

青少年成長歷程計畫(TYP) 流失分析

林桂綉
臺北大學社會學系助理教授

林定香
臺北大學統計學系副教授

謝沛霖
臺北大學社會系碩士生

Panel Study(固定樣本長期追蹤)

- 優點
- 缺點

資料流失(Data attrition)

- 可能造成的問題：
 - 統計檢定力
 - 樣本代表性
 - 參數估計偏誤
- 資料流失的影響視研究目的而定
 - 推估母體分配
 - 推估變項間關係

相關研究方法

Attrition bias

- Ignorable attrition
 - multiple imputation
 - ignorable ML
- Nonignorable attrition
 - selection model
 - pattern-mixture model

相關研究結果

- 採用Selection model分析的研究
 - 選樣偏誤係數檢定顯著，但迴歸係數差異不大
 - 選樣偏誤係數檢定不顯著，無選樣偏誤
- 採用其他方法 (加權或分組分析)的研究
 - 無選樣偏誤，係數差異檢定不顯著
- 其他分析結果不具穩健性的研究

資料與變項測量

資料來源：

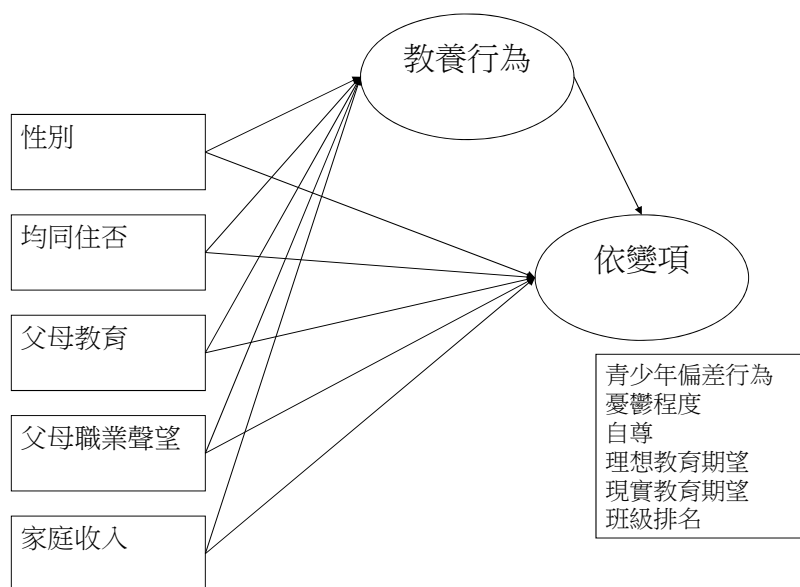
- 使用TYP國一樣本，2000年至2007年調查訪問資料

變項：

- 流失分組(w1~w8)
- Wave 1
 - 控制變項：性別、與父母同住否、父母教育程度、父母職業聲望、家庭月收入
 - 中介變項：父母教養行為
 - 依變項：青少年的偏差行為、憂鬱程度、自尊、理想以及現實教育期望、班級排名

研究方法

- 結構方程模型
 - 測量誤差
- 多群組分析
 - 斜率及截距不同
 - 斜率相同
 - 斜率及截距相同



圖一、跨組比較模型圖

四、研究結果

資料流失狀況

表一. 台灣青少年計畫資料流失狀況

波	持續接受調查者		流失者		再次參加調查者
	人數	% ^a	人數	%	
1	2690	100%	---	---	
2	2583	96%	107	4%	---
3	2512	93%	71	3%	8
4	2215	82%	297	11%	24
5	1839	68%	376	14%	83
6	1431	53%	408	15%	251
7	1176	44%	255	9%	308
8	1008	37%	168	6%	325

^a為佔第一年參與者百分比

資料缺失

- 因分析重點在於樣本流失，因此對資料缺失不多著墨。
- 個別變項缺失以父母職業聲望缺失比例最高(11.2%)，部分缺失是因無對應的社經地位等第資訊。父母親最高教育程度缺失資料有6.7%，其它變項均非常低。

資料流失相關因素

表二.各流失分組基本統計特性

		1 持續參與		2 缺1~2次		3 流失不只兩次		χ^2/F	
		次數	%	次數	%	次數	%		
性別	男	496	49.2	507	52.8	375	52	n.s.	
	女	512	50.8	454	47.2	346	48		
與父母同住否	同住	902	89.6	841	87.8	546	75.8	***	
	其他	105	10.4	117	12.2	174	24.2		
		Mean	sd	mean	sd	mean	sd		
父母教育程度		2.89	1.17	2.9	1.3	2.96	1.2	n.s.	
父母職業聲望		3.77	1.4	3.87	1.39	3.62	1.48	**	2>3
家庭月收入		3.95	2.86	4.2	3.18	3.86	3.15	n.s.	
父母教養行為		3.03	0.46	2.96	0.48	2.93	0.5	***	1>2>3
偏差行為		0.05	0.13	0.06	0.22	0.11	0.27	***	1,2<3
憂鬱程度		0.43	0.46	0.44	0.48	0.48	0.52	n.s.	
自尊		2.69	0.44	2.66	0.44	2.64	0.45	n.s.	
理想教育期望		3.21	1.22	3.15	1.23	2.92	1.3	***	1,2>3
現實教育期望		4.08	1.28	3.95	1.31	3.6	1.42	***	1,2>3
班級排名		2.7	1.14	3	1.17	4.08	1.17	***	1>2>3

附表二、模型比較

依變項		χ^2	df	p-value	CN	CFI	RMSEA
偏差行為	斜率相同	23.37	22	0.381	3304	0.998	0.006
	斜率和截距相同	32.929	24	0.106	2502	0.986	0.014
憂鬱程度	斜率相同	24.003	22	0.347	3213	0.997	0.007
	斜率和截距相同	24.867	24	0.413	3309	0.999	0.004
自尊	斜率相同	23.774	22	0.359	3238	0.997	0.006
	斜率和截距相同	25.267	24	0.391	3250	0.998	0.005
理想教育期望	斜率相同	21.669	22	0.48	3472	1	0
	斜率和截距相同	29.995	24	0.185	2676	0.993	0.012
現實教育期望	斜率相同	23.73	22	0.362	3218	0.998	0.006
	斜率和截距相同	48.781	24	0.002	1671	0.97	0.023
班級排名	斜率相同	18.274	22	0.69	4207	1	0
	斜率和截距相同	74.79	24	0	1098	0.937	0.033

結論

- 樣本流失對於變項間關係研究的影響不大，不影響研究的實質結論
- 但會因樣本數減少使檢定力下降
- **Sensitivity analysis**—不同流失狀況分組

討論

- 納入**CFA**測量模型
- **MI/FIML**等缺失值處理方法
- 國三樣本
- 運用**Panel data**特性

建議

- 未來可探討調查研究設計及執行步驟對於資料流失的影響。
- 此外，建議收集與受訪者流失相關資料。例如由訪員記錄受訪者對接受調查研究的態度、接觸次數、花費的時間等，有助於估計流失概率。

- 以上資料流失分析的結果只適用於特定依變項、特定模型。
- 建議研究者在進行多期資料分析時，檢測是否有資料流失偏誤，以確認研究結論是否具備穩定性。