

# 跨期資料分析： 時期、世代與模型選擇

林宗弘

中央研究院社會學研究所

**2019/05/10**

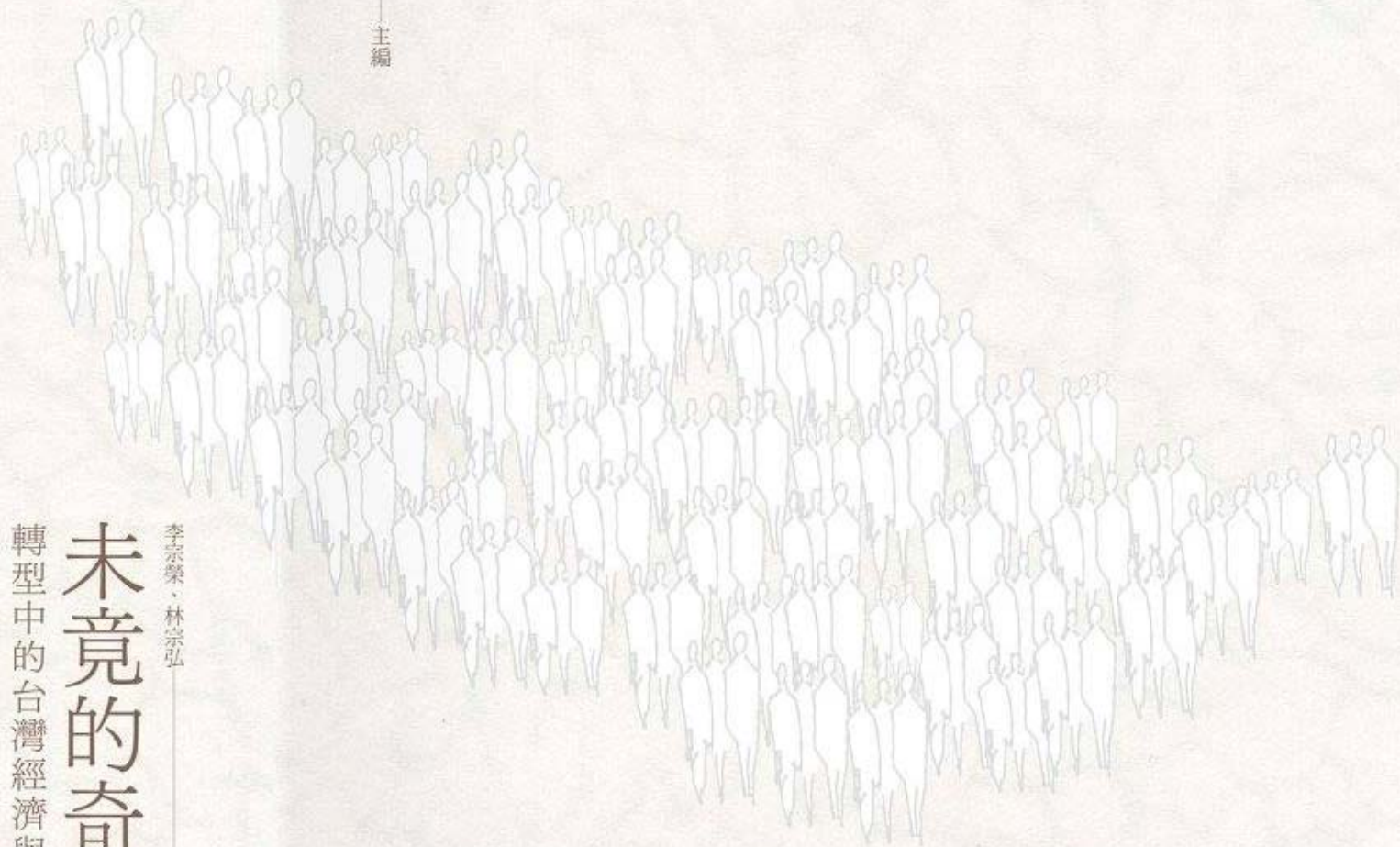
# 未竟的奇蹟

轉型中的  
台灣經濟  
與社會

李宗榮、林宗弘

主編

- 經濟轉型的挑戰
- 產業的網絡與創新
- 企業大型化與隱形冠軍
- 台商與中國效應



李宗榮、林宗弘  
主編  
未竟的奇蹟  
轉型中的台灣經濟與社會

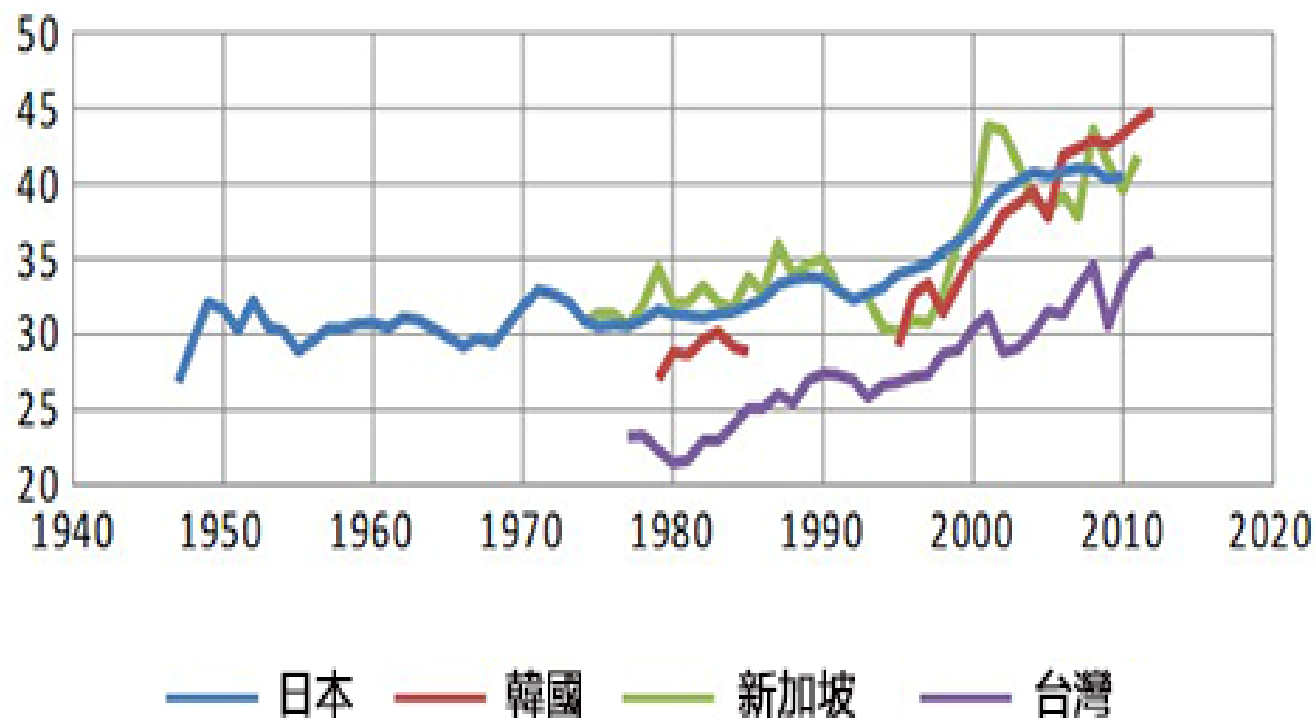
UNFINISHED MIRACLE TAIWAN'S ECONOMY AND SOCIETY IN TRANSITION

# 審視台灣社會變遷：理論的意義

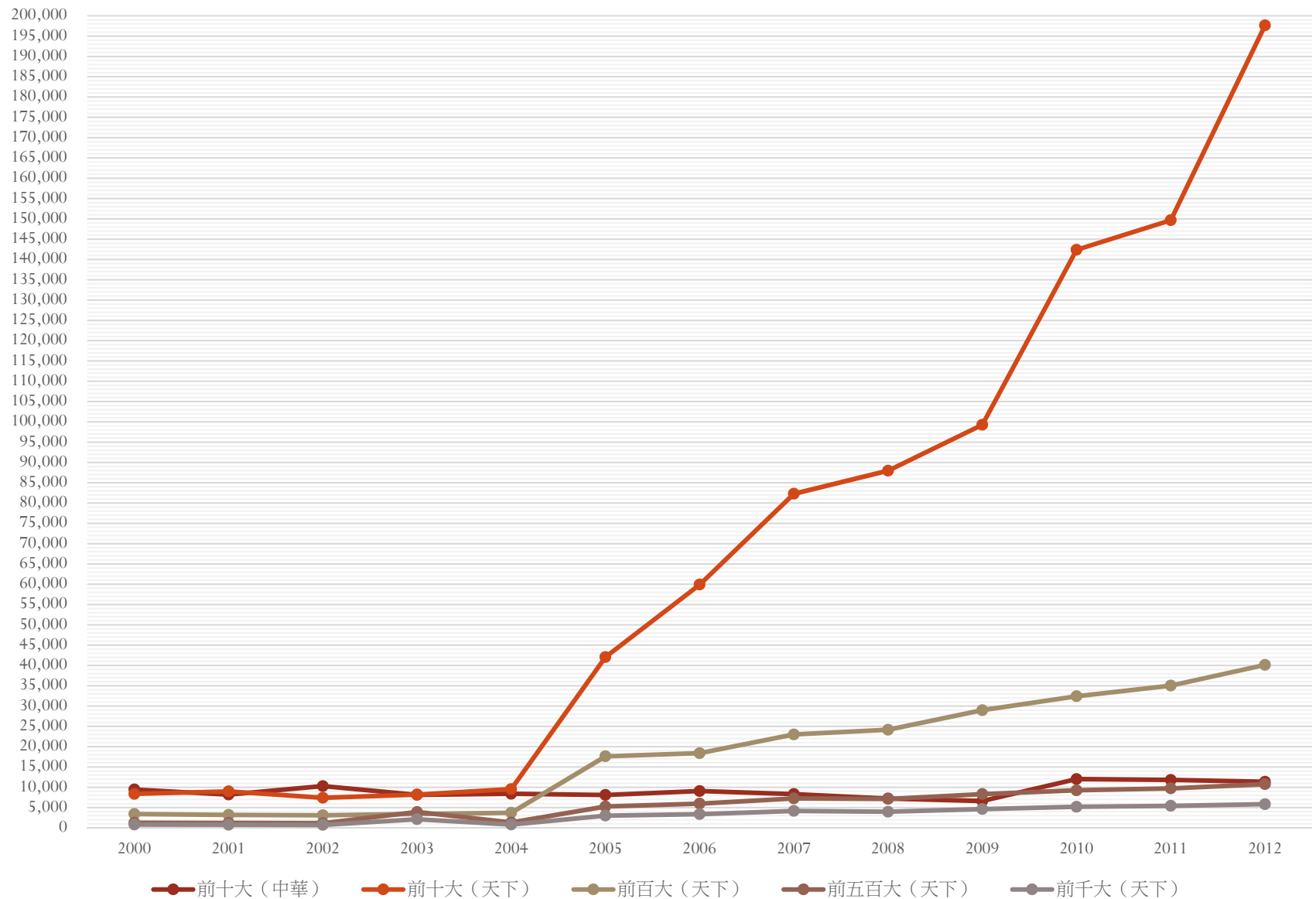
	“奇蹟”典範	“衰退”典範
被解釋項	為何台灣創下極高長期經濟成長率(發展)擺脫邊陲?(台灣/東亞例外論)	為何台灣未來即將面臨經濟與社會危機?(全球後工業國與東亞國家)
國家角色	威權的、自主的、發展導向的	新自由主義全球化衝擊下的失能國家(以及對發展國家的檢討)
產業結構	以中小企業為主的勞力密集與出口導向製造業	西進中國後企業大型化(“進擊的巨人”)與本土產業升級(資本密集)提高財團影響力
金融與財政	國家引導銀行投資的財政與金融政策、國有銀行與民間借貸的雙元結構、1980年代後以散戶為主的股市	銀行私有化與自由化、金改後的“吳辜蔡花”體系、外資與法人主導的股市、股東權益為主的上市公司經營方式
社會流動	中小企業與城市新興中產階級崛起創下國際最高的社會流動比率	產業升級資本外移壓制中小企業、創業衰退、社會流動停滯、失業與彈性工作擴張
性別、家庭與人口	儒教傳統下的父權家庭、犧牲女性(尤其是長女)的教育與勞動市場歧視、早婚/低離婚率與有效的家庭計畫(人口紅利)	少子女化造成性別教育投資平衡、女性勞參率高但仍有歧視，青年貧窮導致推遲婚育年齡、結婚率生育率降低、離婚率升高(父權家庭崩壞)、人口老化(人口負債)
政治衝突與社會分化	國民黨威權國家對抗本土公民社會(與此高度相關的省籍或族群衝突)	民主化(與此相關的中國因素與本土認同深化)後的階級衝突與世代對抗

# 台灣前10%家庭拿走35%所得

TOP10%



# 台灣製造業平均員工數，2000~2012

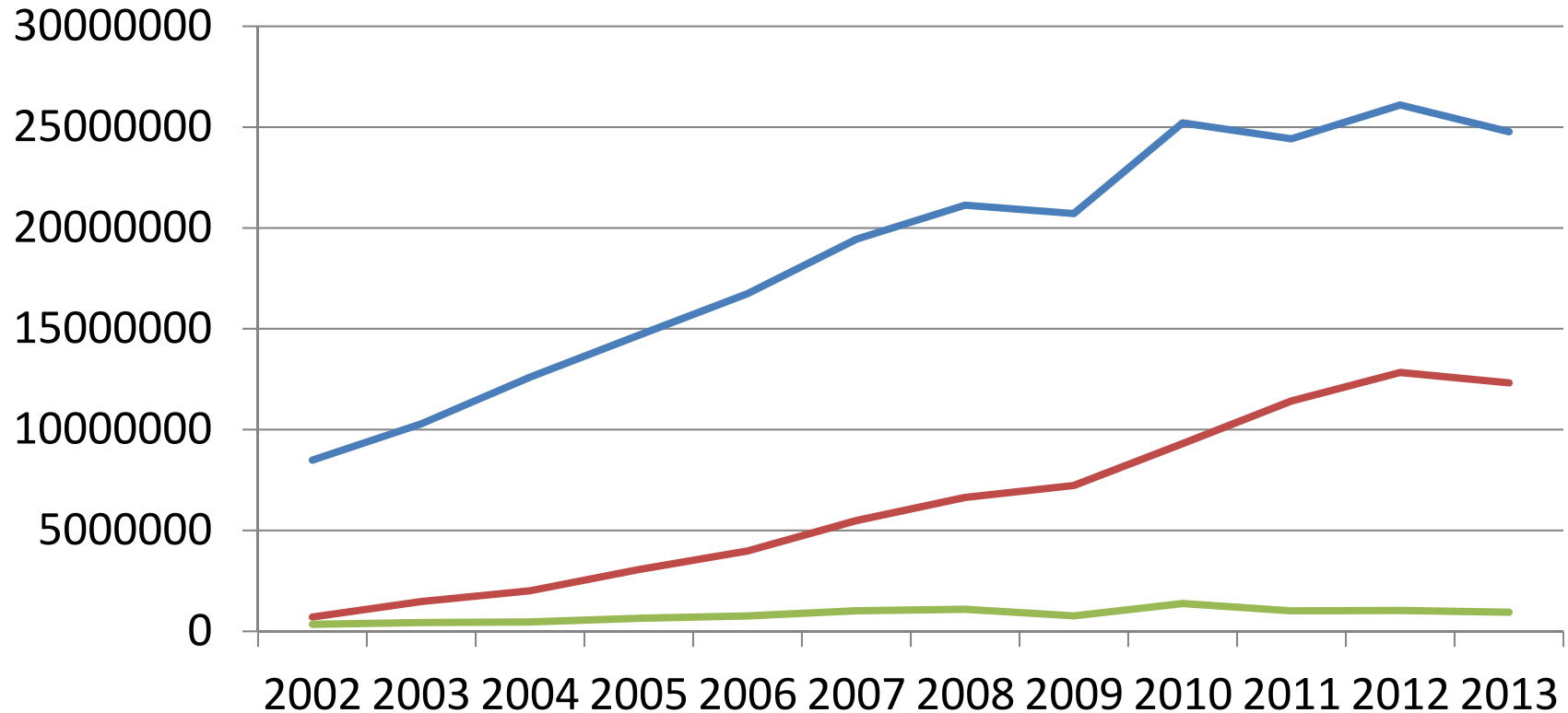


# 三種企業規模巨型化的策略

	三星	鴻海	台積電
組織結構	跨國品牌壟斷資本： 垂直與水平整合	無品牌勞力密集 <b>OEM</b> ，水平整合	資本與技術密集 <b>OEM</b> ，水平整合
雇用人數	約 <b>220,000(2012)</b> ，略 低於一半為韓國雇員	<b>129</b> 萬人，台灣 雇員 <b>6,472</b> 人 (2013)	<b>39,267</b> 人在台灣 (2012)
資本支出	約 <b>120</b> 億美元(2013)	<b>7</b> 億美元(2010)	<b>90</b> 億美元(2013)
平均每人資本支出	<b>54,000</b> 美元	~ <b>540</b> 美元	<b>265,000</b> 美元
營業額佔該國 <b>GDP</b> 比例	<b>22%(2013)</b>	<b>22%</b>	<b>4%</b>
毛利率	<b>18%(2013)</b>	<b>2.4%(2013)</b>	<b>32.8(2012)</b>

# 產業西進：台灣經濟的體外循環

**500大集團營收淨額(單位：百萬新台幣)**



—台灣

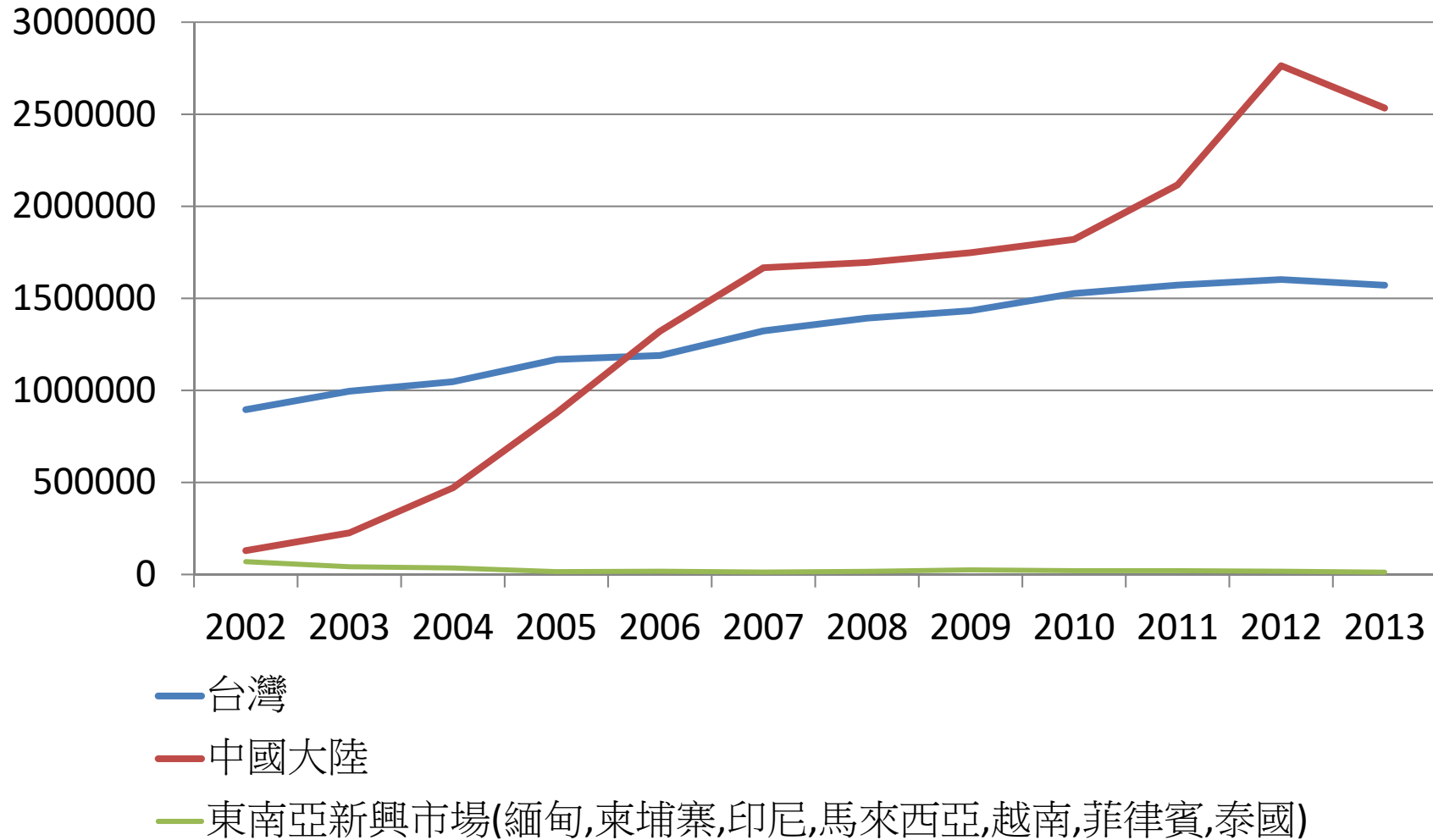
—中國大陸

—東南亞新興市場(緬甸,柬埔寨,印尼,馬來西亞,越南,菲律賓,泰國)



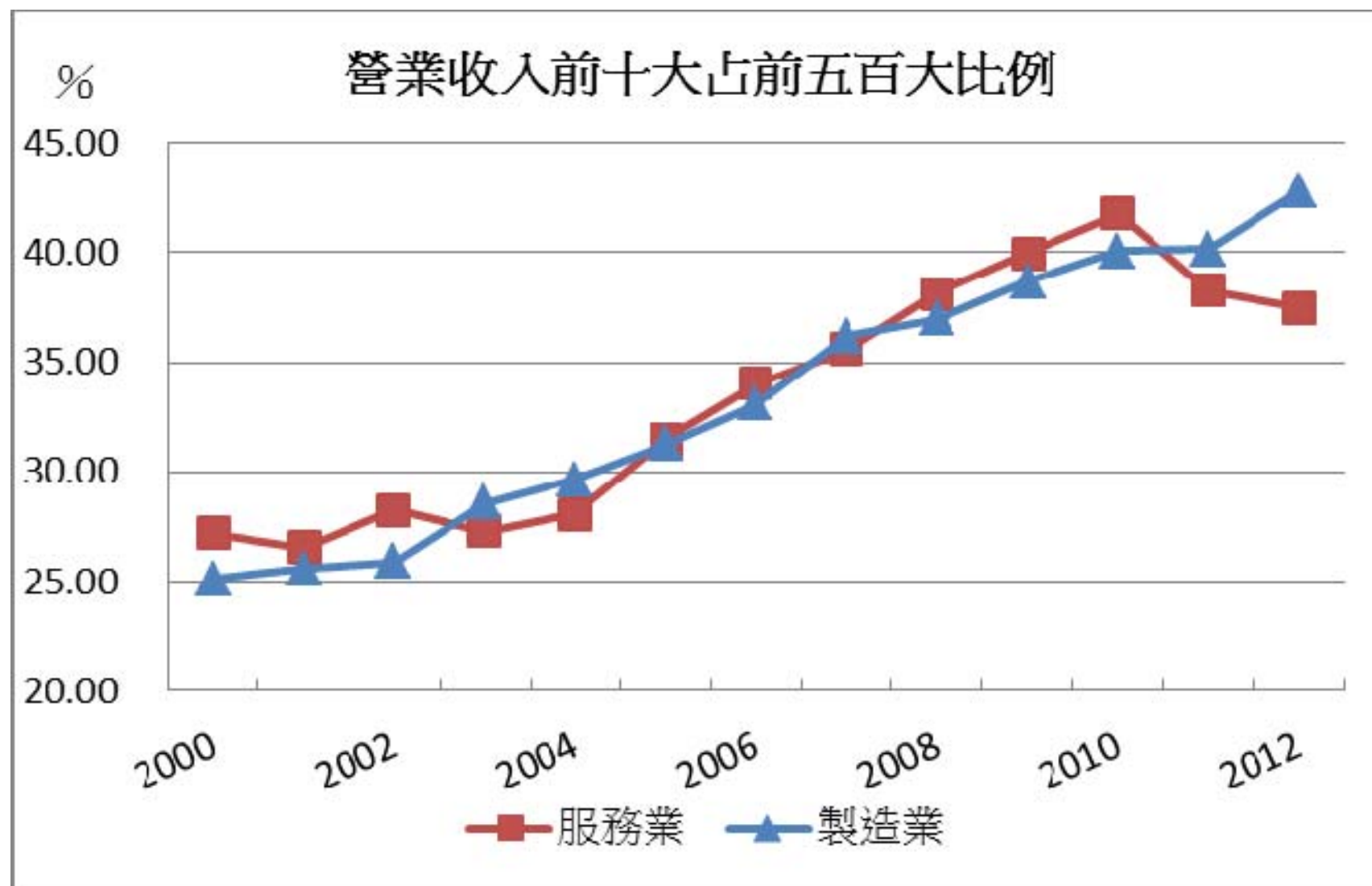
# 產業西進：台灣經濟的體外循環

500大集團員工人數

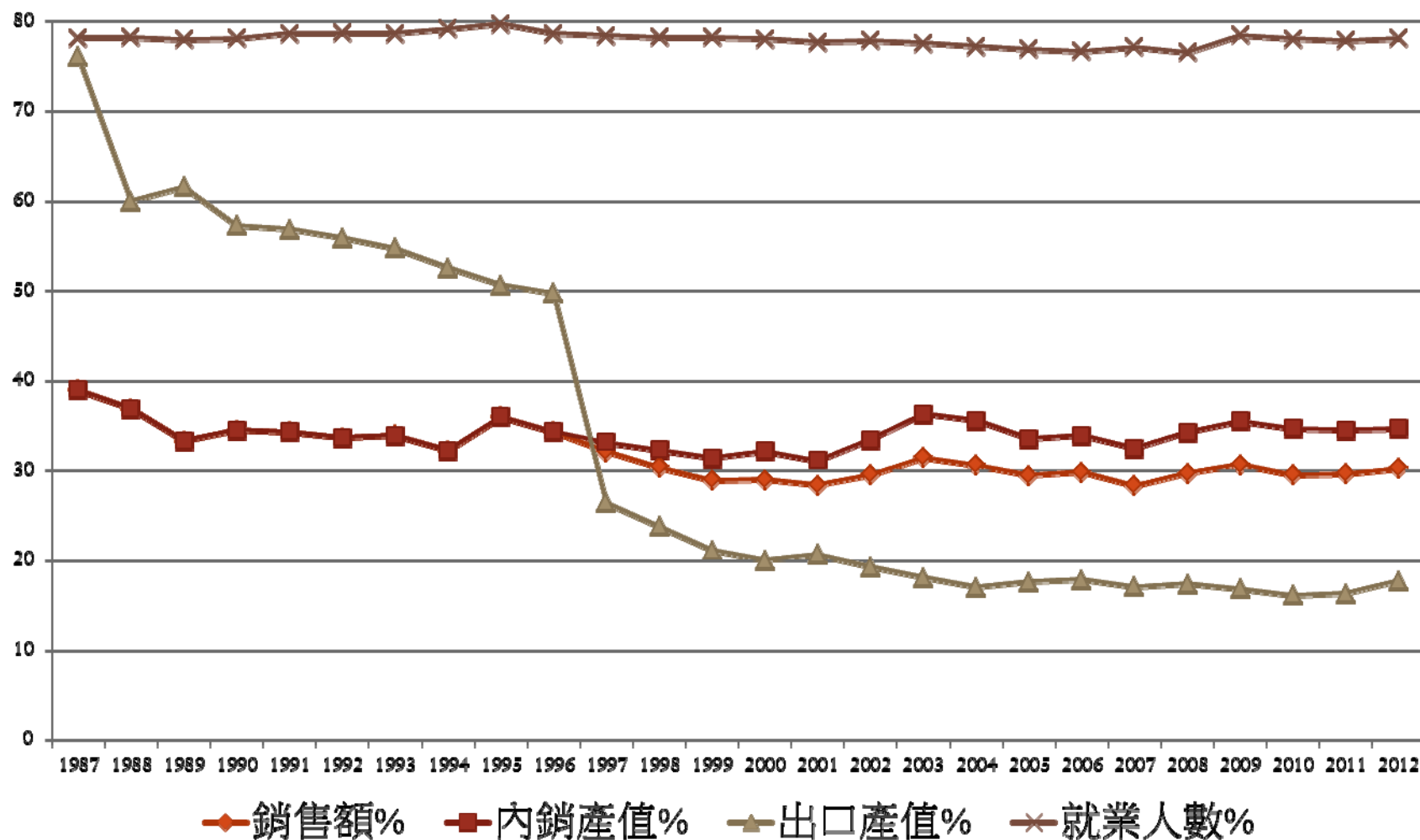




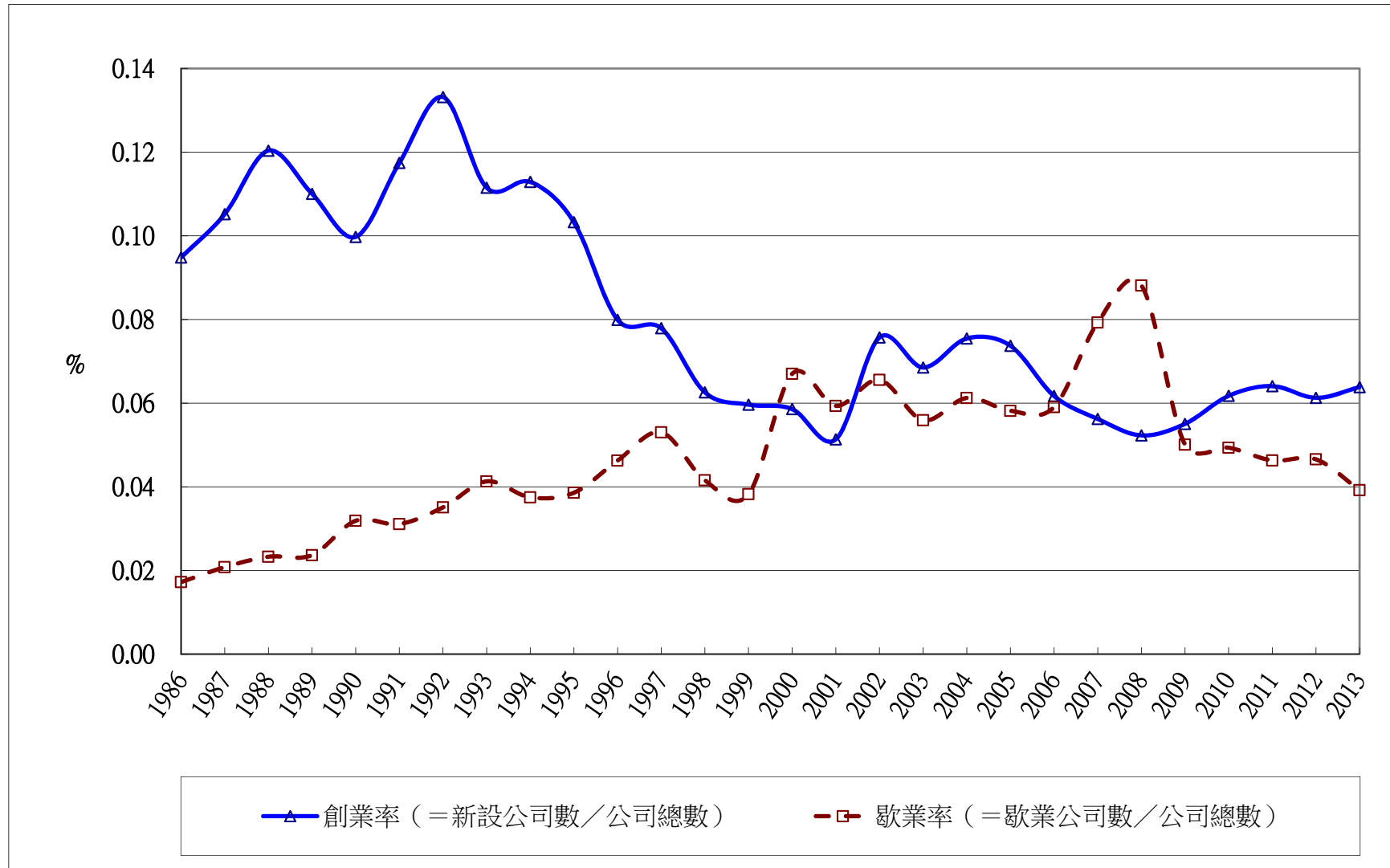
# 資本集中、創業無望



# 台灣中小企業的敗退，1987-2012

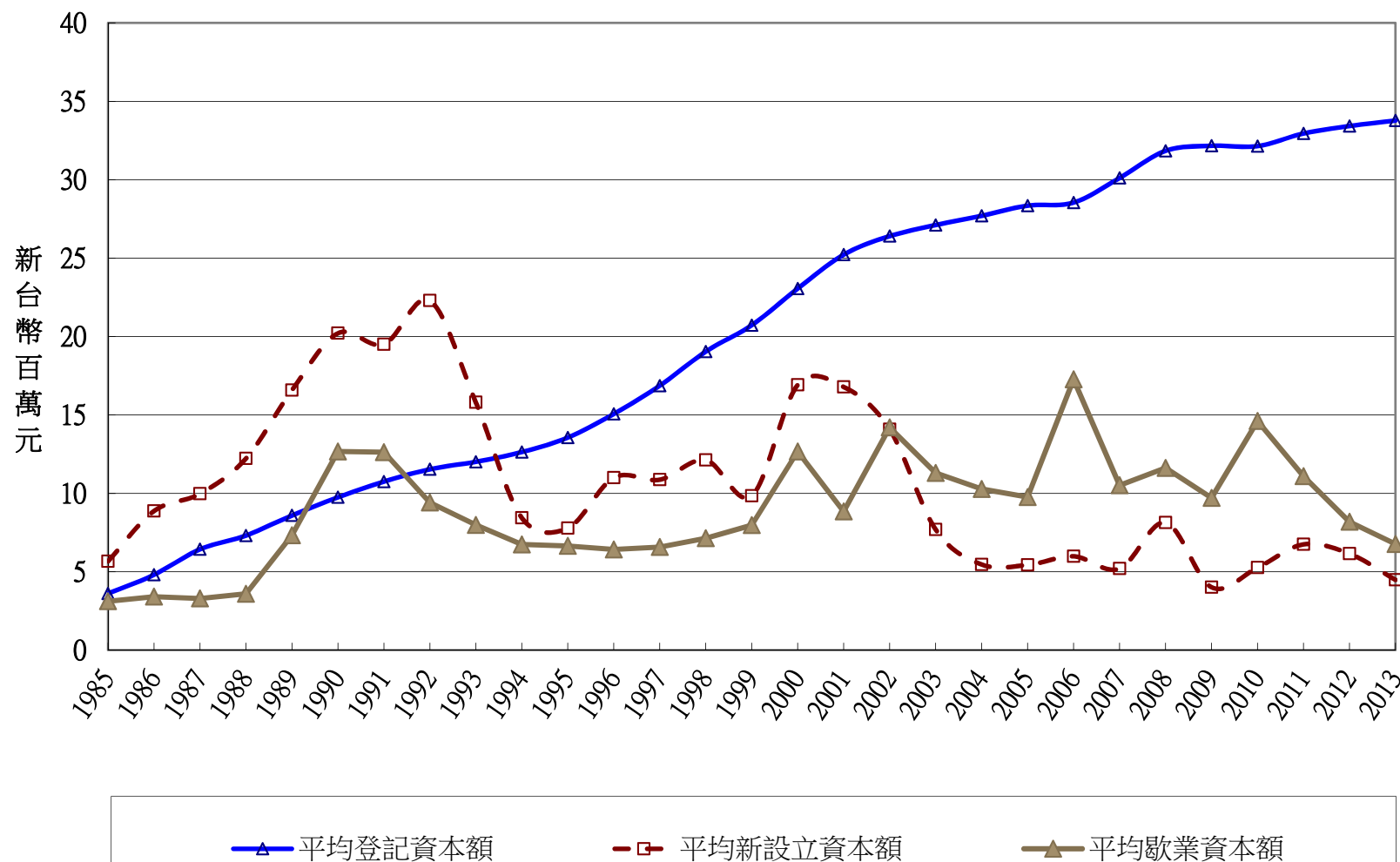


# 資本集中與創業困難



台灣公司創業及歇業趨勢，1986-2013

# 公司平均資本額提高與集中化



台灣公司登記資本額的成長，1985-2013

# 台灣社會轉型二十年：產業轉型與資本西進

相關社會指標（%）	1981	1986	1991	1996	2001	2006	2011	2016
<u>後工業化/專業化</u>								
服務業佔就業人口%	38.8	41.4	47.1	52.4	56.5	58.5	58.6	59.2
大學以上佔就業者%	5.2	5.8	7.2	9.5	12.5	19.3	27.4	33.9
<u>階級結構/從業身分</u>								
雇主	4.5	4.3	5.1	5.3	5.2	5.1	4.5	3.9
自營作業者	20.8	20.2	18.6	17.0	15.8	13.9	12.4	11.7
無酬家屬工作者	10.4	10.9	9.1	8.4	7.3	6.4	5.4	5.0
私營受僱者	51.9	52.8	55.7	58.0	61.5	65.3	68.1	70.2
政府受僱者	12.4	11.9	11.4	11.3	10.2	9.2	9.7	9.1
<u>性別分工轉型</u>								
大學就業者女性%	26.3	30.5	34.0	39.2	41.5	45.2	48.7	55.4
女性勞動參與率	38.8	45.5	44.4	45.8	46.1	48.7	50.0	50.8
女性/男性平均工資%	64.2	65.1	66.2	71.0	75.8	79.2	79.9	83.0
<u>社會不平等</u>								
五分位可支配所得%	23.8	21.8	20.1	18.6	15.6	16.6	16.2	16.4
中間三等分/總所得%	54.2	53.5	53.6	53.9	52.5	53.3	53.2	53.2
吉尼係數	28.1	29.6	30.8	31.7	35.0	33.9	34.2	33.8

# 數據來源

1. 各年度的官方統計資料: 限制在於時間序列不夠長. 相關性太高. 缺乏個體層級的社會機制分析
2. 《台灣社會變遷調查》資料庫。我們做了兩組合併檔: 綜合組問卷(1995-2015)與階層組問卷(1992-2017)

### 3. 相關論文:

張宜君、林宗弘，2017，〈時勢造英雄：台灣所得差異的世代不平等，1992-2017〉

張宜君、林宗弘，2015，〈台灣的高等教育擴張與階級複製：混合效應維繫的不平等(1992-2012)〉，《臺灣教育社會學研究》，第15卷第2期，頁85-129。

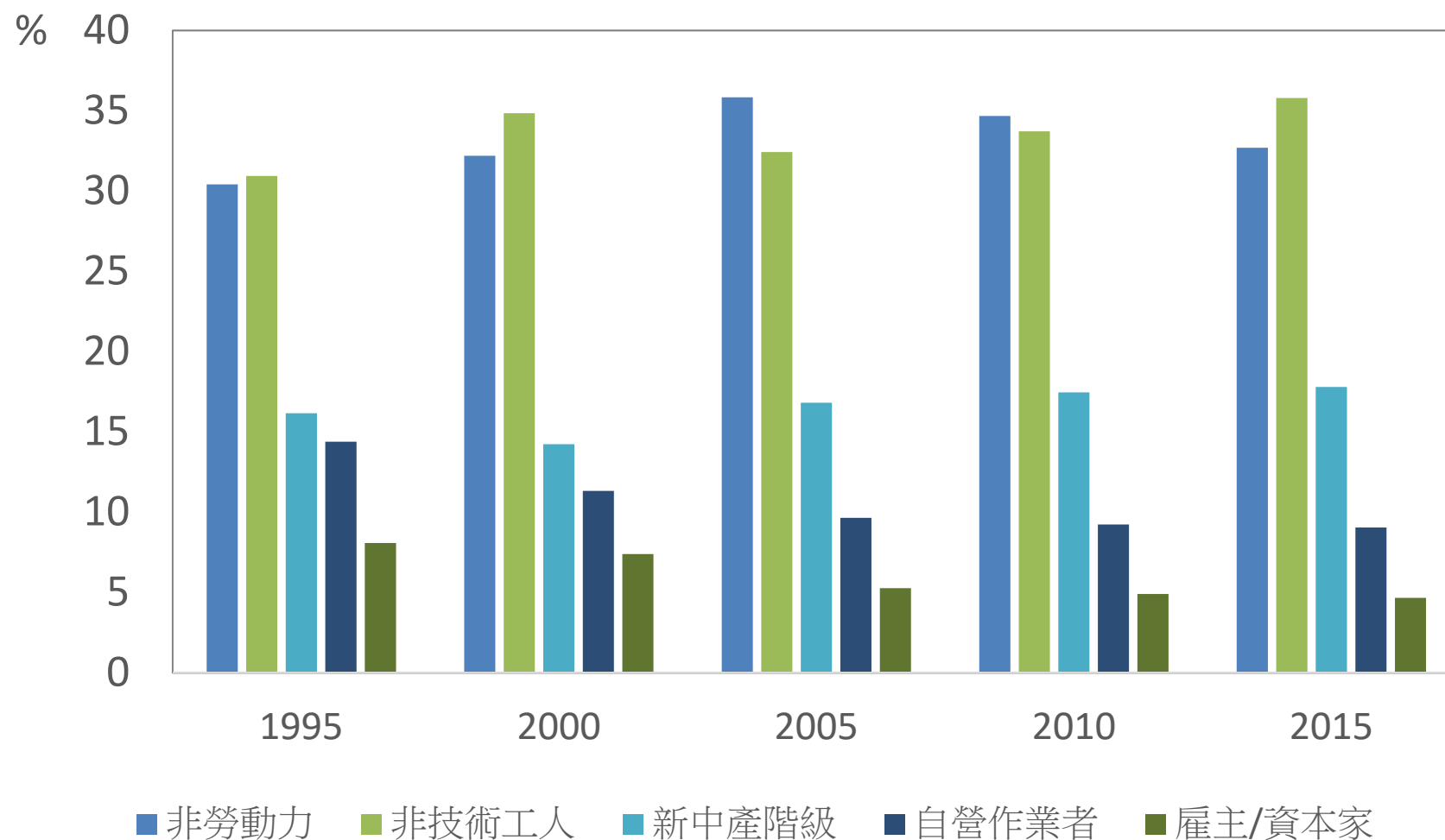
林宗弘，2015，〈再探台灣的世代政治：交叉分類隨機效應模型的應用，1995~2010〉，《人文及社會科學集刊》，第27卷第2期，頁395-436。

林宗弘，2013，〈失落的年代：台灣民眾階級認同與意識形態的變遷(2000-2010)〉，《人文與社會研究集刊》，第25卷第4期，頁689-734。

林宗弘，2009，〈台灣的後工業化：階級結構的轉型與社會不平等，1992-2007〉，《台灣社會學刊》，第43期，頁93-158。

		資本/生產工具	
		有	無
組織權威 或技術	有	雇主/資本家	新中產階級
	無	自營作業者	非技術工人

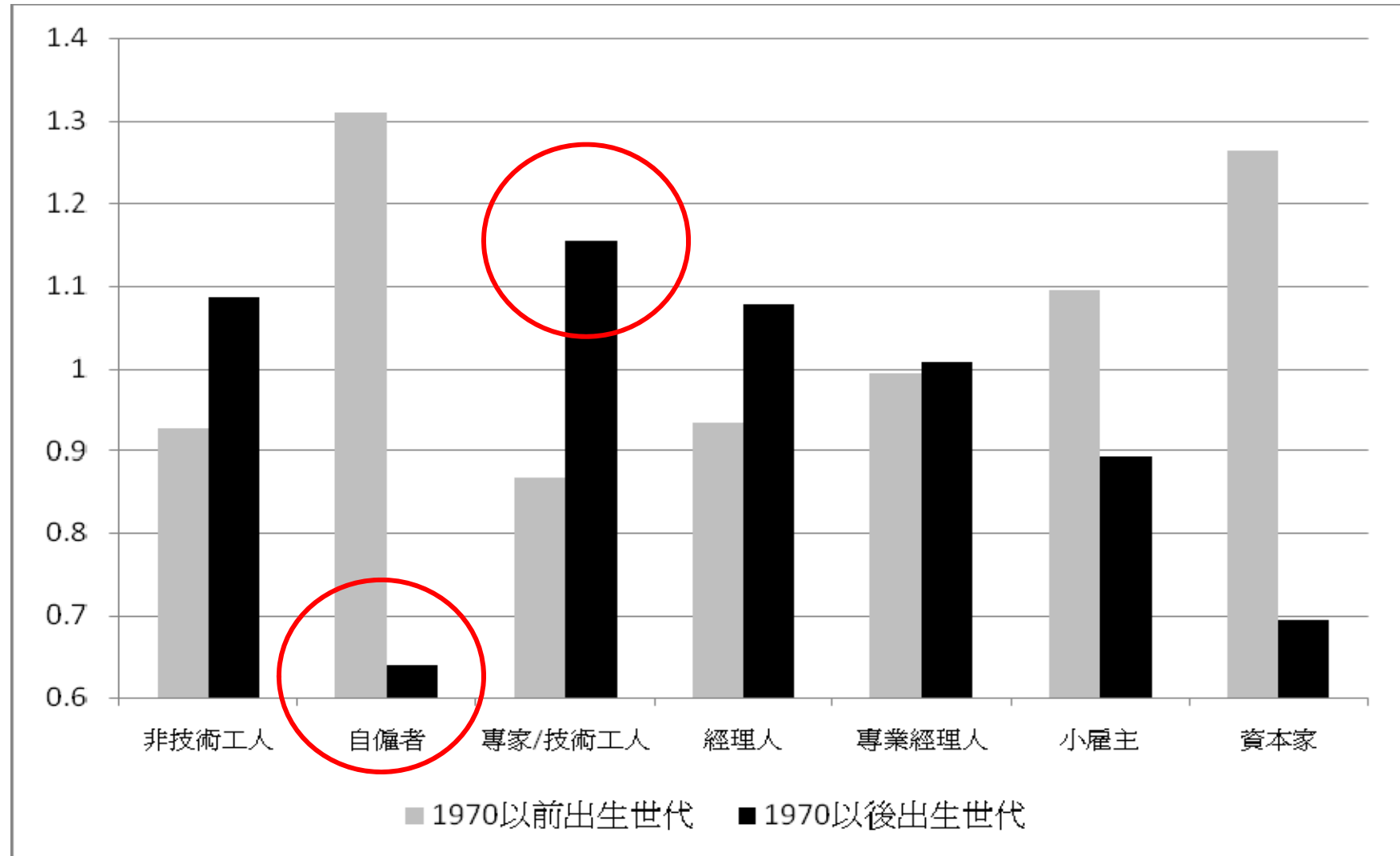
# 量化台灣社會變遷：階級結構





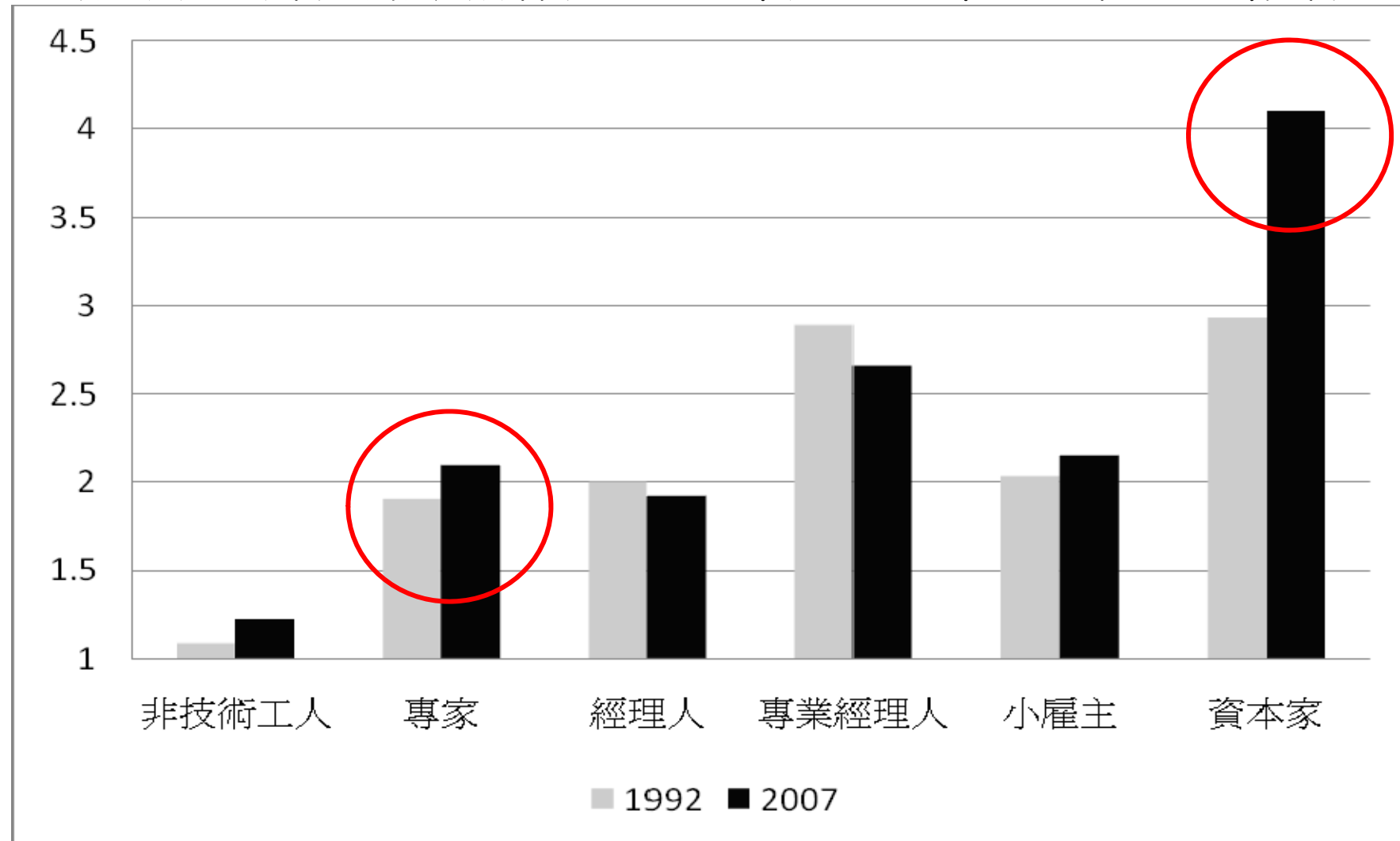
# 量化台灣社會變遷：階級結構(2009)

階級流動發生率：1970後世代黑手變頭家逐漸消失

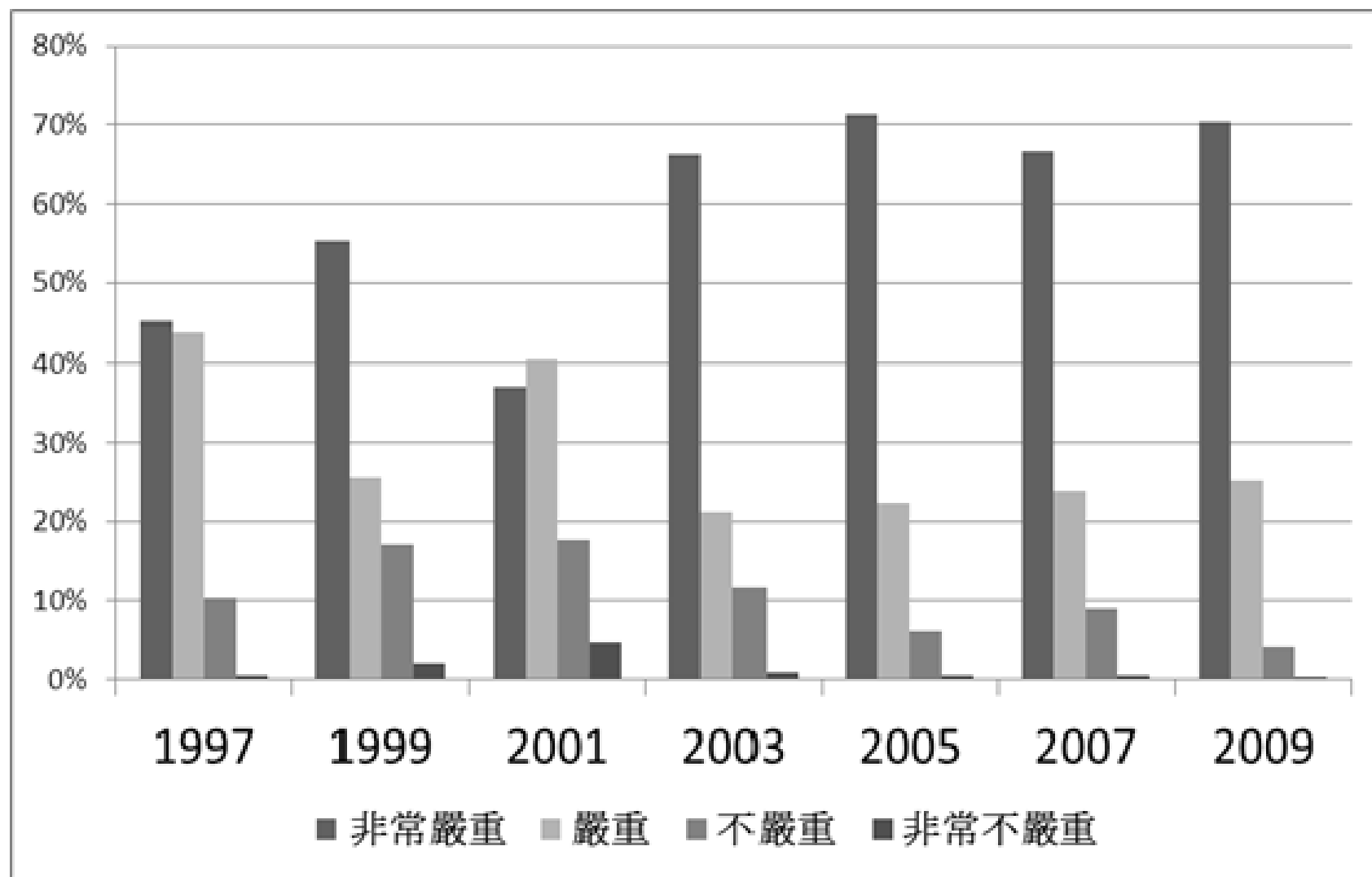


# 量化台灣社會變遷：所得分配

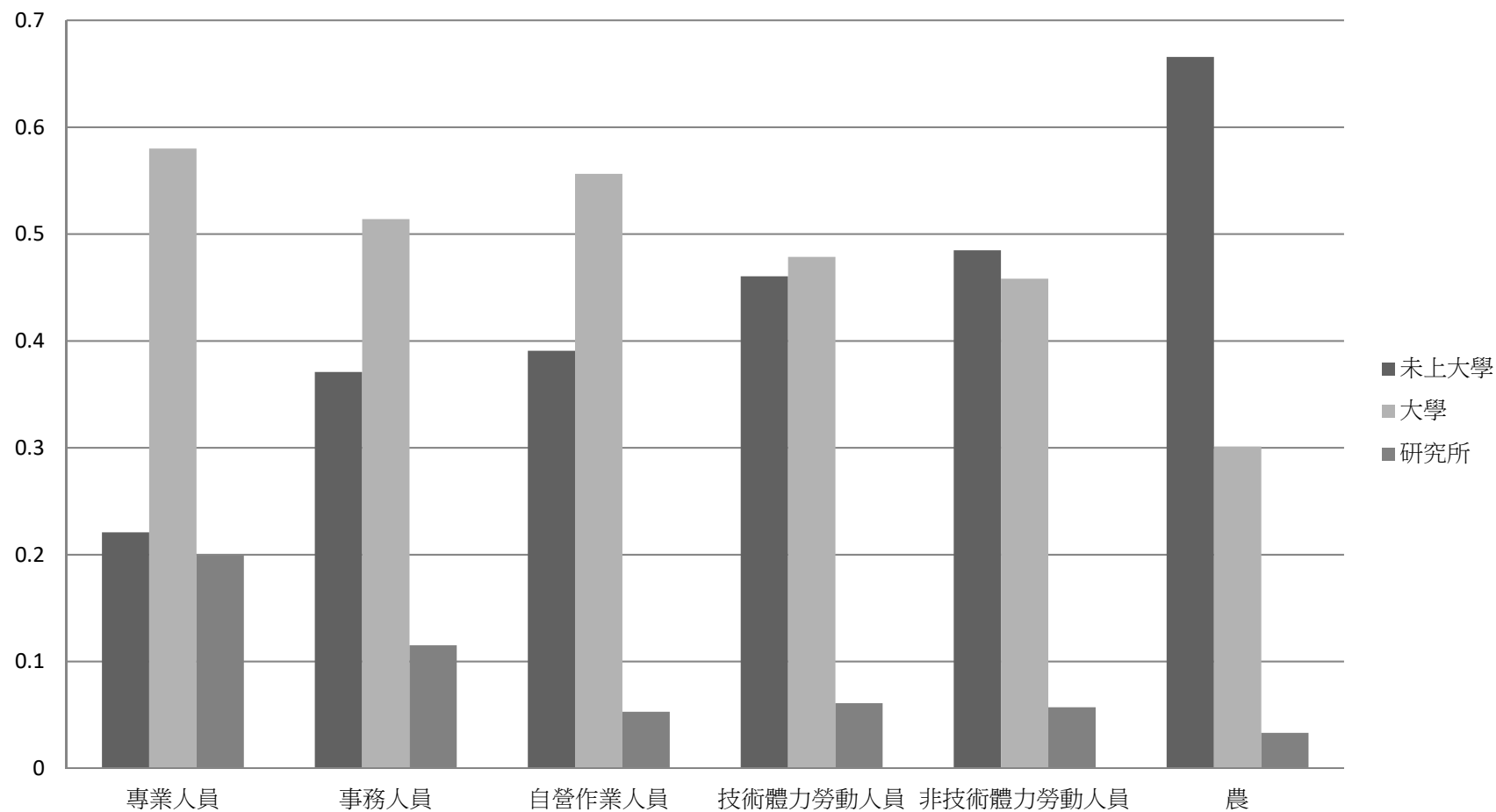
各階級受訪者的相對所得：1992年與2007年的比較，自僱者 = 1



# 對貧富差距的主觀認知，1997-2009

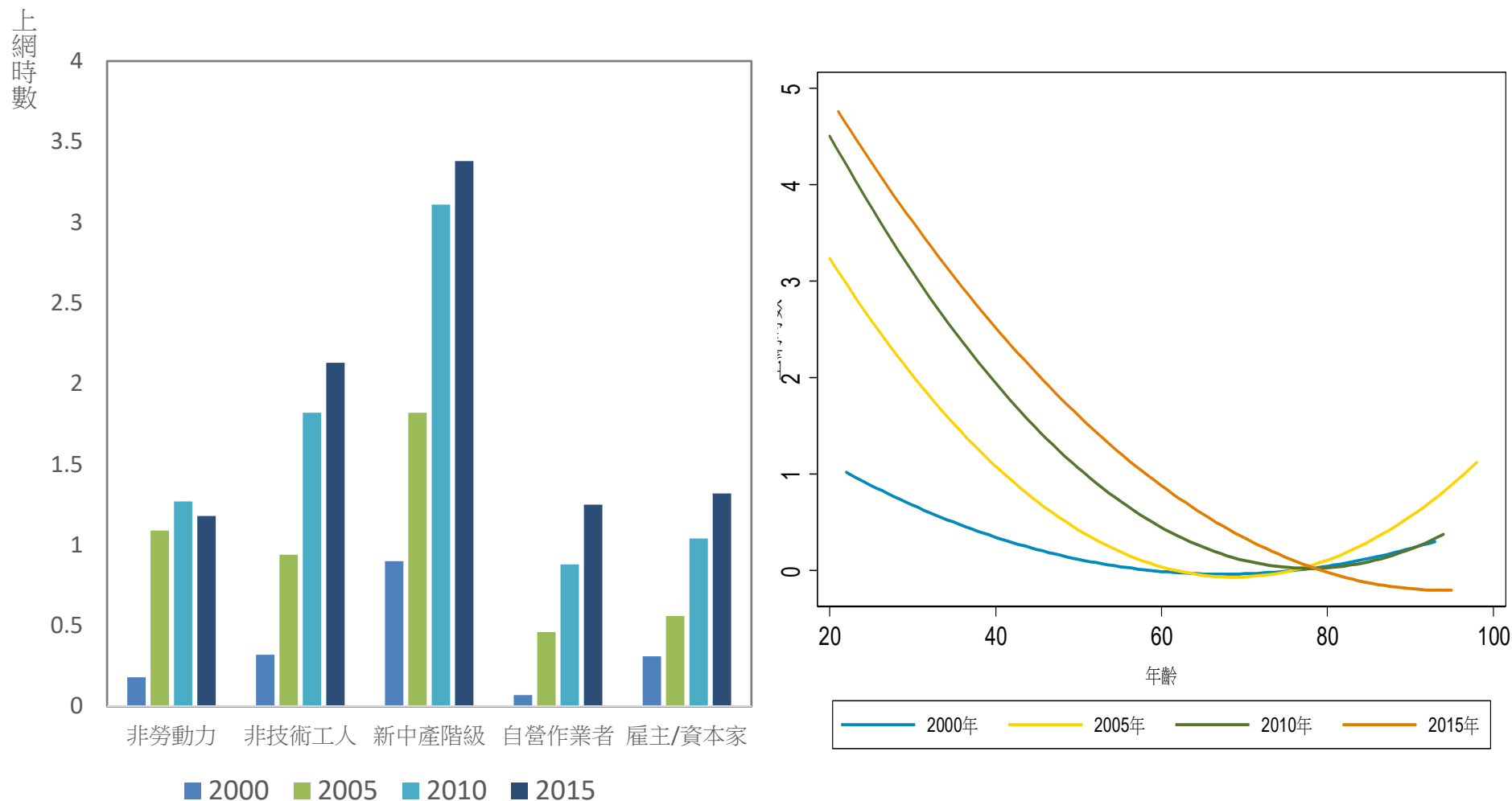


# 量化台灣社會變遷：教育分化



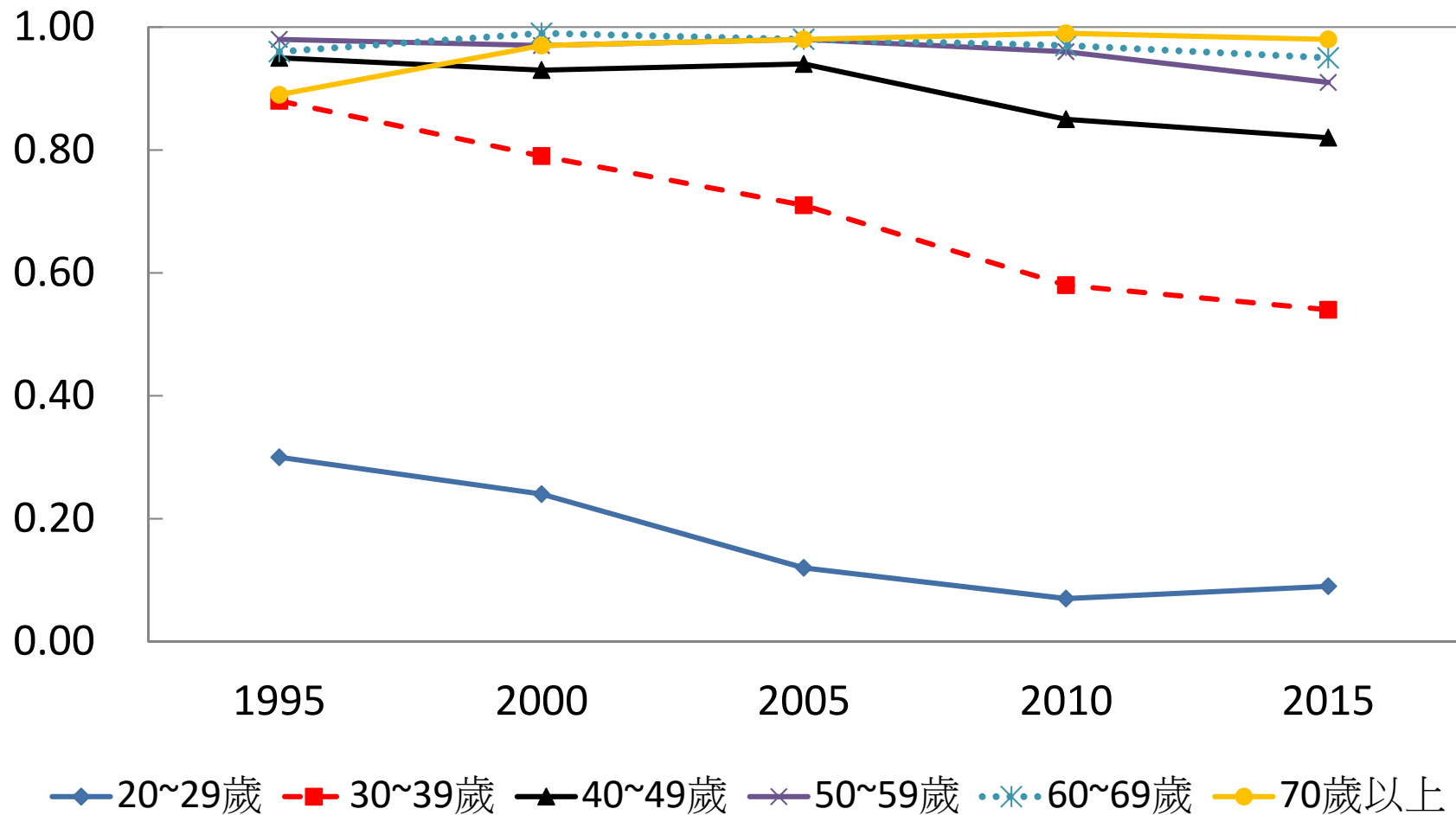
資料來源：社會變遷基本調查階層組問卷（1997年至2012年）

# 量化台灣社會變遷：數位落差



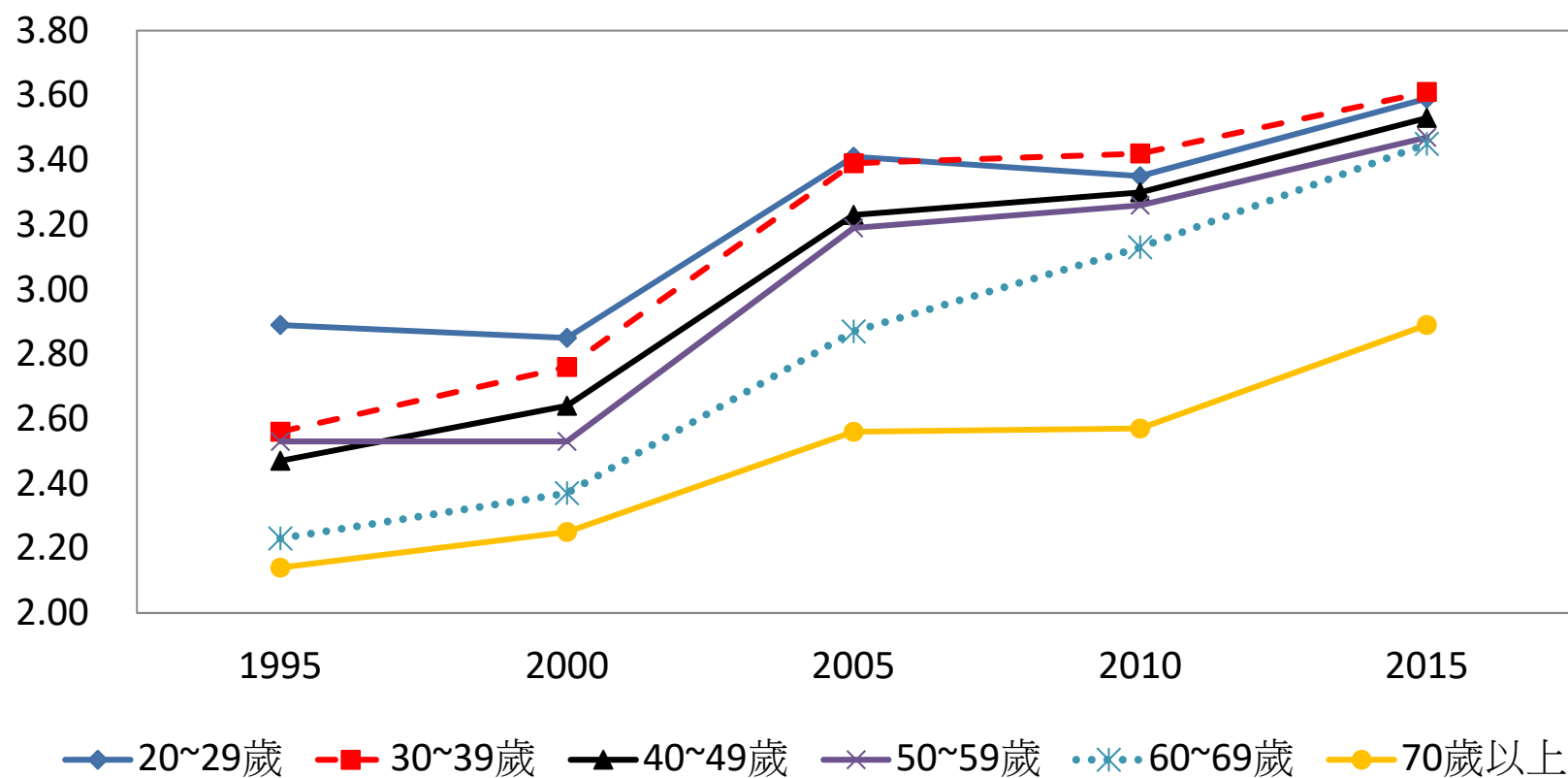
# 量化台灣社會變遷：婚姻狀態

婚姻狀態曾經已婚者所佔比例



# 量化台灣社會變遷：婚姻觀念

主要依變量：「夫妻不合而離婚」

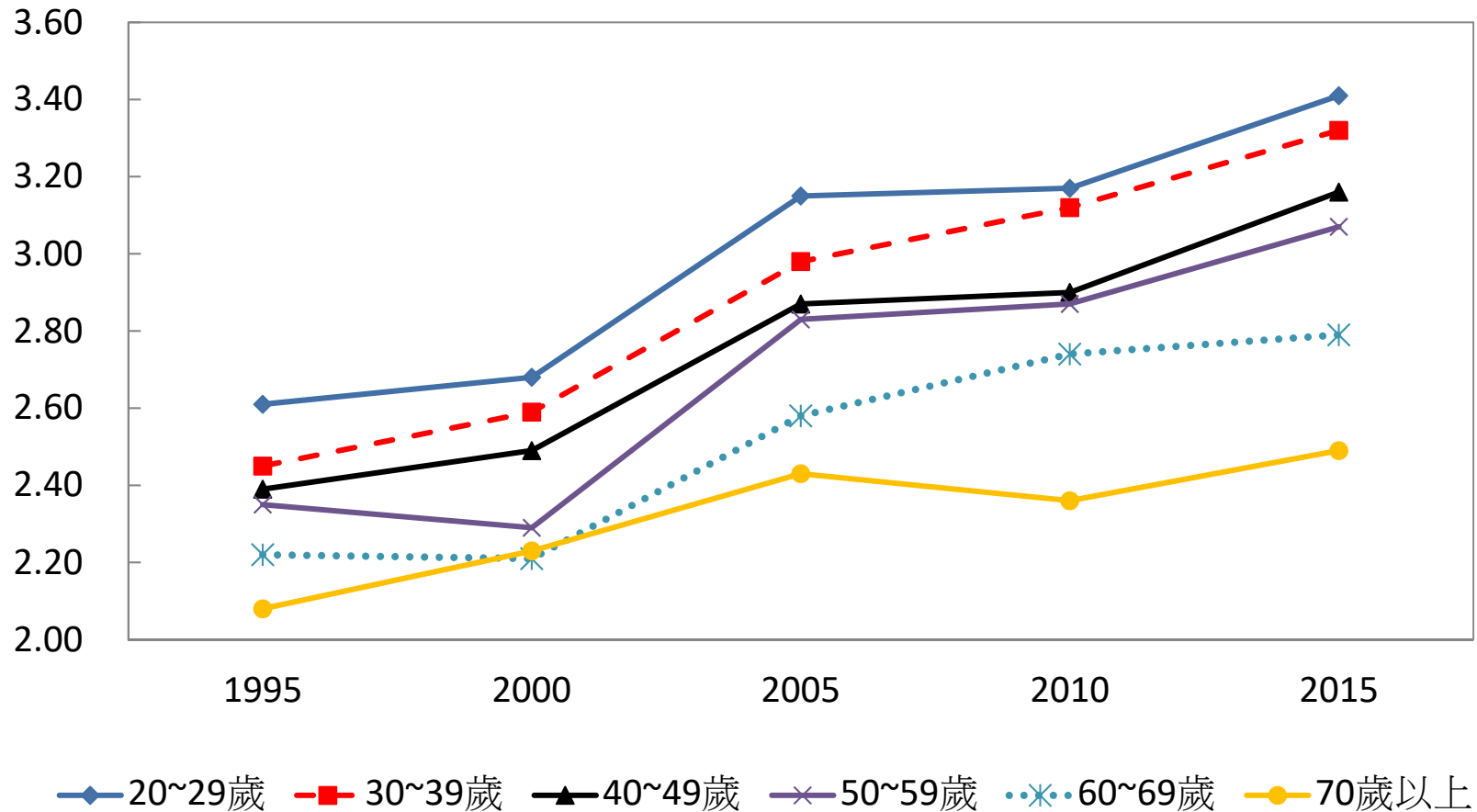


沒有錯=4, 有點錯=3, 相當錯=2, 非常錯=1



# 量化台灣社會變遷：家庭權力解組

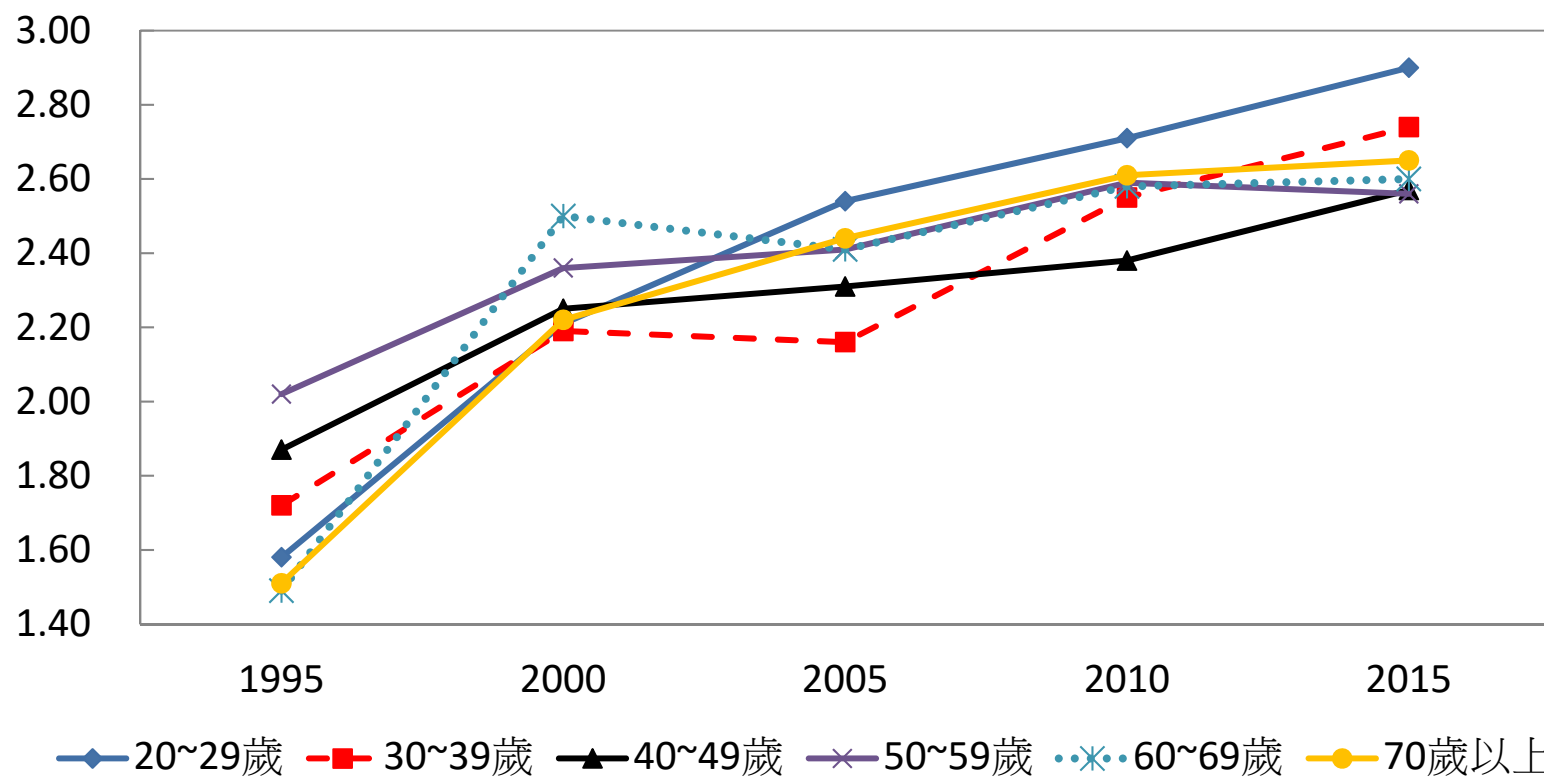
主要依變量：「婆媳意見不合爭吵」



沒有錯=4, 有點錯=3, 相當錯=2, 非常錯=1

# 量化台灣社會變遷：國家認同

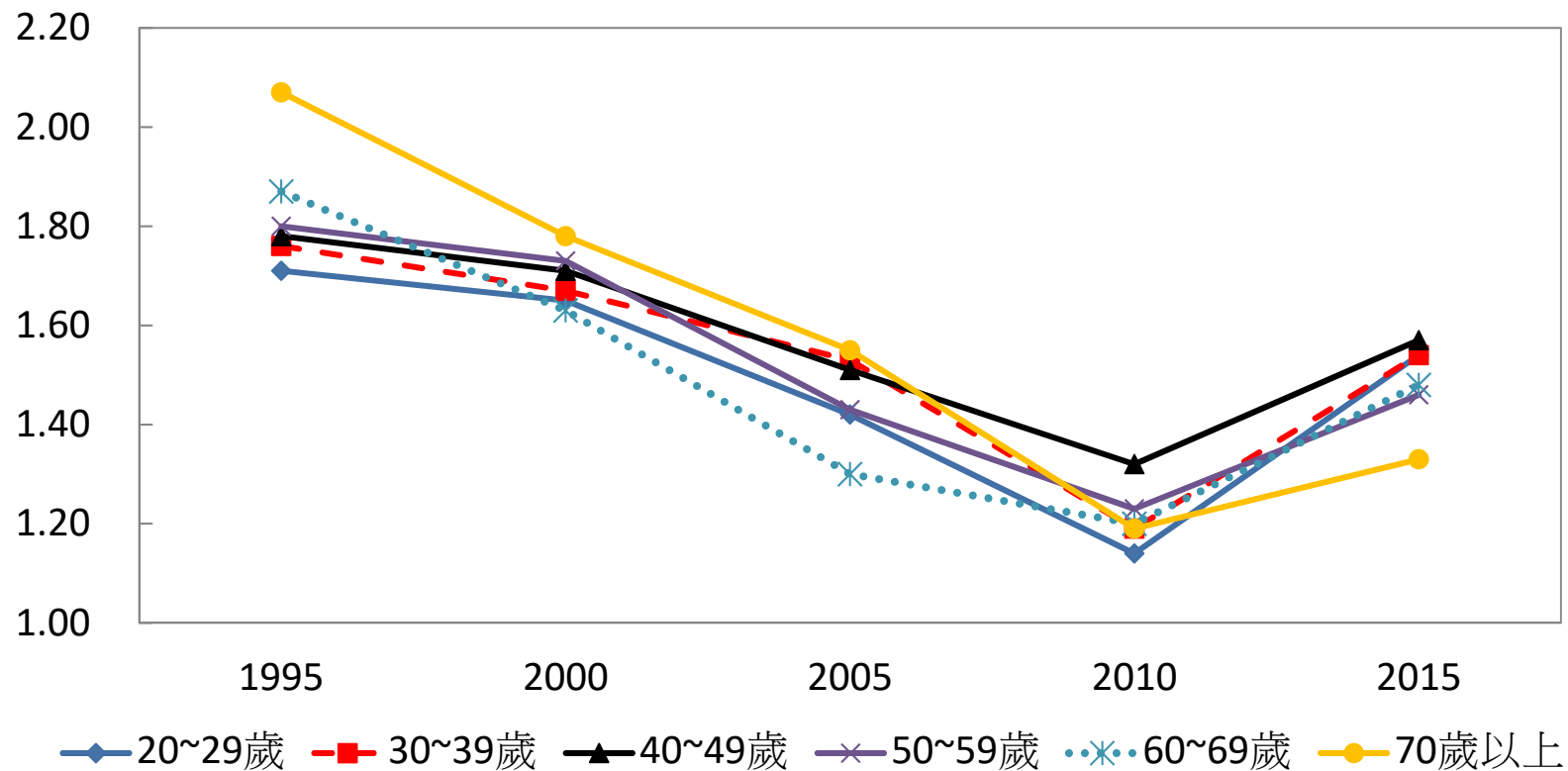
主要依變量(國家認同)：「下面有好幾種稱呼，請問您認為哪一項對您最適合」



中國人(=0)、是中國人也是台灣人(=1)、是台灣人也是中國人(=2)以及台灣人(=3)

# 量化台灣社會變遷：國家認同

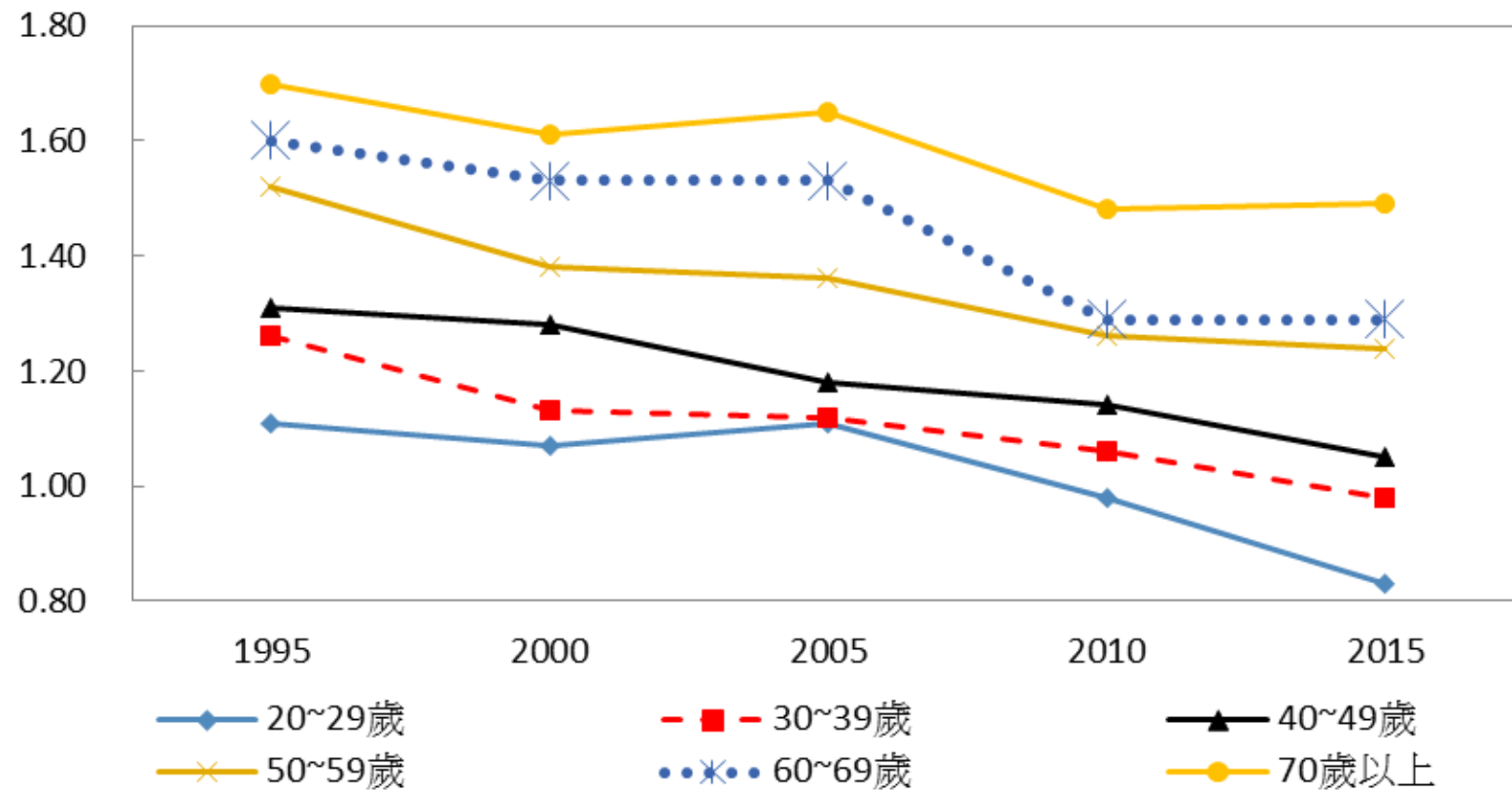
主要依變量(國家認同)：「如果大陸與台灣兩地在經濟、社會與政治各方面的條件相當，則兩岸應該統一」



非常同意=4, 同意=3, 不同意=2, 非常不同意=1

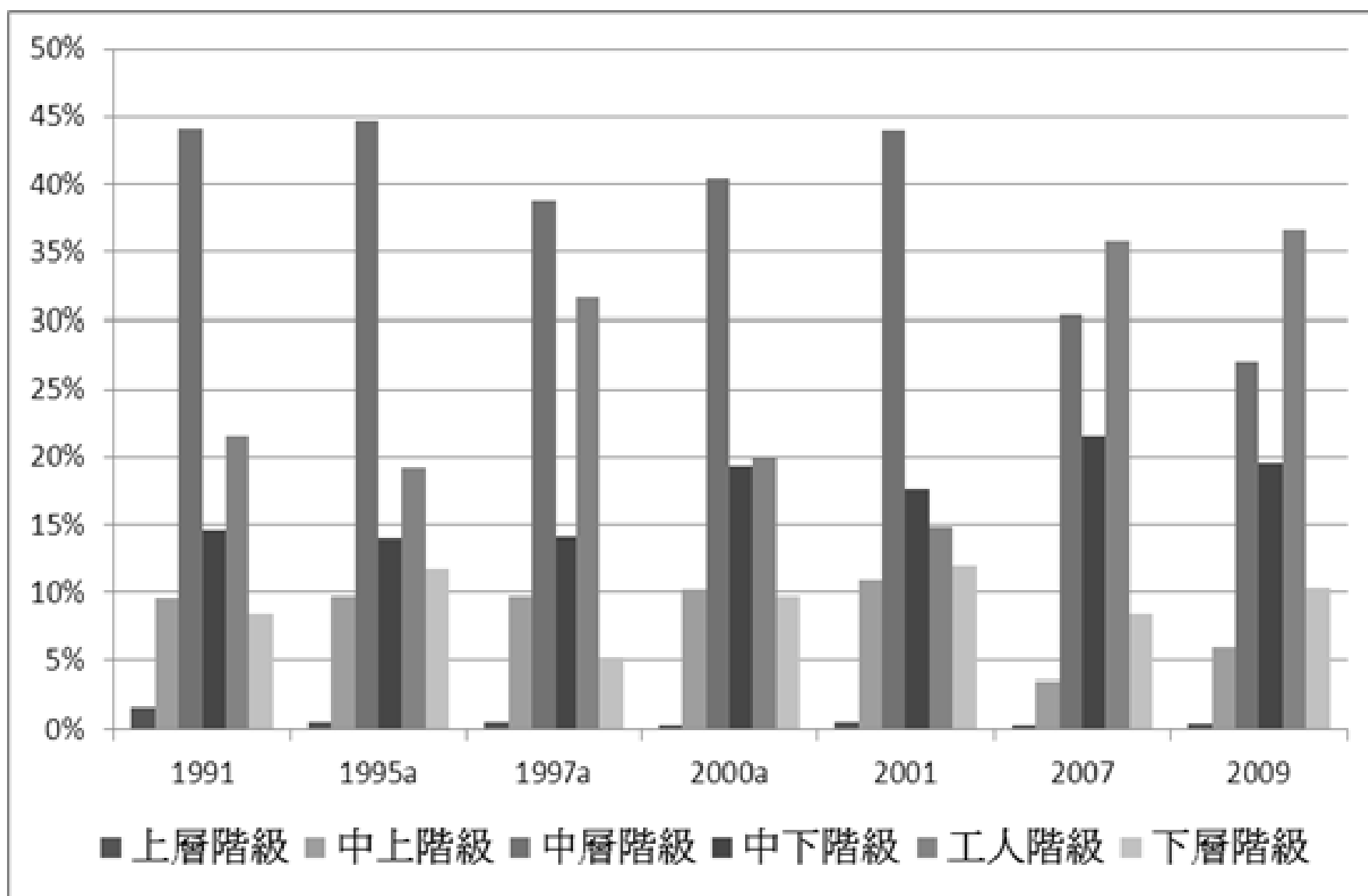
# 量化台灣社會變遷：世代與民主價值

主要依變量(民主價值)：「一種言論或意見可不可以在社會上流傳，應讓政府來決定」

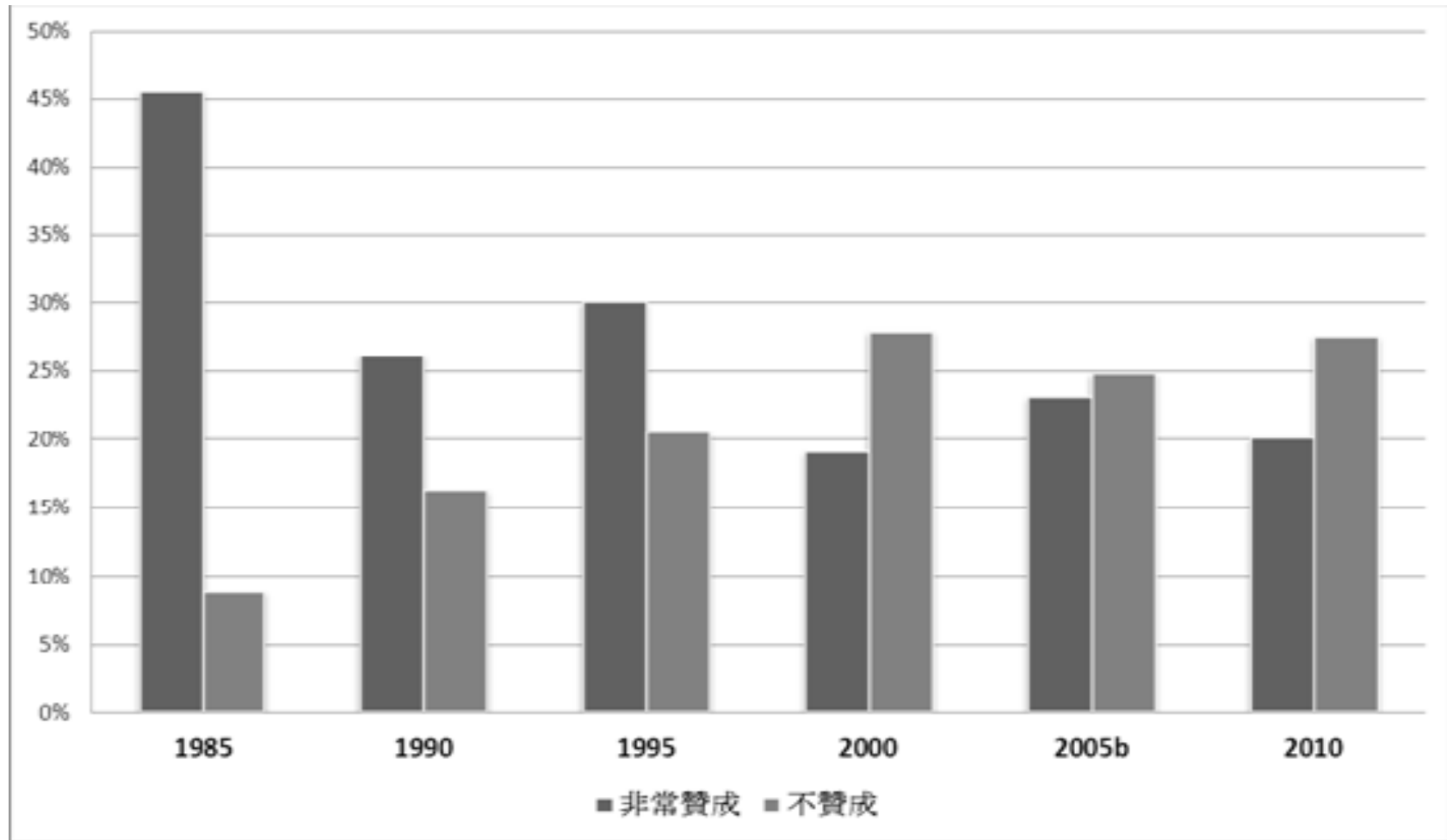


非常同意=3, 同意=2, 不同意=1, 非常不同意=0

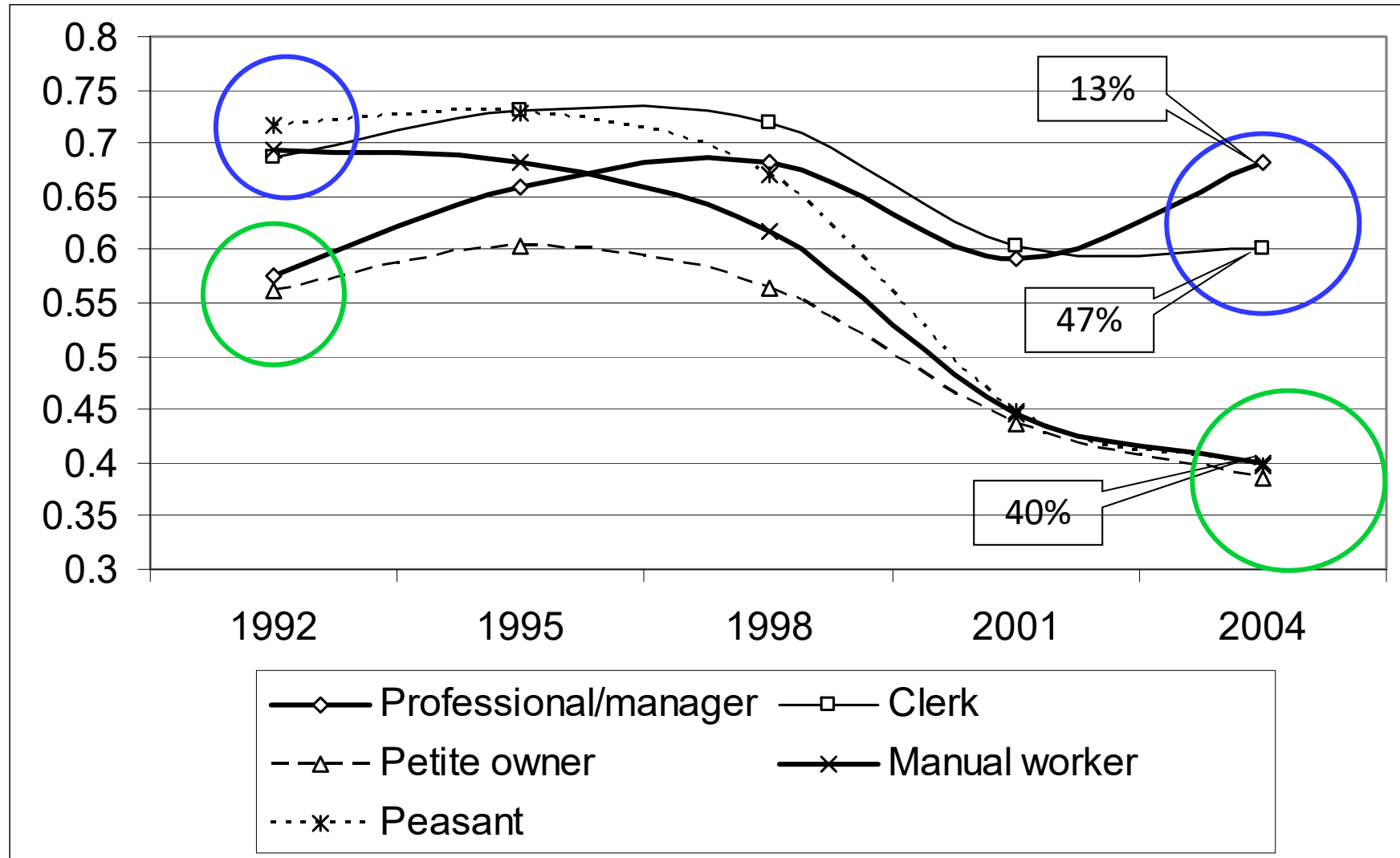
# 主觀階級認同下滑，1991-2009



# 您是否贊成 “人只要肯吃苦就一定會成功



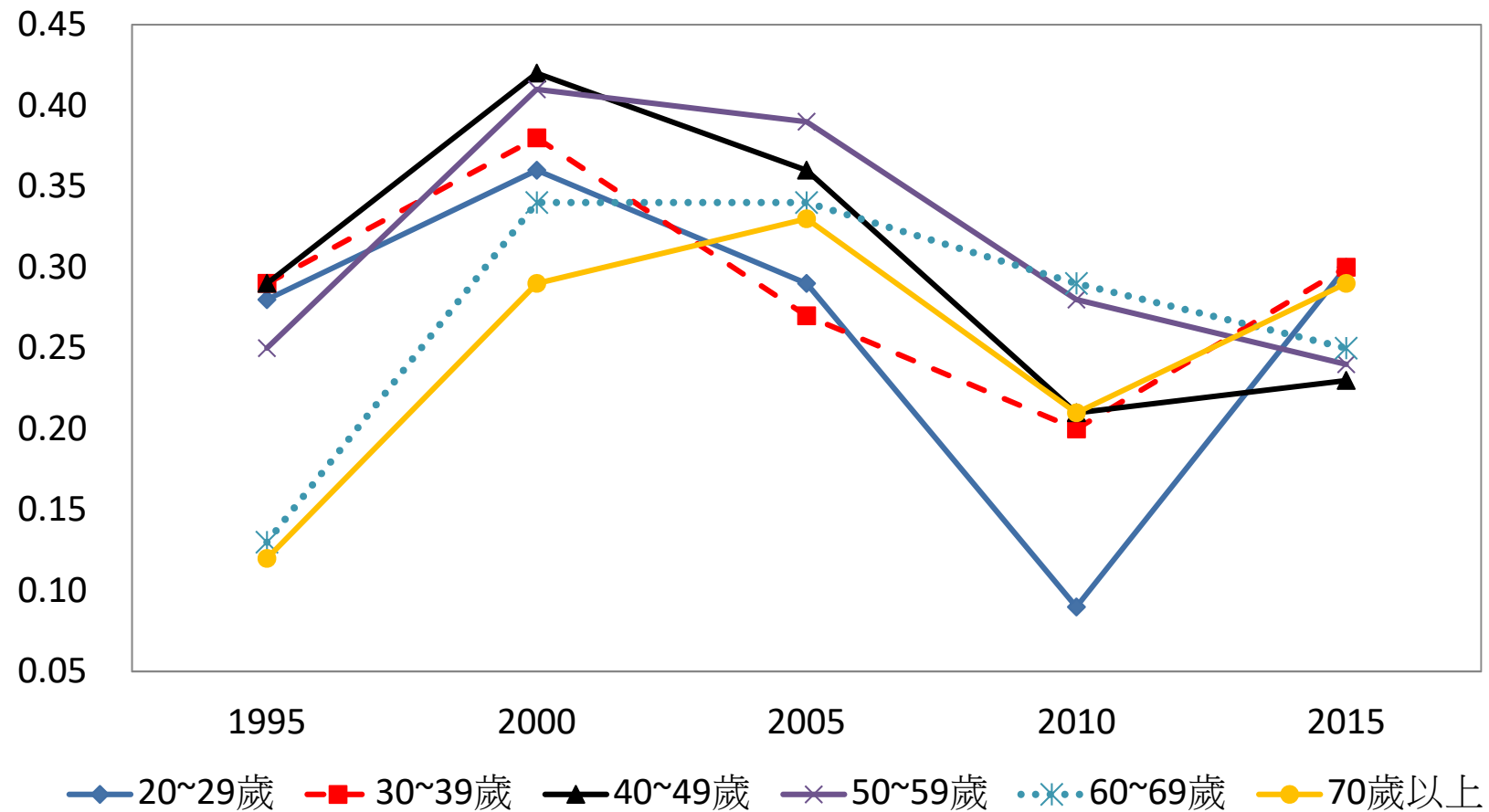
# 台灣階級投票(藍/綠結盟)的趨勢：1992-2004





# 量化台灣社會變遷：

主要依變量：「臺灣不同年齡群體支持泛綠的程度」



1=支持，0=不支持

# 多次橫斷面抽樣的分析方法與限制

- 優勢在哪裡？時間與樣本的隨機性
- 競爭對手：追蹤樣本的非隨機流失問題
- 數據劣勢：因果論證需建立在幾個假設上
  - ① 各期抽樣框架與權數的一致性
  - ② 各期問卷問項與答項的一致性
  - ③ 各期樣本身分的一致性(當成同一群人)
  - ④ 各期變項之間可能有相同的因果關係

**模型選擇：如何處理跨期研究裡的時間因素？**

# 跨期數據裡的時間

跨期合併數據裡處理時間的主要幾種手段：

1. 不理他，合併數據的好處只是提高N
2. 分期切割數據、比較同樣的模型迴歸係數
3. 在合併數據模型裡加入年代的虛擬變量
4. 在合併數據模型裡加入年代的連續變量
5. 在合併數據模型使用多層迴歸分析
6. CCREM模型(隨機效應虛擬變量)
7. CCFEM模型(固定效應虛擬變量)

# 合併數據的描述統計(1992-2017)

時期	1992		1997		2002		2007		2012		2017	
變量	平均 值	標準 差	平均 值	標準 差	平均 值	標準 差	平均 值	標準 差	平均 值	標準 差	平均 值	標準 差
個人每月所得對數	9.97	0.72	10.32	0.68	10.29	0.71	10.31	0.71	10.44	0.72	10.46	0.61
女性	0.43	0.49	0.41	0.49	0.43	0.50	0.44	0.50	0.44	0.50	0.41	0.49
已婚	0.74	0.44	0.71	0.45	0.68	0.47	0.63	0.48	0.60	0.49	0.55	0.50
教育年限	10.03	4.28	10.93	3.90	11.81	3.76	12.36	3.39	13.31	3.30	13.53	3.18
父輩族群身分												
福佬	0.75	0.44	0.78	0.41	0.77	0.42	0.75	0.43	0.79	0.41	0.77	0.42
客家	0.12	0.33	0.12	0.32	0.12	0.32	0.14	0.35	0.12	0.33	0.16	0.37
外省	0.13	0.34	0.10	0.30	0.11	0.31	0.11	0.31	0.09	0.28	0.07	0.25
受訪者階級（EGP）												
控制者	0.25	0.43	0.24	0.42	0.28	0.45	0.27	0.45	0.30	0.46	0.29	0.45
例行非體力勞動者	0.15	0.35	0.18	0.38	0.17	0.38	0.24	0.42	0.25	0.43	0.21	0.41
自雇者	0.11	0.32	0.09	0.29	0.09	0.28	0.03	0.18	0.02	0.15	0.05	0.21
農民	0.13	0.34	0.09	0.29	0.07	0.25	0.04	0.20	0.03	0.17	0.04	0.20
技術體力勞動者	0.22	0.41	0.24	0.43	0.24	0.43	0.18	0.38	0.20	0.40	0.20	0.40
非技術體力勞動者	0.14	0.35	0.16	0.36	0.15	0.36	0.24	0.43	0.20	0.40	0.21	0.40
年齡	38.12	11.06	37.81	10.48	38.52	11.09	39.69	11.28	40.84	11.32	41.35	11.93
N	1,593		2,944		1,047		1,143		1,152		1,074	

# CCREM的APC對照表(階層組)

Period Age	1992	1997	2002	2007	2012	2017	Cohort	Total
Age 20~24	146	294	126	116	82	91		855
Age 25~29	284	445	134	146	147	134	1993-1998	1,290
Age 30~34	252	487	148	157	162	125	1988-1992	1,331
Age 35~39	276	500	159	137	151	142	1983-1987	1,365
Age 40~44	203	470	162	156	151	132	1978-1982	1,274
Age 45~49	135	322	123	184	165	132	1973-1977	1,061
Age 50~54	130	177	106	121	140	133	1968-1972	807
Age 55~59	92	157	44	89	92	113	1963-1967	587
Age 60~64	75	92	45	37	62	72	1958-1962	383
		1928- 1932	1933- 1937	1938- 1942	1943- 1947	1948- 1952	1953-1957	
Total	1,593	2,944	1,047	1,143	1,152	1,074		8,953

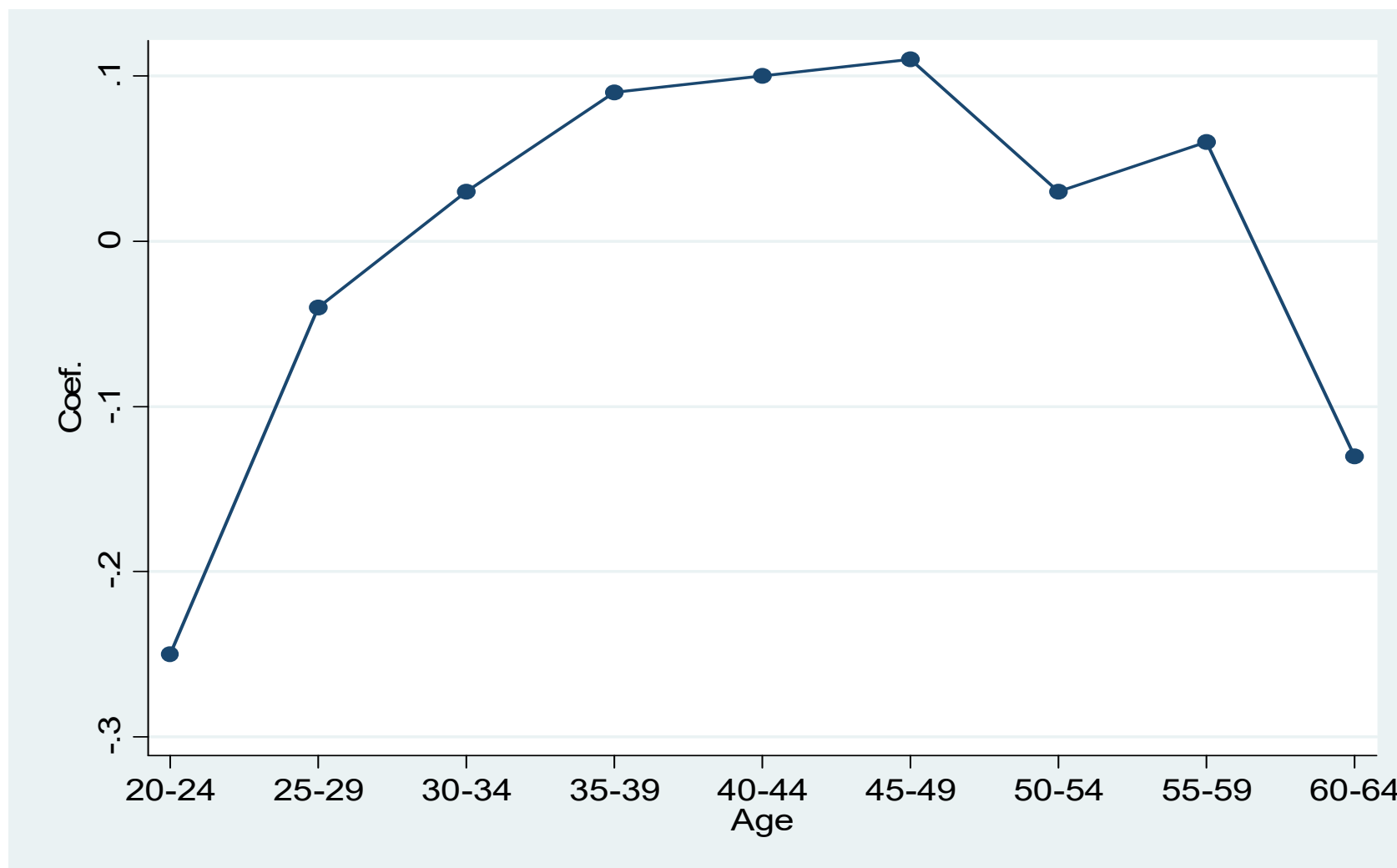
# 模型估計結果（林宗弘 2017）

	物質利益： 個人月平均收入對數 模型(1)		民主價值： 反對國家首長決定 模型(2)		身分認同： 對自己適當的稱呼 模型(3)	
固定效應						
階級位置（以非勞動力為對照組）						
非技術工人	6.353***	(74.857)	0.020	(1.095)	-0.010	(-0.422)
新中產階級	6.490***	(61.354)	0.087***	(3.785)	-0.062*	(-2.022)
自營作業者	6.134***	(55.388)	0.048	(1.948)	0.034	(1.032)
雇主／資本家	6.786***	(49.060)	0.068*	(2.235)	0.074	(1.846)
族群（以福佬族群為對照組）						
客家族群	-0.160	(-1.824)	-0.021	(-1.091)	-0.227***	(-8.666)
外省族群	0.792***	(8.203)	0.007	(0.314)	-0.820***	(-29.142)
女性	-0.674***	(-10.772)	0.003	(0.189)	0.100***	(5.410)
已婚	-0.062	(-0.621)	0.022	(0.994)	-0.008	(-0.281)
教育年限	0.138***	(14.681)	0.030***	(14.161)	-0.028***	(-9.575)
平均每天上網時間	-0.011	(-0.556)	0.010*	(2.472)	0.001	(0.114)
受政府或公營事業僱用	0.223	(1.878)	0.018	(0.686)	-0.086*	(-2.403)
反對國家首長決定					0.049***	(3.488)

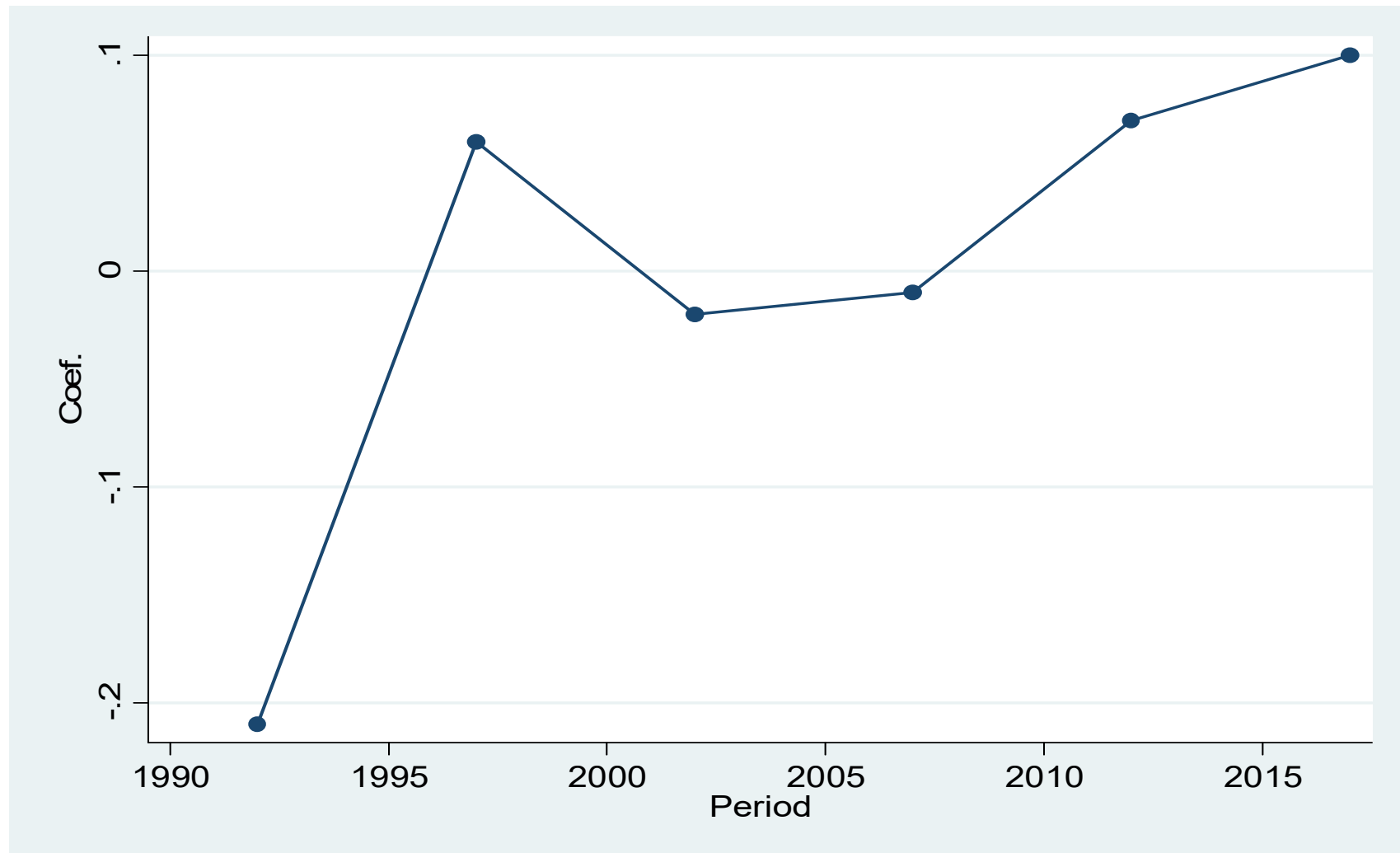
	物質利益： 個人月平均收入對數 模型(1)		民主價值： 反對國家首長決定 模型(2)		身分認同： 對自己適當的稱呼 模型(3)	
隨機效應						
20～24歲	-0.358**	(-2.743)	0.127***	(4.382)	-0.001	(-0.020)
25～29歲	0.050	(0.478)	0.083***	(3.606)	0.049	(1.602)
30～34歲	-0.070	(-0.726)	0.053*	(2.478)	0.034	(1.198)
35～39歲	-0.139	(-1.464)	0.004	(0.170)	0.057*	(2.054)
40～44歲	-0.207*	(-2.147)	-0.011	(-0.502)	0.094***	(3.322)
45～49歲	-0.079	(-0.794)	-0.030	(-1.363)	0.082**	(2.795)
50～54歲	-0.174	(-1.652)	-0.039	(-1.684)	0.035	(1.133)
55～59歲	-0.392***	(-3.482)	-0.085***	(-3.413)	0.040	(1.209)
60～64歲	-0.330**	(-2.837)	0.016	(0.630)	-0.117***	(-3.347)
65～69歲	0.500***	(3.909)	-0.030	(-1.039)	-0.125**	(-3.263)
70歲以上	1.200***	(9.926)	-0.089**	(-3.257)	-0.149***	(-4.058)
調查年份（1995）	-0.385***	(-6.187)	-0.063***	(-4.475)	-0.542***	(-28.682)
調查年份（2000）	-0.111	(-1.785)	-0.057***	(-4.012)	-0.057**	(-3.005)
調查年份（2005）	0.015	(0.244)	-0.018	(-1.329)	0.043*	(2.365)
調查年份（2010）	-0.116	(-1.881)	0.079***	(5.780)	0.239***	(13.098)
調查年份（2015）	0.597***	(8.446)	0.058***	(4.025)	0.317***	(16.478)
世代（～1925）	-1.049***	(-4.750)	-0.061	(-1.191)	-0.075	(-1.074)
世代（1926～1930）	-0.555***	(-3.454)	-0.168***	(-4.312)	-0.166**	(-3.140)
世代（1931～1935）	-0.021	(-0.148)	-0.128***	(-3.718)	0.094*	(2.000)
世代（1936～1940）	0.515***	(3.922)	-0.020	(-0.653)	0.066	(1.608)
世代（1941～1945）	0.527***	(4.229)	0.054*	(1.966)	-0.023	(-0.635)
世代（1946～1950）	0.593***	(5.017)	0.068**	(2.598)	-0.057	(-1.622)
世代（1951～1955）	0.627***	(5.414)	0.091***	(3.591)	-0.056	(-1.649)
世代（1956～1960）	0.588***	(5.257)	0.096***	(3.943)	-0.104**	(-3.198)
世代（1961～1965）	0.234*	(2.144)	0.090***	(3.776)	-0.148***	(-4.659)
世代（1966～1970）	0.251*	(2.362)	0.032	(1.391)	-0.222***	(-7.263)
世代（1971～1975）	0.156	(1.502)	0.016	(0.701)	-0.131***	(-4.379)
世代（1976～1980）	-0.018	(-0.156)	-0.021	(-0.849)	0.061	(1.877)
世代（1981～1985）	-0.210	(-1.726)	-0.059*	(-2.238)	0.187***	(5.359)



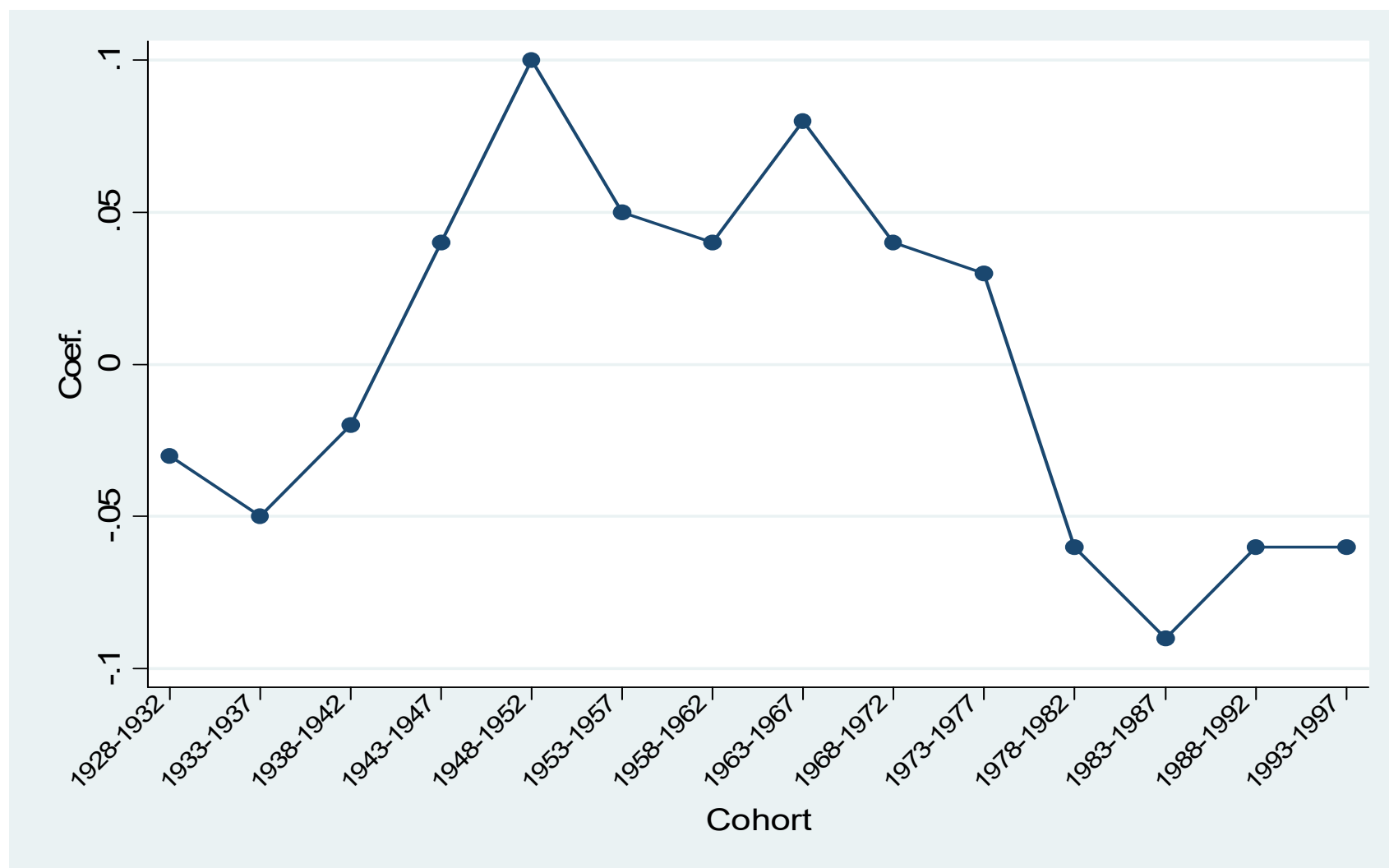
# 年齡群體與每月所得對數的迴歸係數 (張宜君/林宗弘 2019)



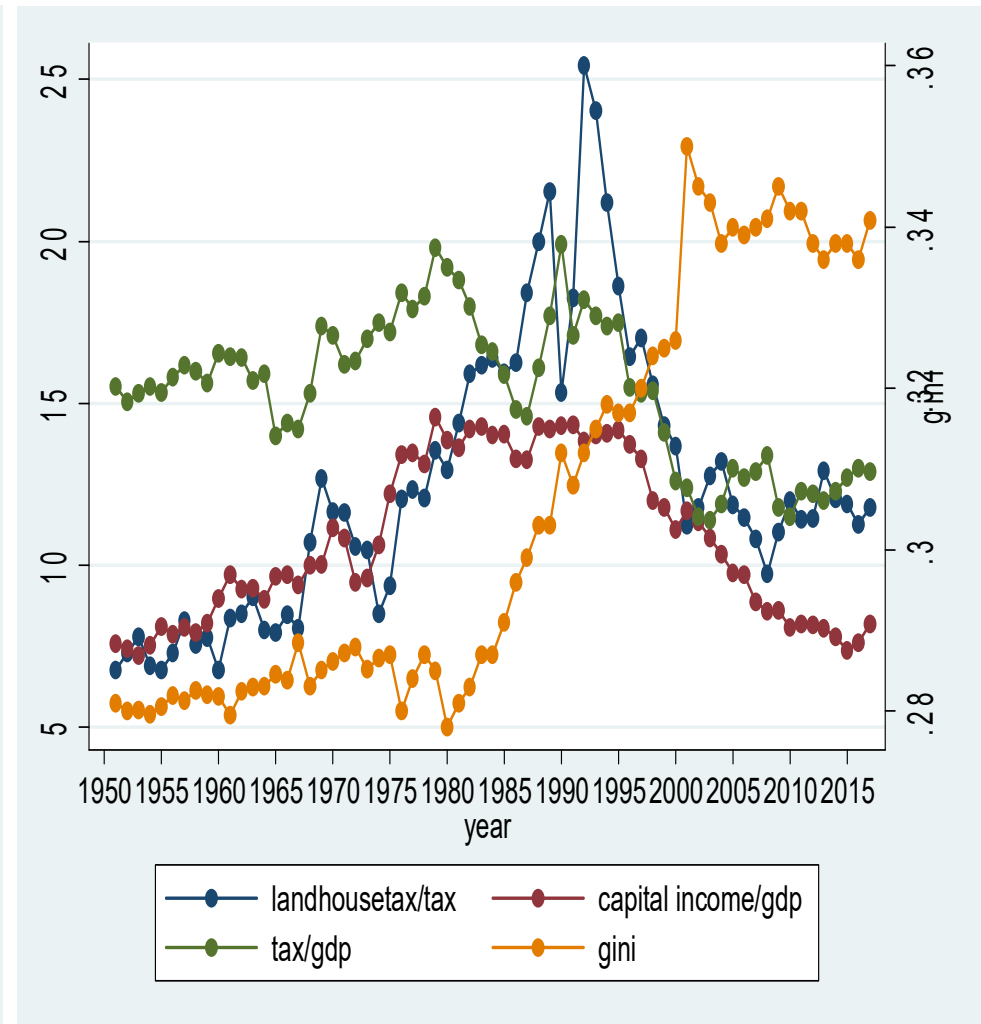
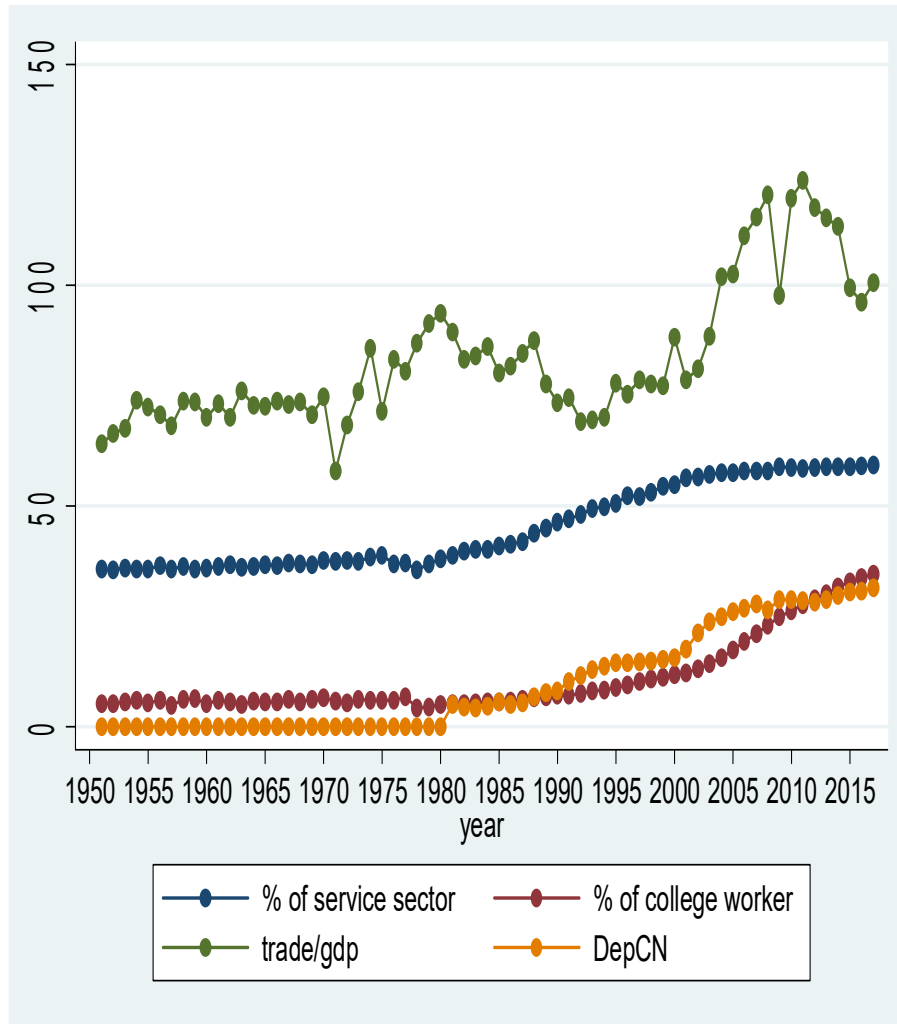
# 調查時期與每月所得對數的迴歸係數 (張宜君/林宗弘 2019)



# 出生世代與每月所得對數的迴歸係數 (張宜君/林宗弘 2019)



# 加總層級變量與個體樣本合併



# CCREM+總體層級(受訪者25)資料

世代總體變量(該世代25歲時的數據)

GDP成長率	-0.0021 (0.0026)	-0.0002 (0.0025)	-0.0012 (0.0026)	0.0009 (0.0025)	-0.0015 (0.0028)	
服務業就業人口%	-0.02** (0.01)				-0.0012 (0.01)	
大專畢業勞動者%	0.0018 (0.01)				0.0029 (0.01)	
全球貿易依賴度 (Trade/GDP)		-0.0002 (0.0025)			-0.0001 (0.0014)	
對中國貿易依賴度		-0.0080+ (0.0044)			-0.0058 (0.0076)	
所得貧富差距(Gini)			-3.10* (1.41)		-1.47 (1.77)	
土地與房屋稅% (除以總稅收)			0.01+ (0.0032)		0.01+ (0.0036)	
資本收益% (收入面除以GDP)			0.01 (0.01)		0.01 (0.01)	
國家能力 (稅收/GDP)				0.02*** (0.01)	0.01* (0.01)	
截距	9.49*** (0.03)	10.26*** (0.28)	9.60*** (0.12)	10.24*** (0.49)	9.14*** (0.10)	9.66*** (0.75)
N	8953	8953	8953	8953	8953	8953

# 合併跨期數據的使用方式：

## 給研究者的具體建議

- 當**APC**問題是研究重點(人口學或世代衝突議題)，使用**CCREM**的時機為：至少有五期以上的數據為好，最好是七期以上，否則：
- 當只有三期以下合併數據時，使用**HLM**或固定效應的時期虛擬變量
- 在4~7期之間可以使用**CCFEM**或**CCREM**，看兩者的統計差異是否非常顯著(Ken Land)
- 如果**APC**效果很不穩定或考慮放棄討論世代議題，用時期與其他變量的交互作用來呈現社會變遷

# 合併跨期數據的使用方式： 給研究者的具體建議

在理論上：呈現文獻對立典範進行假設檢驗，避免單純的數據探勘與無分析的描述

在數據處理上：尋找長期(定期)連貫的變量，有許多問題都還乏人問津(宗教/家庭價值觀)

在統計方法上發揮跨期分析優勢、避免技術錯誤：

1. 在最終模型裡盡量不要依各期切割樣本：**(a)**可以保持樣本的總量；**(b)**有助於跨時段比較
2. 當年齡與世代不是關注重點時，不必輕易動用CCREM，建議使用連續變量控制A與P，或以固定效應控制A，以HLM控制時期效果
3. 宏觀統計資料可以引入，當成層級資料或個體資料使用

# 衰退典範的一些理論問題

## 1. 變遷已經發生，因果關係卻未必明確：

- ① 內生因素：現代化的延續？後物質主義？民主轉型？寧靜的性別革命？制度路徑依賴與失敗？
- ② 外生因素：中國效應、地緣政治變遷、資本主義全球化與金融危機、新科技與環境災難
- ③ 多元理論：因果機制解釋上的不一致

## 2. 政策意義：對抗自我實現、逃離衰退典範 台灣社會變遷調查仍然是最好的資訊來源！