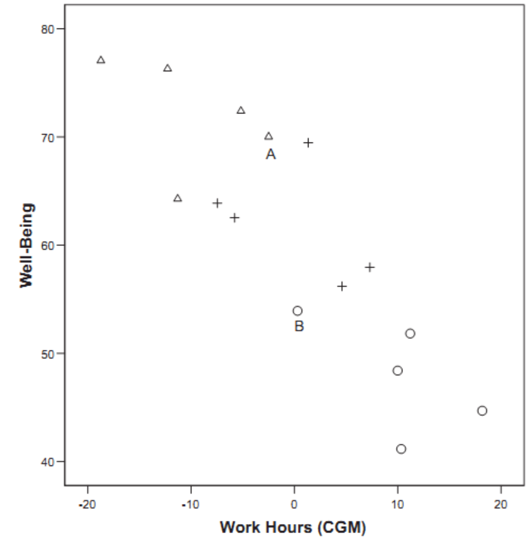
LIVELLO 2 $$\beta_{0j} = \mu_{WELLBEING_j} - \beta_{1j}(\bar{x}_{HOURS_j} - \bar{x}_{HOURS})$$

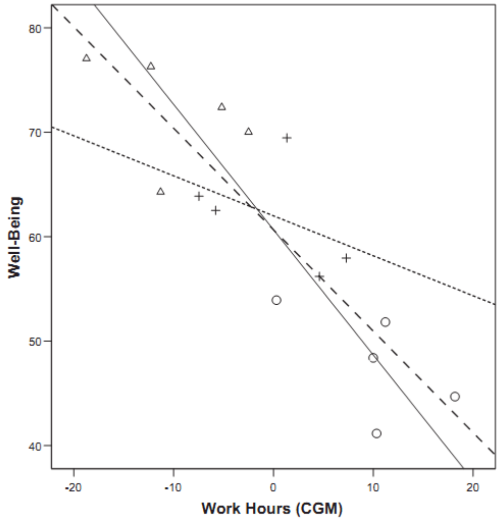
In questo caso, l'intercetta ($\mu_{wellbeingj}$ o γ_{00}) è **la media per il soggetto *i* nell'outcome all'interno del cluster *j* "aggiustata" per la slope (fissa) e la deviazione del suo gruppo di appartenenza dalla media generale;**

La CGM cambia l'interpretazione delle componenti della varianza: la variabilità intorno all'intercetta (nei modelli RI rappresentata con u_{0j}) quantifica la variabilità intorno a tale media "aggiustata" (cioè la variabilità intorno a $\mu_{wellbeingj}$ o γ_{00} dopo aver controllato per l'effetto del predittore). La variabilità intorno alle slope (intorno a B_{1j} o u_{ij}) tende quindi a 0, perché u_{0j} a u_{1j} sono termini correlati.

Grand Mean Centering - CGM







- - - - Regressione OLS

.... Regressione "Within"(Pooled)

——— Regressione "Between"

Group Mean Centering - CGC

1) CWC **influenza l'ordine di rango di un soggetto rispetto al predittore, poiché è espresso in termini di deviazione dal proprio gruppo di appartenenza**;

2) Il predittore di livello 1 viene considerato "al netto" della sua *nestedness* (i.e., il predittore considerato al livello between NON è correlato con la sua controparte between);

3) L'intercetta è interpretabile come la media dell'outcome per un qualsiasi cluster *j* ***quando il valore del predittore al livello individuale è indentico alla media del gruppo di appartenenza***.

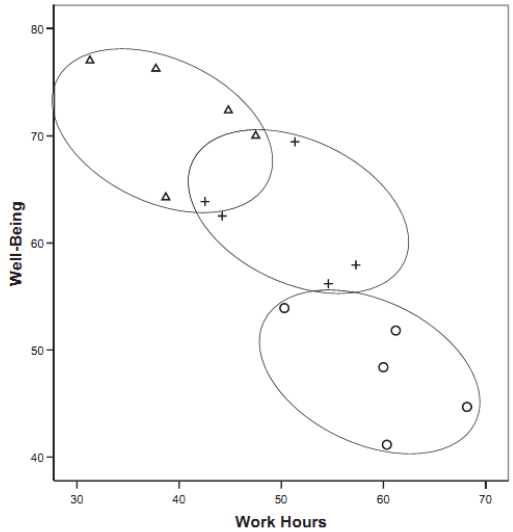
Group Mean Centering - CWC

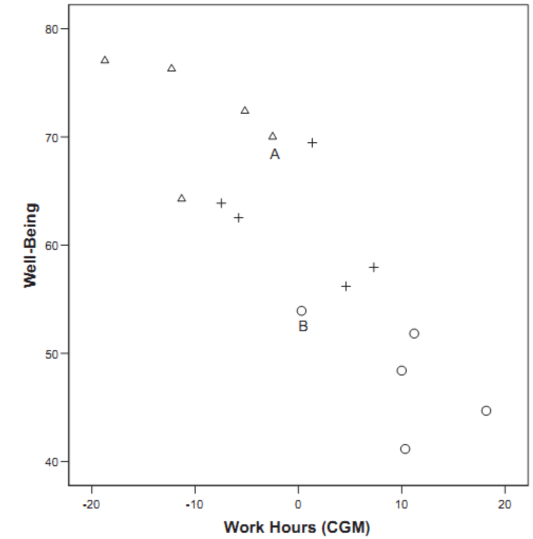
LIVELLO 1 $$WELLBEING_{ij} = \beta_{0j} + \beta_{1j}(HOURS_{ij} - \bar{x}_{HOURS_j}) + r_{ij}$$

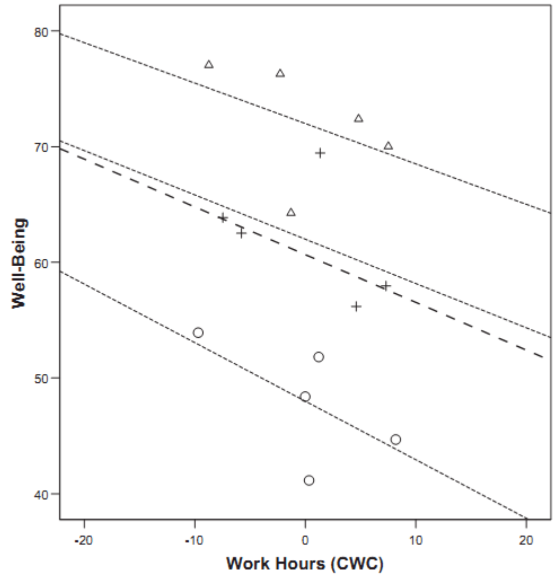
In questo caso, l'intercetta è **la media nell'outcome per il soggetto *i* quando la sua media è identica a quella del proprio cluster *j*;**

La CGM cambia l'interpretazione delle componenti della varianza: la variabilità intorno all'intercetta (nei modelli RI rappresentata con u_{0j}) quantifica la variabilità intorno a tale media NON aggiustata per lo scarto dalla media generale. La variabilità intorno alle slope (intorno a B_{1j} o u_{ij}) tende quindi a variare di più rispetto a CGM, perché u_{0j} a u_{1j} NON sono termini correlati.

Grand Mean Centering - CGM







- - - - Regressione OLS

.... Regressioni "Within" (pooled)

CGM o CWC? (Enders & Tofighi, 2007, p. 136)

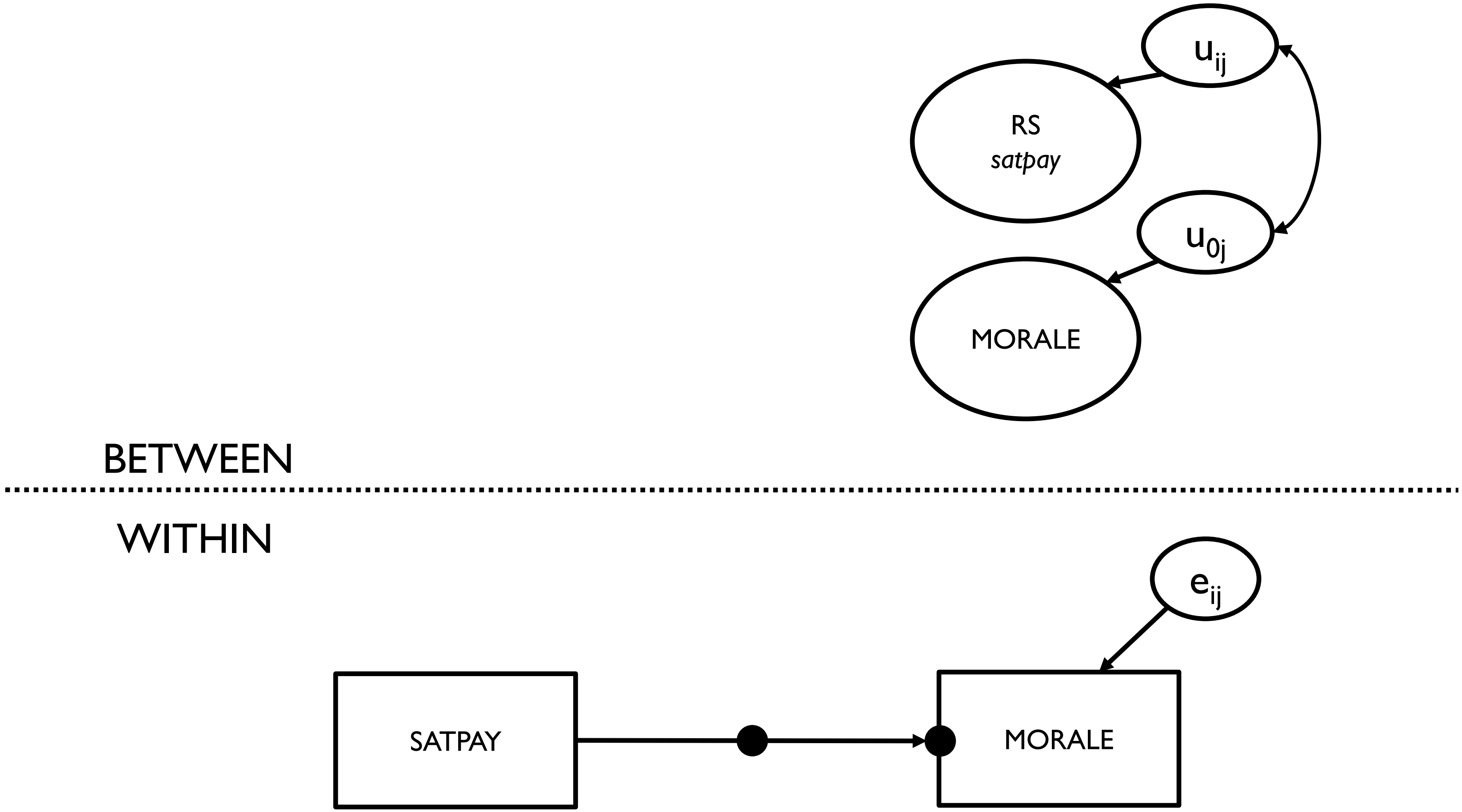
a) CWC è appropriata quando si è particolarmente interessati ai predittori di livello 1;

b) CGM è appropriata quando si è interessati principalmente ai predittori di livello 2 e vuole "controllare" gli effetti di alcune variabili di controllo al livello 1;

c) Usare CGM o CWC è relativamente intercambiabile nel caso in cui si è interessati agli effetti di predittori locati su entrambi i livelli;

d) Usare CGM nel caso di interazioni tra variabili di livello 2 e CWC nel caso di interazioni cross-livello o interazioni tra variabili di livello 1.

EG2 – I&S RANDOM



EG1 – M1CGM.inp

```
TITLE:      I&S random
DATA:       FILE IS ch3new.dat;
                 Format is 5f8.0, 3f8.2;
VARIABLE:        Names are deptid morale satpay female white pctbelow lev1wt
        lev2wt;
                 Usevariables are deptid morale satpay ;
                 Cluster is deptid;
                 Within = satpay;

Define:          Center satpay  (grandmean);

ANALYSIS:        Type= Twolevel random;

Model:           %Between%
                   morale; S ;

                   %Within%
                 S | morale on satpay;

OUTPUT:            SAMPSTAT TECH1;
```

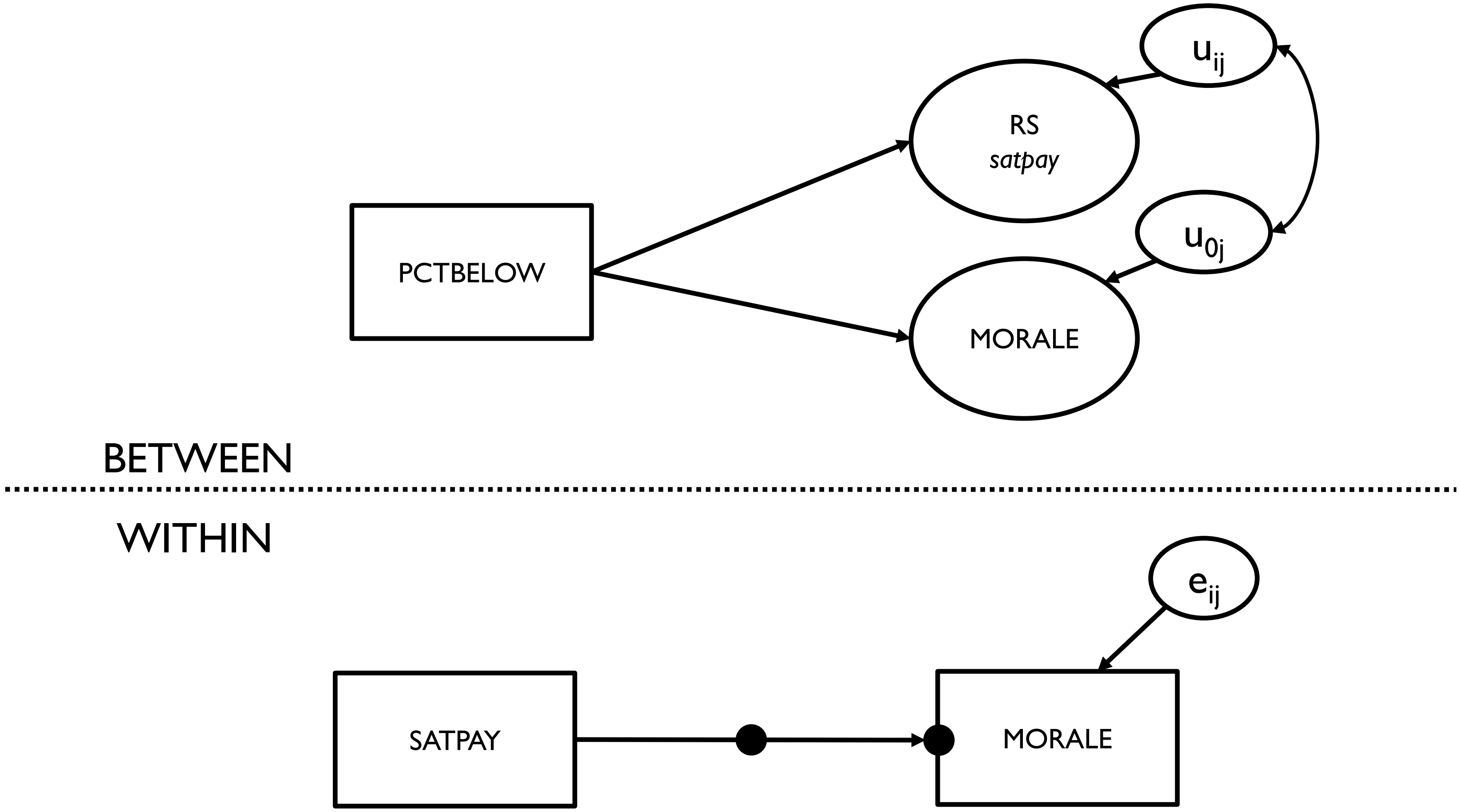
EG1 – M1CGM.inp

```
TITLE:      I&S random
DATA:       FILE IS ch3new.dat;
                  Format is 5f8.0, 3f8.2;
VARIABLE:         Names are deptid morale satpay female white pctbelow lev1wt
            lev2wt;
                  Usevariables are deptid morale satpay ;
                  Cluster is deptid;
                  Within = satpay;

Define:           Center satpay  (groupmean);

ANALYSIS:         Type= Twolevel random;

Model:            %Between%
                  morale; S ;

                  %Within%
                  S | morale on satpay;

OUTPUT:           SAMPSTAT TECH1;
```

Riportare i risultati (I&S RANDOM)

Model:	EG1 - CGM	EG1 - CWC	EG2 - CGM	EG2 - CWC
Fixed Part	**Coeff. (s.e.)**	**Coeff. (s.e.)**	**Coeff. (s.e.)**	**Coeff. (s.e.)**
γ_{oo}	26.42 (.11)	26.43 (.19)		
$B_{1J}satpay_{1J}$	1.21 (.01)	1.20 (.01)		
$\gamma_{01}pctbelow_J$				
$\gamma_{i1}pctbelow_J$				
Random Part				
u_{oj}	**1.85 (.23)**	**5.64 (.61)**		
u_{ij}	**.008 (.006)**	**.009 (.003)**		
e_{ij}	17.64 (.29)	17.63 (.30)		
Deviance (NEP)	75692.23(5)	75858.18(5)		

EG2 – I&S AS OUTCOMES



EG2 – M2CGM.inp

```
TITLE:      Model 4 - intercept-and-slope as outcomes
DATA:       FILE IS ch3new.dat;
                 Format is 5f8.0, 3f8.2;
VARIABLE:        Names are deptid morale satpay female white pctbelow
        lev1wt lev2wt;
                 Usevariables are deptid morale satpay  pctbelow;
                 Cluster is deptid;
                 Between = pctbelow;
                 Within = satpay;

Define:          Center satpay  (grandmean)
                 pctbelow (grandmean);

ANALYSIS:        Type= Twolevel random;

Model:           %Between%
                    morale S on pctbelow;
                 morale with s;

                    %Within%
                  S | morale on satpay;

OUTPUT:             SAMPSTAT TECH1;
```

EG2 – M2CWC.inp

```
TITLE:      Model 4 - intercept-and-slope as outcomes
 DATA:      FILE IS ch3new.dat;
                    Format is 5f8.0, 3f8.2;
 VARIABLE:          Names are deptid morale satpay female white pctbelow
          lev1wt lev2wt;
                    Usevariables are deptid morale satpay  pctbelow;
                    Cluster is deptid;
                    Between = pctbelow;
                    Within = satpay;

 Define:            Center satpay  (groupmean)
                    pctbelow (grandmean);

 ANALYSIS:          Type= Twolevel random;

 Model:             %Between%
                      morale S on pctbelow;
                    morale with s;

                      %Within%
                     S | morale on satpay;

 OUTPUT:              SAMPSTAT TECH1;
```

Riportare i risultati (I&S as outcomes)

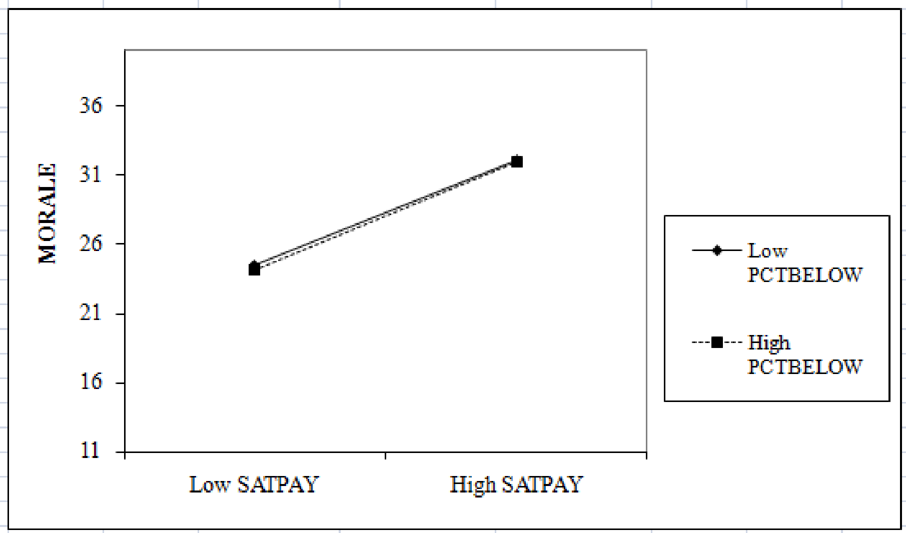
Model:	EG1 - CGM	EG1 - CWC	EG2 - CGM	EG2 - CWC
Fixed Part	**Coeff. (s.e.)**	**Coeff. (s.e.)**	**Coeff. (s.e.)**	**Coeff. (s.e.)**
γ_{00}	26.42 (.11)	26.43 (.19)	27.19 (.23)	28.31 (.36)
$B_{1J}satpay_{1J}$	1.21 (.01)	1.20 (.01)	1.18 (.03)	1.16 (.03)
$\gamma_{01}pctbelow_J$			**-.03 (.007)**	**-.07 (.01)**
$\gamma_{i1}pctbelow_J$			.001 (.001)	.001 (.001)
Random Part				
u_{oj}	1.85 (.23)	5.64 (.61)	**1.67 (.21)**	**4.46 (.52)**
u_{ij}	.008 (.006)	.009 (.003)	.008 (.011)	.008 (.003)
e_{ij}	17.64 (.29)	17.63 (.30)	17.64 (.30)	17.63 (.30)
Deviance (NEP)	75692.23(5)	75858.18(5)	75674.18(8)	75819.01(8)

Plottare l'interazione cross-level (EG2 – CWC)

www.jeremydawson.co.uk/slopes.htm

```
UNIVARIATE HIGHER-ORDER MOMENT DESCRIPTIVE STATISTICS

     Variable/         Mean/     Skewness/   Minimum/ % with                Percentiles
    Sample Size      Variance    Kurtosis    Maximum  Min/Max      20%/60%    40%/80%    Median

 MORALE               26.360      -0.195      11.000    0.64%      21.000     25.000     27.000
       13189.000      38.206      -0.420      40.000    0.26%      28.000     32.000
 SATPAY                0.000      -0.292     -10.915    0.01%      -2.874     -0.639      0.252
       13189.000      10.510      -0.403       8.015    0.03%       1.123      2.905
 PCTBELOW             27.091       1.065       2.400    0.61%      13.700     22.600     26.970
         165.000     242.324       1.595      82.900    0.61%      26.970     37.600
```

Enter information from your regression in the shaded cells	
Variable names:	
Name of independent variable:	SATPAY
Name of moderator:	PCTBELOW
Unstandardised Regression Coefficients:	
Independent variable:	1,16
Moderator:	-0,007
Interaction:	0,001
Intercept / Constant:	28,31
Means / SDs of variables:	
Mean of independent variable:	0
SD of independent variable:	3,24
Mean of moderator:	27,091
SD of moderator:	15,57

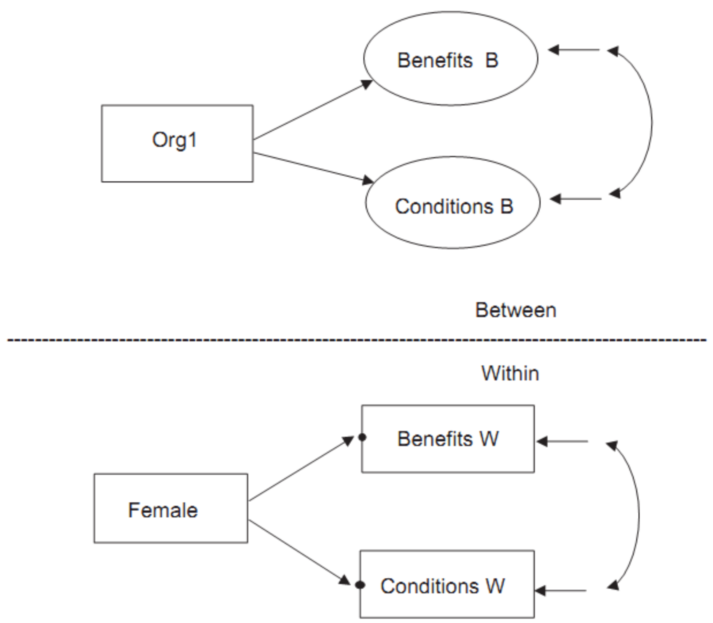


LA PATH ANALYSIS A 2 LIVELLI

2 outcomes – 1 predittore LEV1 & LEV2

12.445 impiegati *nested* in
160 organizzazioni produttive e di servizi

FEMALE (Within) = 0 maschi, 1 femmine
ORG1 (Between) = percezione di contesto, riscalata da 0 a 1
BENEFITS & CONDITIONS = Misurati su 10 punti (1=minimo, 10 massimo).



Modello matematico di base

$$Y_{Wij} = \eta_j + B_W Y_{Wij} + \zeta_{Wij}$$

$$\eta_j = \alpha_B + B_B Y_{Bj} + \zeta_{Bj}$$

EG2 – PATH1.inp

```
TITLE:          Model 1: Two-level model with multivariate outcomes;
DATA:        FILE IS C:\mplus1\ch4mv.dat;
                Format is 7f8.0,6f8.2;

VARIABLE:       Names are orgid female white satpay morale org1 org2
        benefit cond zresour zproduct lev1wt lev2wt;
                Usevariables are benefit cond female org1;
                within =  female;
                between = org1;
                CLUSTER IS orgid;
ANALYSIS:       TYPE = twolevel;
                Estimator is mlr;

Model:  %Between%
        benefit cond on org1;
        benefit with cond;
        %Within%
        benefit cond on female;
        benefit with cond;
OUTPUT:         SAMPSTAT Standardized TECH1;
```

TECH1 - WITHIN

BETA

	BENEFIT	COND	FEMALE	ORG1
BENEFIT	0	0	1	0
COND	0	0	2	0
FEMALE	0	0	0	0
ORG1	0	0	0	0

PSI

	BENEFIT	COND	FEMALE	ORG1
BENEFIT	3			
COND	4	5		
FEMALE	0	0	0	
ORG1	0	0	0	0

TECH1 - BETWEEN

ALPHA

	BENEFIT	COND	FEMALE	ORG1
1	6	7	0	0

BETA

	BENEFIT	COND	FEMALE	ORG1
BENEFIT	0	0	0	8
COND	0	0	0	9
FEMALE	0	0	0	0
ORG1	0	0	0	0

PSI

	BENEFIT	COND	FEMALE	ORG1
BENEFIT	10			
COND	11	12		
FEMALE	0	0	0	
ORG1	0	0	0	0

OUTPUT - WITHIN

STDYX Standardization

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
Within Level				
BENEFIT ON				
FEMALE	0.237	0.009	27.603	0.000
COND ON				
FEMALE	0.245	0.009	26.231	0.000
BENEFIT WITH				
COND	0.678	0.007	93.127	0.000
Residual Variances				
BENEFIT	0.944	0.004	231.782	0.000
COND	0.940	0.005	205.866	0.000

1) Le femmine hanno punteggi più alti in benefit e conditions;
2) A livello *within*, benefit e conditions sono piuttosto correlate.

OUTPUT - BETWEEN

STDYX Standardization

```
Between Level

 BENEFIT    ON
    ORG1               0.225      0.089      2.544      0.011

 COND       ON
    ORG1               0.276      0.085      3.237      0.001

 BENEFIT  WITH
    COND               0.927      0.015     63.519      0.000

 Intercepts
    BENEFIT            8.835      0.605     14.612      0.000
    COND              10.352      0.678     15.259      0.000

 Residual Variances
    BENEFIT            0.949      0.040     23.760      0.000
    COND               0.924      0.047     19.612      0.000
```

1) L'incremento di una unità standard in ORG1 determina l'aumento di .22 e .28 unità standard, rispettivamente, in benefits & conditions.
2) A livello *between*, benefit e conditions sono sostanzialmente indistinguibili.

VARIANZA SPIEGATA

R-SQUARE

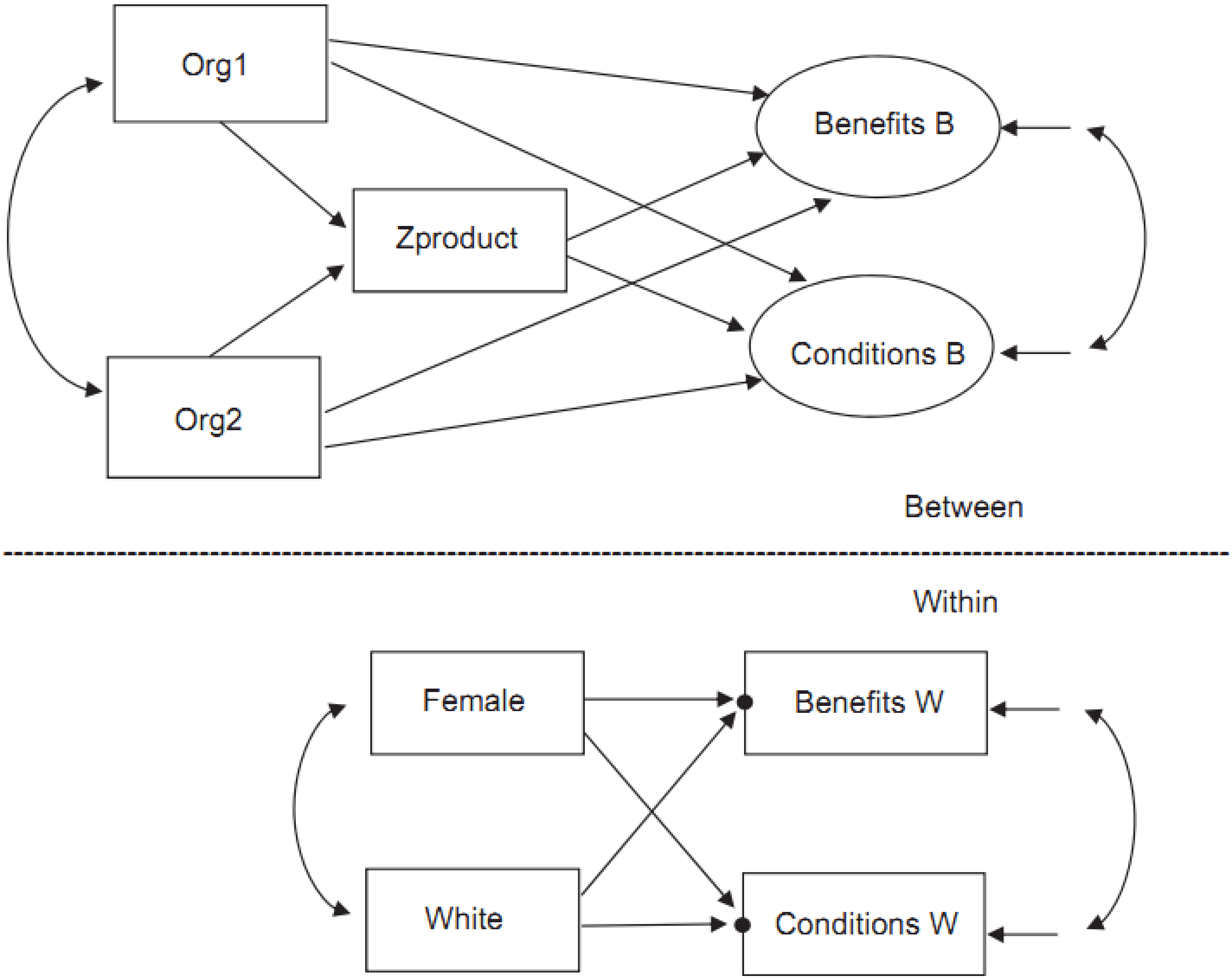
Within Level

Observed Variable	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
BENEFIT	0.056	0.004	13.801	0.000
COND	0.060	0.005	13.115	0.000

Between Level

Observed Variable	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
BENEFIT	0.051	0.040	1.272	0.203
COND	0.076	0.047	1.618	0.106

PATH ANALYSIS



EG3 – PATH2.inp

```
TITLE:    Model 2: Multivariate outcomes and mediating variable;
DATA:     FILE        IS C:\mplus1\ch4mv.dat;
          Format is 7f8.0,6f8.2;

VARIABLE:   Names are orgid female white satpay morale org1
          org2 benefit cond resour zproduct lev1wt lev2wt;
          Usevariables are benefit cond female white org1 org2 zproduct;
           within =  female white;
          between = org1 org2 zproduct;
           CLUSTER IS orgid;

ANALYSIS:   TYPE = twolevel;
           Estimator is mlr;

Model:     %Between%
           benefit cond on org1 org2;
          zproduct on org1 org2;
          benefit cond on zproduct;
          benefit with cond;
           %Within%
           benefit cond on female white;
          benefit with cond;
OUTPUT: SAMPSTAT Standardized TECH1;
```

TECH1 - WITHIN

BETA

	ZPRODUCT	BENEFIT	COND	FEMALE	WHITE
ZPRODUCT	0	0	0	0	0
BENEFIT	0	0	0	1	2
COND	0	0	0	3	4
FEMALE	0	0	0	0	0
WHITE	0	0	0	0	0
ORG1	0	0	0	0	0
ORG2	0	0	0	0	0

PSI

	ZPRODUCT	BENEFIT	COND	FEMALE	WHITE
ZPRODUCT	0				
BENEFIT	0	5			
COND	0	6	7		
FEMALE	0	0	0	0	
WHITE	0	0	0	0	0
ORG1	0	0	0	0	0
ORG2	0	0	0	0	0

TECH1 - BETWEEN

ALPHA

	ZPRODUCT	BENEFIT	COND	FEMALE	WHITE
1	8	9	10	0	0

BETA

	ZPRODUCT	BENEFIT	COND	FEMALE	WHITE
ZPRODUCT	0	0	0	0	0
BENEFIT	13	0	0	0	0
COND	16	0	0	0	0
FEMALE	0	0	0	0	0
WHITE	0	0	0	0	0
ORG1	0	0	0	0	0
ORG2	0	0	0	0	0

BETA

	ORG1	ORG2
ZPRODUCT	11	12
BENEFIT	14	15
COND	17	18
FEMALE	0	0
WHITE	0	0
ORG1	0	0
ORG2	0	0

PSI

	ZPRODUCT	BENEFIT	COND	FEMALE	WHITE
ZPRODUCT	19				
BENEFIT	0	20			
COND	0	21	22		
FEMALE	0	0	0	0	
WHITE	0	0	0	0	0
ORG1	0	0	0	0	0
ORG2	0	0	0	0	0

OUTPUT - WITHIN

STDYX Standardization

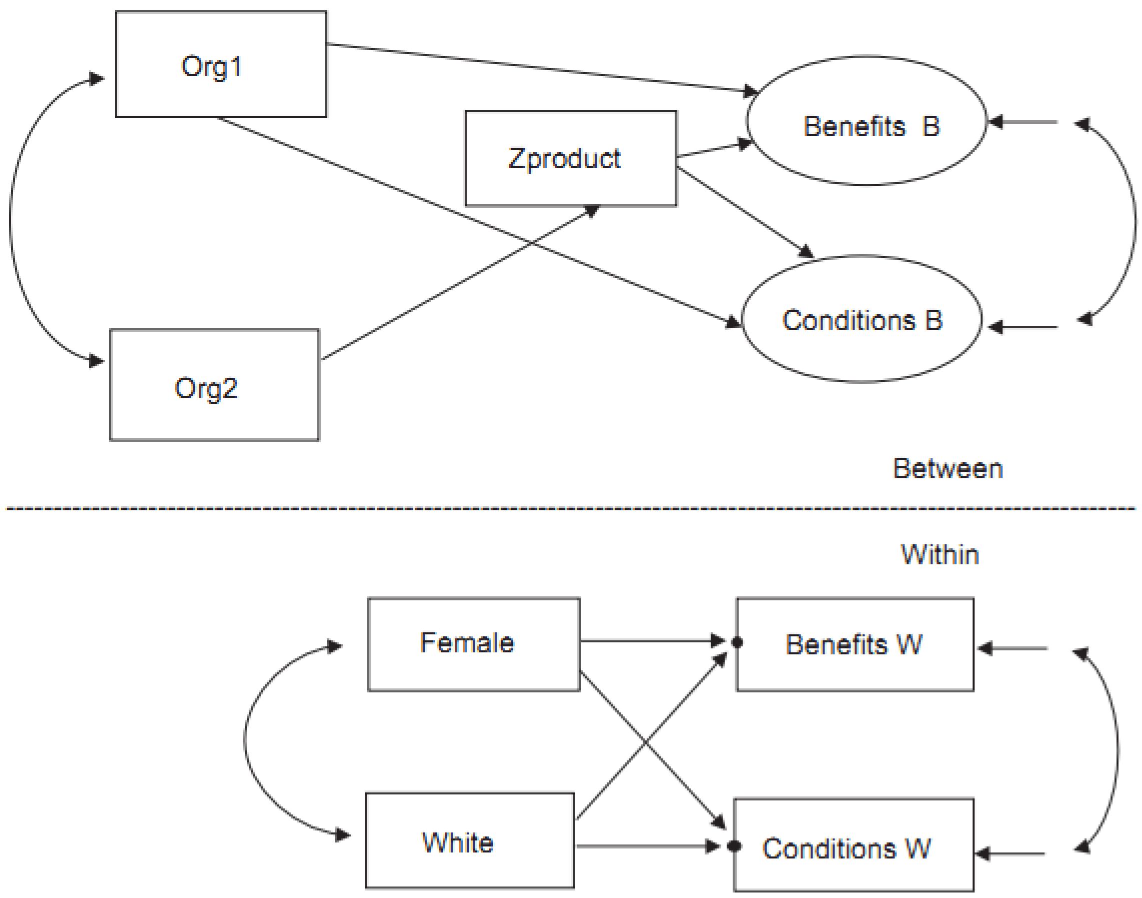
```
                                                    Two-Tailed
                    Estimate       S.E.  Est./S.E.    P-Value

Within Level

 BENEFIT    ON
    FEMALE             0.237      0.009     27.795      0.000
    WHITE              0.076      0.010      7.900      0.000

 COND       ON
    FEMALE             0.244      0.009     26.483      0.000
    WHITE              0.085      0.010      8.913      0.000

 BENEFIT  WITH
    COND               0.676      0.007     92.536      0.000

 Residual Variances
    BENEFIT            0.938      0.004    213.208      0.000
    COND               0.933      0.005    195.755      0.000
```

OUTPUT - BETWEEN

STDYX Standardization

```
Between Level

 BENEFIT    ON
    ORG1               0.241      0.064      3.776      0.000
    ORG2               0.028      0.064      0.431      0.666
    ZPRODUCT           0.651      0.055     11.885      0.000

 COND       ON
    ORG1               0.296      0.060      4.933      0.000
    ORG2               0.022      0.065      0.338      0.735
    ZPRODUCT           0.679      0.051     13.356      0.000

 ZPRODUCT   ON
    ORG1              -0.045      0.088     -0.510      0.610
    ORG2              -0.180      0.075     -2.410      0.016

 BENEFIT  WITH
    COND               0.859      0.029     30.091      0.000

 Intercepts
    ZPRODUCT           0.113      0.099      1.148      0.251
    BENEFIT            8.790      0.614     14.325      0.000
    COND              10.404      0.697     14.923      0.000

 Residual Variances
    ZPRODUCT           0.970      0.025     38.921      0.000
    BENEFIT            0.526      0.058      9.055      0.000
    COND               0.458      0.055      8.257      0.000
```

Rimuovere i path non significativi (PATH3.INP)



Confrontare il fit

Modello PATH2.inp

Chi-Square Test of Model Fit

Value	0.002*
Degrees of Freedom	0
P-Value	0.0000
Scaling Correction Factor for MLR	1.0000

Modello PATH3.inp

Chi-Square Test of Model Fit

Value	0.471*
Degrees of Freedom	3
P-Value	0.9252
Scaling Correction Factor for MLR	1.0359

*** The chi-square value for MLM, MLMV, MLR, ULSMV, WLSM and WLSMV cannot be used for chi-square difference testing in the regular way. MLM, MLR and WLSM chi-square difference testing is described on the Mplus website. MLMV, WLSMV, and ULSMV difference testing is done using the DIFFTEST option.**

Satorra-Bentler Scaled Chi Square	df	*p*-value
0,469069312	3	0,926

Analisi Multilivello con M*plus*

Seminario 3 – 23 Maggio 2016

Valerio Ghezzi

Dipartimento di Psicologia

Sapienza – Università di Roma