Some Applications of Panel Data in Health Economics

Patrick GAGLIARDINI
Università della Svizzera Italiana

June 2015

M. Kerkhofs and M. Lindeboom (1997): “Age Related Health Dynamics and Changes in Labour Market Status”, *Health Economics*, 6, 407-423.

Goal: Study the effect of age and labour market status on changes in health

Model:

$$h_{it} = \alpha_0 + \alpha_1 L_{it} + \alpha_2 S_{it} + \beta' x_{it} + \gamma_i + \varepsilon_{it}$$

h_{it}=health status (a score)
L_{it}=labour market status ($= 1$ if individual works)
S_{it}=health shock ($= 1$ if a disease, accident ... in $(t-1, t)$)
x_{it}=characteristics (age, education, marital status, ...)

Data: Dutch CERRA panel dataset, two waves: 1993, 1995
4727 individuals in 1993, about 70% responding in 1995

Table 1. First stage estimates: linear regression of change of HSCL score

	HSCL on 7-point scale		Total HSCL score	
Variable	Estimate	*t*-value[a]	Estimate	*t*-value[a]
Constant	–0.6245	2.27 (2.35)	–6.0405	2.42 (2.49)
Age in 1993[b]	0.0146	2.83 (2.89)	0.1336	2.84 (2.82)
Dummy female	–0.1052	1.48 (1.46)	–0.8357	1.30 (1.19)
First differences of:				
Dummy partner	–0.2907	1.52 (1.43)	–2.3907	1.38 (1.19)
Dummy work	0.2543	1.87 (1.77)	1.2567	1.02 (0.97)
Dummy disabled	0.2205	1.35 (1.00)	0.3730	0.25 (0.15)
Dummy early retired	0.0845	0.79 (0.79)	0.0904	0.09 (0.10)
Dummy self employed	0.4692	1.79 (1.77)	3.5749	1.51 (1.66)
Dummy work (–2 yrs)	0.0987	0.60 (0.60)	0.3687	0.25 (0.23)
Dummy disabled (–2)	0.3841	1.75 (1.55)	5.6067	2.82 (1.68)
Dummy early ret (–2)	0.3396	2.20 (2.38)	2.0833	1.49 (1.57)
Dummy self empl (–2)	–0.2587	0.75 (0.62)	–1.7471	0.56 (0.48)
Months worked in last:				
2 years	0.0062	0.91 (0.86)	0.0757	1.21 (1.09)
5 years	–0.0101	2.01 (1.97)	–0.1218	2.67 (2.56)
10 years	0.0101	2.72 (2.49)	0.1034	3.06 (2.65)
Negative health shock	0.4040	3.49 (2.97)	3.2101	3.06 (2.67)
Positive health shock	–0.3342	1.28 (1.43)	–2.5756	1.09 (1.35)
R^2 Square	0.0231		0.0229	
F	3.5595		3.5174	

[a]Absolute *t*-values and White heteroscedasticity corrected *t*-values in parentheses.
[b]Age in 1993 in the difference equation can be related to the effect of age squared in the health level equation.

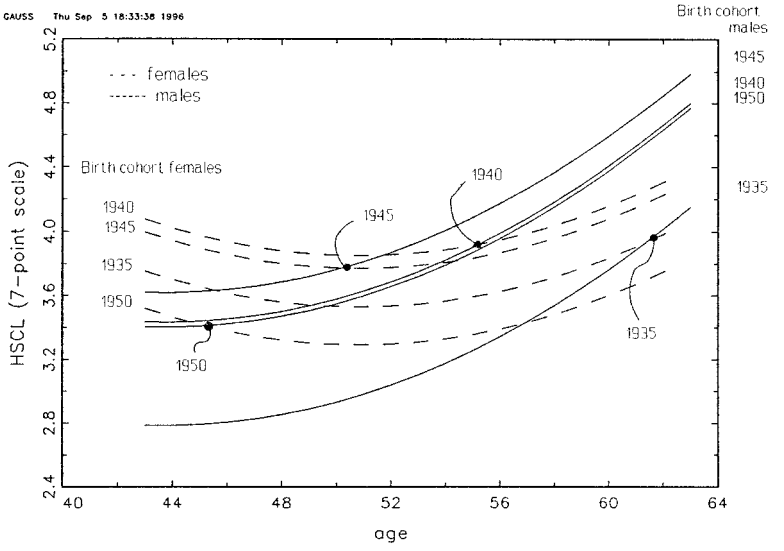

Figure 1. Age–health profiles for males and females from different birth cohorts (HSCL on seven-point scale; low = healthy).

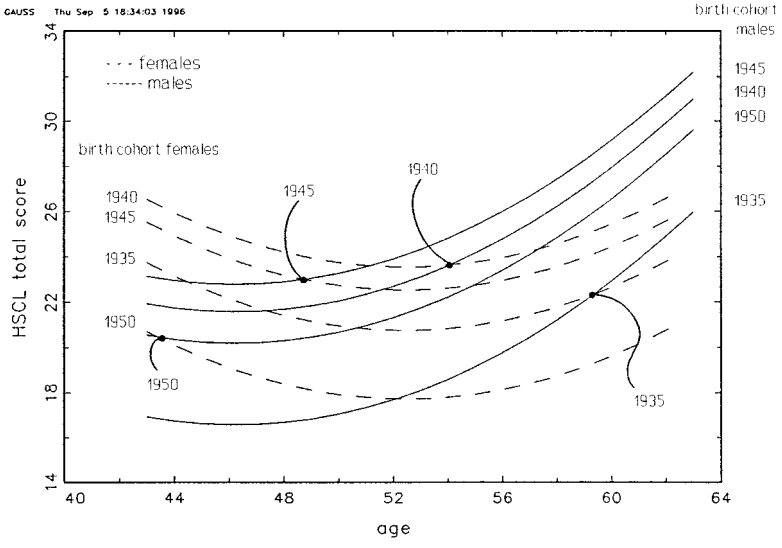

Figure 2. Age–health profiles for males and females from different birth cohorts (HSCL total score; low = healthy).

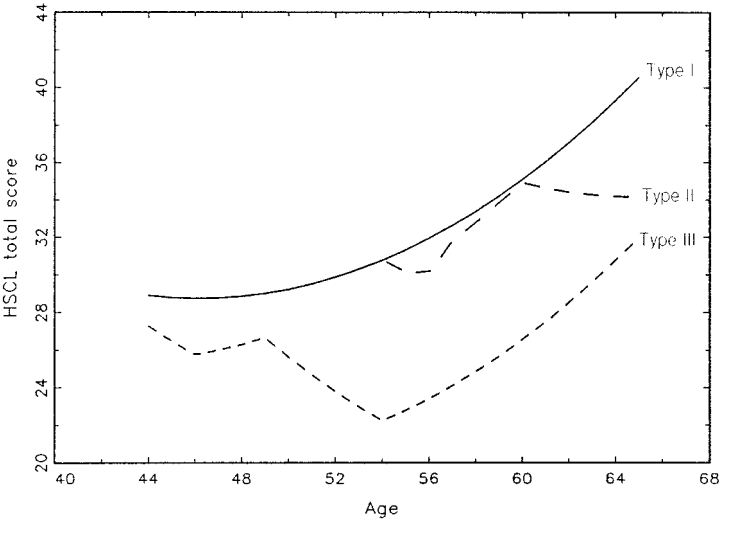

Figure 3. Age–health profiles for three different labour market patterns (HSCL total score; low = healthy).

M. Tamm, H. Tauchmann, J. Wasem and S. Gress (2007): “Elasticities of Market Shares and Social Health Insurance Choice in Germany: A Dynamic Panel Data Approach”, *Health Economics*, 16, 243-256

Goal: Estimate the price elasticities of insurers’ market shares

Model:

$$\log(s_{it}) = \alpha \log(s_{i,t-1}) + \beta \log(p_{it}) + \delta_t + \gamma_i + \varepsilon_{it}$$

s_{it}=market share of insurer i at time t
p_{it}=contribution rate (premium = contribution rate $\times$ wage)

Short-run price elasticity = β

Long-run elasticity to permanent price shock = $\dfrac{\beta}{1-\alpha}$

Data: Panel of 7 inequally spaced waves between 2001 and 2004 for German social health insurers

Table III. Fixed-effects estimates for static model

	Fixed-effects model		IV fixed-effects model	
	Coefficient	Std. error	Coefficient	Std. error
Contribution rate	−0.0045	0.0274	−0.1885*	0.0961
Within-R^2	0.1546			
F-/Wald Test	12.32***		2.01e+06***	
Observations	1960		1589	
Test for endogeneity (t-statistic)			1.77*	

Note: Regression includes time dummies for each wave. Huber–White robust standard errors given. *** indicates significance at the 1% level, ** at 5%, * at 10%.

Table IV. GMM estimates for dynamic panel data model

	First-differenced GMM				System GMM			
	x_{it} predetermined GMM1		x_{it} endogenous GMM2		x_{it} predetermined GMM3		x_{it} endogenous GMM4	
	Coef.	Std. error	Coef.	Std. error	Coef.	Std. error	Coef.	Std. error
Market share in $t-1$	0.9798***	0.0751	0.9525***	0.0378	1.0123***	0.0191	1.0453***	0.0220
Contribution rate	−0.0034	0.0545	−0.0413	0.0451	−0.1187***	0.0387	−0.1715***	0.0307
Observations	1221		1221		1588		1588	
AR(1)	−3.32***		−3.87***		−4.16***		−4.41***	
AR(2)	0.12		0.16		−0.05		−0.12	
Sargan statistic	57.65***		40.27*		73.42***		57.34**	
Diff.-Sargan test (fewer instruments)	45.76*** (25)		21.83 (20)		32.12 (25)		24.69 (20)	
Diff.-Sargan test (system vs first-dif. GMM)					15.77 (14)		17.07* (10)	

Note: Regression includes time dummies for each wave. Two-step GMM estimates with corrected standard errors (Windmeijer 2005). AR(1) and AR(2) are tests for first- and second-order serial correlation in the first-differenced residuals (Arrelano and Bond 1991). (Difference) Sargan statistics are χ^2 distributed; number in brackets behind difference Sargan test provides the number of restrictions/degrees of freedom. *** indicates significance at the 1% level, ** at 5%, * at 10%.

Table VII. Estimates of short-run premium elasticity

	GMM2		UR2	
Mean premium elasticity	−0.55		−1.09	
95% confidence interval	−1.74	+0.64	−1.43	−0.75

Note: Elasticity estimated for sample mean. Estimates based on results from Tables IV and V.

B. Gannon (2005): “A Dynamic Analysis of Disability and Labour Force Participation in Ireland 1995-2000”, *Health Economics*, 14, 925-938

Goal: Analyse the effect of disability on participation in the labour force

Model:

$$y_{it} = 1\{\beta_0 + \beta_1 y_{i,t-1} + \beta_2 D_{it} + \beta_3 D_{i,t-1} + \beta_4' z_{it} + \alpha_i + \varepsilon_{it} \geq 0\}$$

y_{it}=indicator of labour force participation ($= 1$ if i works at t)
D_{it}=disability dummy
z_{it}=individual characteristics (age, education, unearned income, ...)

Distinguish state dependence (via $y_{i,t-1}$) vs. unobserved heterogeneity (via α_i)

Data: Living Ireland Survey, 1995-2000

Likelihood function from:

$$f(y_{i,1},\ldots,y_{i,T}|y_{i,0},x_i) = \int \prod_{t=1}^{5} f(y_{i,t}|y_{i,t-1},x_{i,t},\alpha_i) f(\alpha_i|y_{i,0},\bar{x}_i) d\alpha_i$$

where:

$$\begin{aligned} f(y_{i,t}|y_{i,t-1},x_{i,t},\alpha_i) &= [\Phi(\beta_0+\beta_1 y_{i,t-1}+\tilde{\beta}' x_{i,t}+\alpha_i)]^{y_{i,t}} \\ &\quad [1-\Phi(\beta_0+\beta_1 y_{i,t-1}+\tilde{\beta}' x_{i,t}+\alpha_i)]^{1-y_{i,t}} \end{aligned}$$

with $x_{i,t} = (D_{it}, D_{i,t-1}, z_{i,t})$ and:

$$\alpha_i \sim N(\delta_0+\delta_1 y_{i,0}+\delta_2' \bar{x}_i, \sigma_\alpha^2)$$

Account for correlation of random effects with initial observations and explanatory variables!

Table 1. Sample size and composition at each wave, age 15–64, Living in Ireland Survey 1995–2000

	1995	1996	1997	1998	1999	2000
Men	50.4	50.5	50.4	49.8	49.9	49.1
Women	49.6	49.5	49.6	50.2	50.1	50.9
Age 15–24	24.9	24.7	24.2	23.7	22.8	23.1
24–34	20.5	20.2	20.3	20.5	20.0	18.7
35–44	20.6	20.7	21.1	20.9	21.4	21.3
45–54	19.1	19.4	19.3	19.7	19.8	19.5
55–65	14.8	14.9	15.0	15.2	15.9	17.4
Education						
Primary	26.9	26.3	26.2	24.6	23.8	21.8
Secondary	59.8	60.7	60.7	58.7	58.3	60.7
Third level	13.2	13.1	13.1	16.6	17.9	17.6
Married	59.1	58.7	59.2	58.5	58.6	56.9
N	7254	6337	5782	5273	4482	3670

Table 3. Variable definitions for dependent and independent variables

Variable	Definition
LFP	= 1 if participating in the labour market, = 0 otherwise
Disabled with severe limitation	= 1 if disabled and severely limited in daily activities, = 0 otherwise
Disabled with some limitation	= 1 if disabled and limited to some extent in daily activities, = 0 otherwise
Disabled with no limitation	= 1 if disabled and not limited in daily activities, = 0 otherwise (Base category=no disability)
Age 15–24	= 1 if aged 15–24 years, = 0 otherwise
Age 25–34	= 1 if aged 25–34 years, = 0 otherwise
Age 35–44	= 1 if aged 35–44 years, = 0 otherwise
Age 45–54	= 1 if aged 45–54 years, = 0 otherwise (Base category=aged 55–64 years)
BMW	= 1 if living in border, midlands, west region, = 0 otherwise (Base category=rest of country)
Secondary education	= 1 if highest level of education completed is secondary, = 0 otherwise
Third level education	= 1 if highest level of education completed is third level, = 0 otherwise (Base category=no qualifications or highest level of education completed is primary)
Married	= 1 if married or living with a partner, = 0 otherwise
Age youngest child < 4	= 1 if age of youngest child is less than 4, = 0 otherwise
Age youngest child > = 4 and < 12	= 1 if age of youngest child is greater than or equal to 4 and less than 12, = 0 otherwise
Age youngest child > = 12 and < 18	= 1 if age of youngest child is greater than or equal to 12 and less than 18, = 0 otherwise (Base category=no children)
Unearned income	=Net household income − net individual disposable income (Net individual disposable income includes net incomes from work, social welfare payments and child benefit. Net household income aggregates individual data to household level)

Note: The regional classifications are based on the NUTS (Nomenclature of Territorial Units) classification used by Eurostat.

Table 6. Panel model results

	Men (coefficients)			Women (coefficients)		
	[1] Pooled static	[2] Random effects dynamic (re-scaled)	[3] Pooled dynamic	[4] Pooled	[5] Random effects dynamic (re-scaled)	[6] Pooled dynamic
Lag LFP		0.7511**	1.687**		0.7494**	1.7974**
		(0.1194)	(0.0918)		(0.0835)	(0.0623)
Disabled with severe limitation	−1.2368**	−0.6639**	−0.5653**	−0.9173**	−0.8256**	−1.1359**
	(0.1314)	(0.2653)	(0.2218)	(0.1736)	(0.2827)	(0.2393)
Disabled with some limitation	−0.7886**	−0.5159**	−0.4757**	−0.3296**	−0.3137**	−0.4210**
	(0.0814)	(0.1594)	(0.1285)	(0.0755)	(0.1283)	(0.1106)
Disabled with no limitation	−0.2066**	−0.3464**	−0.3397**	−0.0175	−0.1811**	−0.2732**
	(0.1042)	(0.2161)	(0.1380)	(0.0928)	(0.1497)	(0.1326)
Lagged disability						
Disabled with severe limitation	−1.0555**	−0.2534	−0.0765	−0.6203**	−0.1470	0.0102
	(0.1275)	(0.2593)	(0.2465)	(0.1626)	(0.2863)	(0.2643)
Disabled with some limitation	−0.5802**	0.0259	0.1796	−0.2742**	−0.0056	0.0514
	(0.0783)	(0.1592)	(0.1302)	(0.0714)	(0.1303)	(0.1177)
Disabled with no limitation	−0.0925	0.0887	0.1298	−0.0290	−0.0495	−0.0464
	(0.1175)	(0.2254)	(0.1461)	(0.0962)	(0.1566)	(0.1363)
Initial condition						
LFP in 1995		1.2059**	0.6399**		0.8984**	0.6315**
		(0.2096)	(0.0944)		(0.1353)	(0.0626)
Random effect (*time averages*)						
Disabled with severe limitation		−0.8815**	−0.9013**		−0.3077	−0.2653
		(0.5948)	(0.4588)		(0.7211)	(0.5607)
Disabled with some limitation		−0.7265**	−0.7146**		−0.1387	−0.1209
		(0.3237)	(0.2371)		(0.2744)	(0.2041)
Disabled with no limitation		0.3616	0.2146		0.4464*	0.5171*
		(0.5068)	(0.3297)		(0.3844)	(0.3087)
Constant	0.4642**	−0.8210**	−1.0449**	−0.5446**	−0.1118**	−1.5214**
	(0.1332)	(0.2167)	(0.1332)	(0.1074)	(0.1595)	(0.0945)
N	5930	5930	5930	6330	6330	6330
Pseudo R^2	0.2772		0.5371	0.1700		0.5303
Rho		0.4684**			0.3984**	

**$p \leqslant 0.05$, *$p \leqslant 0.10$.
(Significance in random effects models are based on *t*-stats on base coefficients, not on the rescaled coefficients reported in this table). Estimation was carried out using the xtprobit command in Stata Version 7.0.

Table 7. Average partial effects

	Men			Women		
	[1] Pooled static	[2] Random effects dynamic (rescaled)	[3] Pooled dynamic	[4] Pooled static	[5] Random effects dynamic (rescaled)	[6] Pooled dynamic
Disabled with severe limitation	−0.3346**	−0.1111**	−0.0865**	−0.3377**	−0.2557**	−0.3979**
	(0.0504)		(0.0471)	(0.0502)		(0.0598)
Disabled with some limitation	−0.1680**	−0.0746**	−0.0654**	−0.1308**	−0.0787**	−0.1666**
	(0.0238)		(0.0230)	(0.0295)		(0.0428)
Disabled with no limitation	−0.0330**	−0.0461**	−0.0438**	−0.0069	−0.0435**	−0.1086**
	(0.0187)		(0.0221)	(0.0369)		(0.0524)
Lag LFP*		0.1292**	0.3927**		0.1296**	0.6286**

**$p \leqslant 0.05$, *$p \leqslant 0.10$.
(Significance in random effects models are based on t-stats on base coefficients, not on the rescaled coefficients reported in this table). Estimation was carried out using the xtprobit command in Stata Version 7.0.

REFERENCES

Review articles

Jones, A. (2000): “Health Econometrics”, in *Handbook of Health Economics*, A. Culyer and J. Newhouse (eds.), North-Holland.

Jones, A. (2007): “Panel Data Methods and Applications to Health Economics”, in *The Palgrave Handbook of Econometrics*, Volume 2: Applied Econometrics, T. Mills and K. Patterson (eds.).

Papers using linear static panel data models

Kerkhofs, M., and M., Lindeboom (1997): “Age Related Health Dynamics and Changes in Labour Market Status”, Health Economics, 6, 407-423.

Lindeboom, M., Portrait, F., and G., van den Berg (2002): “An Econometric Analysis of the Mental-Health Effects of Major Events in the Life of Older Individuals”, Health Economics, 11, 505-520.

Papers using linear dynamic panel data models

Bishai, D. (1996): “Quality Time: How Parents’ Schooling Affects Child Health Through Its Interaction with Childcare Time in Bangladesh”, Health Economics, 5, 383-407.

Brown, T., Coffman, J., Quinn, B., Scheffer, R., and D., Schwalm (2005): “Do Physicians Always Flee from HMO’s? New Results Using Dynamic Panel Estimation Methods”, Health Services Research, 40, 357-73.

Tamm, M., Tauchmann, H., Wasem, J., and S., Gress (2007): “Elasticities of Market Shares and Social Health Insurance Choice in Germany: A Dynamic Panel Data Approach”, Health Economics, 16, 243-256.

Papers using discrete choice and sample selection panel models

Askildsen, J., Baltagi, B., and T., Holmas (2003): “Wage Policy in the Health Care Sector: A Panel Data Analysis of Nurses’ Labour Supply”, Health Economics, 12, 705-719.

Bjorklund, A. (1985): “Unemployment and Mental Health: Some Evidence from Panel Data”, Journal of Human Resources, 20, 469-483.

Contoyannis, P., Jones, A., and N., Rice (2004): “The Dynamics of Health in the British Household Panel Survey”, Journal of Applied Econometrics, 19, 473-503.

Gannon, B. (2005): “A Dynamic Analysis of Disability and Labour Force Participation in Ireland 1995-2000”, Health Economics, 14, 925-938.