

家庭動態調查：樣本結構、問卷內容、資料外釋與應用

于若蓉[†] 黃奕嘉

中央研究院人文社會科學研究中心

摘 要

這篇論文的主要目的，在介紹華人家庭研究計畫在臺灣、中國大陸東南沿海省市所進行的家庭動態調查。家庭動態調查是以家庭為主要議題的追蹤調查，臺灣調查的訪問對象除 25 歲以上的成年人口外，在前述主樣本的子女滿 16 歲時，亦納入訪問樣本進行逐期追蹤調查。自 1999 年進行首次調查迄今，訪問最久的主樣本群，已完成了十七個波次的問卷訪問。堪稱臺灣規模最大、完訪波次最多的追蹤調查。至於在中國大陸上海、福建、浙江三個省市所做的調查，是以臺灣調查為藍本的追蹤調查；自 2004 年進行至今，已完成五個波次的問卷訪問。研究者除可運用臺灣的調查資料進行分析外，可與中國大陸蒐集的調查資料進行跨地域的比較研究。對於家庭、勞動、人口等領域有興趣的研究者，家庭動態調查資料是非常難能可貴的研究素材。此外，家庭動態調查所蒐集的豐富周邊資料，可配合問卷資料進行方法研究。在這篇論文中，除介紹家庭動態調查的樣本結構、問卷內容、資料蒐集方式之外，也說明如何取得資料、查詢資料的研究成果，並說明可能的研究課題。非常期盼這項珍貴的調查資料，能更廣為統計學者或其他領域的研究者應用。

關鍵詞：家庭動態調查、追蹤調查、存活分析、隨機效果模型、樣本流失。

JEL classification: C81.

[†]通訊作者：于若蓉
E-mail: yurr@gate.sinica.edu.tw

1. 前言

家庭是構成社會的核心單位，伴隨著婚姻狀態、代間居住安排、生育與子女成長離巢等變化，家庭成員的組成會隨之變動，成員間的關係也隨之變動。由於家庭的組成與分合關係人口、社會、經濟的變化，其動態變化趨勢與影響家庭動態變化的因素都非常值得關注。以家庭為核心議題的長期追蹤調查，具有學術研究與政策層面的重要價值。放眼國際，不少國家都有以家庭為核心議題的大型長期追蹤調查計畫，如美國 Panel Study of Income Dynamics (PSID)、英國 Understanding Society: The UK Household Longitudinal Study (UKHLS)、德國 German Socio-Economic Panel (SOEP)、澳洲 Household, Income and Labour Dynamics in Australia survey (HILDA) 等。¹

在傳統的華人社會中，個人作為家庭的一份子，與家庭的關係較西方社會更為緊密，且背負著結婚生子、奉養父母等義務。隨著時代變遷，臺灣的生育率降低、晚婚甚或不婚的現象愈益普遍、兩代的居住安排或資源交換等關係亦有所轉變，以家庭為核心議題的長期追蹤調查有愈益重要的趨勢。自 1998 年展開的華人家庭研究計畫，即是在朱敬一院士的號召下，召集了社會、心理、經濟、方法研究等領域的研究者共同參與，希望藉由家庭動態調查 (Panel Study of Family Dynamics, PSFD) 追蹤資料的蒐集，探討家庭的面貌及動態變化 (章英華、朱敬一，2005)。

在計畫草創階段，參與問卷設計、調查規劃的研究者除朱敬一院士外，包含于若蓉 (經濟、方法研究)、朱瑞玲 (心理)、胡克威 (社會、方法研究)、章英華 (社會)、蔡淑鈴 (社會)、簡錦漢 (經濟)、鍾經樊 (經濟) 等。在資料庫後續建置的十多年間，陸續邀請周玉慧 (心理)、張峯彬 (社會)、陶宏麟 (經濟)、鄭保志 (經濟)、蔡文禎 (經濟)、林常青 (經濟)、鄭雁馨 (人口)、李瑞中 (人口)、蔡明璋 (社會)、王盈婷 (社會)、胡立中 (社會) 等學者加入。其後幾經更迭，目前參與計畫的研究人員包括 (依姓名筆劃序)：于若蓉、王盈婷、朱瑞玲、周玉慧、胡立中、章英華、陶宏麟、張峯彬、蔡明璋、簡錦漢等。

在 1998 至 2002 年間，計畫隸屬於中央研究院經濟研究所，並由院內的調查研究工作室 (調查研究專題中心的前身) 負責田野調查。其後隨著中央研究院組織調整，計畫於 2004 年改隸於中央研究院蔡元培人文社會科學研究中心；2004 年因應蔡元培人文社會科學研究中心與中山人文社會科學研究所合併為人文社會科學研究中心，改隸於該中心。2013 年，基於中心項下僅設專題中心，不設計畫的組織調整方向，華人家庭研究計畫改隸於中心項下的調查研究專題中心。華人家庭研究計畫的經費來源，目前主要來自中央研究院。在歷年的運作中，曾得到國家科

¹以 PSID 為例，可參考李唯君 (1996) 一文的說明。

學委員會、蔣經國國際學術交流基金會、國科會社會科學研究中心等單位的經費挹注，及中央研究院主題計畫的經費補助。

在下一節中，將說明華人家庭研究計畫在臺灣與中國大陸東南沿海地區所做的追蹤調查。由於臺灣調查除針對 25 歲以上的成年人口進行追蹤訪問外，也針對其子女進行問卷資料蒐集，第 2 節將以較多的篇幅說明臺灣調查的概況。第 3 節將說明調查資料的釋出情況，及如何取得資料。在第 4 節中，將說明調查資料目前的應用情況，及在方法研究等方面的潛在應用方向。最末一節則為結論。

2. 臺灣與中國大陸調查概況說明

華人家庭研究計畫自 1999 年起展開臺灣的追蹤調查，訪問最久的主樣本群，迄今已完成十七個波次的追蹤調查。主樣本涵蓋 1935-91 年出生的臺灣民衆；除主樣本外，也將訪問對象擴及主樣本的子女，自主樣本子女滿 16 歲起納入訪問樣本，進行長期追蹤訪問。包含主樣本、子女樣本在內，目前持續追蹤的有六千多位受訪者，是國內追蹤時間最長、規模最大的追蹤調查。

在中國大陸的調查，2003 年起與中國社會科學院人口與勞動經濟研究所（以下簡稱人口所）合作，於上海、浙江、福建三個省市進行調查，迄今已完成五個波次的追蹤調查。選定上海、浙江、福建三個省市為調查地區的原因是，這三個省市位於中國大陸東南沿海地區，與臺灣的社會、文化、經濟背景較為相似，便於進行跨地域的比較研究。以下將針對臺灣調查、中國大陸調查的概況進一步說明。

2.1 臺灣調查概況說明

2.1.1 臺灣調查的樣本結構與調查頻率

臺灣調查在規劃之初，是以 25-64 歲民衆為主要的訪問對象，分三年逐步完成首波調查資料蒐集。在 1999 年進行首波調查的訪問對象是 1953-46 年出生的人口，2000 年擴及 1935-54 年出生人口，2003 年再擴及 1964-76 年出生人口。隨著追蹤調查的進行，原有的主樣本年齡逐漸增長，同時有部分樣本因健康、去世、接觸不到等因素而流失，因而於 2009、2016 年調查分別增補 1977-83、1984-91 年出生的年輕樣本。

不論哪一年啟動調查的主樣本，首波調查均是以符合出生年次的戶籍人口資料作為抽樣清冊，採用分層多階段抽樣 (stratified multi-stage sampling) 方法，於各層內採用抽取率與單位大小成比例 (probability proportional to size, PPS) 等距抽樣法逐步抽出鄉鎮市區、村里、

個人。首先依據各鄉鎮市區都市化發展程度的分層，利用內政部臺閩地區人口統計所列的鄉鎮市區人口數，依據人口比例計算各層應抽出的鄉鎮市區數，以等距抽樣法抽出鄉鎮市區；其後，在中選的個別鄉鎮市區中，依人口數多寡採等距抽樣原則抽出兩個村里；最後，對中選村里，依等距原則抽出訪問對象。²主樣本於完成第一次抽樣調查訪問後，即在後續的追蹤調查中持續進行追訪。

自 2000 年調查起，計畫將訪問對象擴及主樣本的子女，將 1977 年以後出生且滿 16 歲的子女納入訪問，並於其後進行追蹤訪問；至子女樣本滿 25 歲時，則視同主樣本，在以主樣本首波調查問卷進行訪問後，之後就隨同主樣本一併進行追蹤訪問。子女樣本不同於主樣本，不是透過隨機抽樣抽取出來，而是根據主樣本於問卷中所回答的子女訊息，於子女符合訪問年次範圍時，在徵詢主樣本及其子女同意後，進行子女的問卷訪問。由於符合訪問年次的主樣本子女不見得能順利接觸並完成訪問，並不是所有符合訪問年次的主樣本子女都能完成首次訪問並納入追蹤；此外，完訪子女樣本的首次受訪年齡也可能高於 16 歲。

對於前述的主樣本與子女樣本追蹤調查，主樣本是採每年進行追蹤訪問的方式進行追訪，子女樣本則以每兩年訪問一次的方式進行追蹤。為減輕受訪者的負擔，在 2012 年的調查完成後，主樣本改為每兩年追訪一次。在歷次調查中，對於首次訪問完訪的樣本，計畫會在後續調查中嘗試接觸並進行追訪，不會因部分波次未成功完訪而停止；只有在受訪者明確表示拒訪，或沒有任何有效聯絡資訊能與受訪者接觸，或在受訪者過世的情況下，才會停止後續訪問。

至於調查模式 (mode)，自 1999 年至 2016 年的調查，都是以面訪的方式，由接受過完整訓練的訪員接觸受訪者並進行訪問。至 2018 年調查，在經費的考量下，採取混合模式 (mixed mode) 的方式，部分樣本改採網路調查，由受訪者上網進行問卷填寫；其餘樣本仍採面訪的方式進行問卷資料蒐集。

2.1.2 臺灣調查的主樣本與子女樣本問卷

臺灣調查所使用的問卷，依樣本類別，區分為主樣本首次訪問問卷、主樣本追蹤問卷、子女樣本問卷三種。以下分別說明。

²五群主樣本首次調查的抽樣方式略有差異。其中，在各鄉鎮市區的分層部分，1999、2000、2003 年採用羅啟宏 (1992) 的分層原則；2009、2016 年則採用侯佩君等 (2008) 提出的分層方式。此外，為了避免拒訪或其他干擾因素而無法達成預定完成數，1999 及 2000 年調查依各地過去的完訪經驗，在正取樣本外，另抽出 1.5 倍到 2 倍的備取樣本；若正取樣本無法達成預定完成數，則訪問備取樣本。2003、2009、2016 年調查則採用比例膨脹原則，抽出預定完成數 1.6 倍至 2.1 倍的樣本，每個抽出樣本皆為必訪對象。抽樣設計的細節，可參考「學術調查研究資料庫」對釋出資料的相關說明 (https://srda.sinica.edu.tw/browsingbydatatype_result.php?category=surveymethod&type=2&csid=5)。

- (1) 主樣本首次訪問問卷：用於主樣本首次抽樣調查 (問卷代號 RI)，以及子女樣本轉為主樣本 (滿 25 歲) 後的第一次訪問 (問卷代號 RCI)。問卷的主要內容包含受訪者及家庭的基本資訊，主要題組包含：個人基本資料、教育經驗與最高教育程度、工作經驗與現況、婚姻及配偶資料、家庭價值與態度、親屬 (含父母、配偶父母、兄弟姊妹、子女) 基本資料及與親屬的互動關係、住宅與居住安排、家庭決策與收支、子女生育與教養、家庭關係與和諧等。

其中的教育經驗題組，以回溯的 (retrospective) 方式詢問受訪者從小學到最高學歷各階段的求學歷程。2009 年起增加教育史題組，詢問受訪者各求學階段的起迄時間。在工作經驗與工作現況部分，除了受訪當時的工作資訊外，也詢問受訪者第一份工作的內容；2014 年起，進一步詢問受訪者完成最高學歷後的第一份工作情形。在婚姻題組中，除了配偶的基本資訊及工作現況外，也詢問受訪者交往及剛結婚時的情形。父母及配偶父母題組中，除基本資料外，也詢問從事最久的工作、受訪者 16 歲時父母的工作、父母的居住安排，及受訪者與父母的聯絡情形等。藉此瞭解受訪者家人的個人基本資訊，及互動關係等進一步的資訊。

- (2) 主樣本追蹤問卷：受訪者在完成前述的主樣本首次訪問問卷後，則以追蹤問卷逐期進行追蹤訪問。在原始規劃中，這類問卷的代號以「R」開頭，後以羅馬數字標示訪問波次，如 RII、RIII、RIV ...；2005 年起，不再區分訪問波次，統一以「RR」作為問卷代號。主樣本追蹤問卷的主要內容，包括受訪者的生活與工作狀況、婚姻及配偶資料 (如配偶工作現況)、居住安排、親屬 (如子女、父母、配偶父母) 資訊、家庭決策與收支、子女生育與教養等題組。

前述核心題組納入每次的調查問卷中，藉以對相關題項進行長期資料蒐集。另有部分題組，則採單次加掛或隔數年加掛一次的方式納入問卷。歷年問卷中，曾多次加掛的題組包含：家人互動關係、家庭經濟管理、非正式教育進修情形、語言使用、宗教信仰、網絡與社會活動、健康問題、社會態度、情緒感受；單次加掛題組包含：風險展望、家人關係、借貸問題、父母經濟與分產情況、運動與健康問題 (患病憂慮、運動習慣、生活作息、生活感受)、受詐騙經驗、勞保相關題組、借貸往來、公益捐款、人格題組、家人照顧、生活態度、家務分工、性別角色與婚姻態度、算命與消災祈福、同居行為等。

為擴大參與，計畫團隊在每次進行調查問卷設計前，均對外徵求加掛題組。自外部徵求而來的加掛題組以一次性加掛為原則，若計畫團隊認為有需要，也會間隔幾年後再次加掛。公開徵求加掛題組的資訊，除了於計畫網站公告外，也透過相關學會公告。

- (3) 子女樣本問卷：適用對象為主樣 16-24 歲間的子女。訪問主題以子女樣本的求學及生活現況為主，主要題組包含：個人基本資料、教育經驗、工作經驗與現況、婚姻與配偶資料、住宅與居住安排、家庭決策與收支。其中，在教育經驗部分，與主樣本首次訪問問卷不同的是，子女樣本問卷著重於求學現況相關訊息的蒐集，且所包含的相關題項內容較廣；至於主樣本首次訪問問卷，則對各個求學階段的資訊都加以蒐集。

2.1.3 臺灣調查的執行

每次調查的田野調查執行期間，大約於當年的農曆年前 14 天展開，至同年六月底結束。訪問期間包含了農曆新年、清明節等華人傳統重要節日，希望藉此提高接觸到受訪者的機會。問卷設計等前置工作，大約在田野調查開始的一年之前展開。在問卷設計階段，由研究團隊檢視既有的問卷題組，在維持核心題組儘量不做更動的前提下，考量過去訪問結果、時代變遷等因素，進行問卷修訂。對於加掛題組，則考量可容許的問卷長度、加掛題組的重要性，及與核心題組的關聯等因素，決定加掛題組的內容。問卷初步設計完成後，則視需要進行認知訪談 (cognitive interview) 或預試 (pretest)，以測試新增或修改的題目是否恰當。為避免影響正式訪問，預試或認知訪談的對象並非取自正式訪問樣本。研究團隊會視該年問卷設計的實際情形，決定採用整份問卷或部分問卷題目進行預試，並決定以面訪或電訪進行預試。在完成認知訪談、預試之後，研究團隊再就訪談結果進行討論與問卷題目修訂，直到問卷內容底定為止。自 2014 年起，對於問卷內容、調查執行方式與流程、調查資料保管與對外提供方式、樣本名單的管理等，研究團隊均依循學術研究倫理規範，於每次調查的前置作業階段，即提送中央研究院人文會科學研究倫理委員會審查，並於田野調查執行前取得審核通過證明。

對預試、正式調查的田野調查執行工作，是由研究團隊與調查研究專題中心的調查組合作進行。對研究團隊預定進行訪問的樣本，調查組會預先郵寄訪函，於訪函中敘明訪問目的、受訪者參與的重要性、預定訪問期間、可資查證的管道、計畫與調查組的聯絡資訊等；另會寄發公文給內政部警政署反詐騙專線、地區派出所、村里長等，以便受訪者查證。訪員的召募與分派，也是由調查組負責。每次調查至少有半數以上的訪員，在過去曾參與過家庭動態調查訪問。在進行田野調查之前，會進行為期四天左右的訪員訓練，除由研究團隊講解問卷內容、問卷記錄原則外，調查組會講解訪問原則等其他訪員應注意的事項。

自 1999 年至 2016 年調查，幾乎所有問卷都是採面訪方式，由訪員以一對一訪問的方式完成。直到 2018 年，才有部分樣本改採網路自填問卷的方式進行調查，其餘的樣本仍採面訪方式進行調查。在面訪方面，1999 年至 2011 年是採用紙本問卷進行訪問；2012 年起改以調查研

究專題中心開發的電腦輔助面訪 (computer-assisted personal interviewing, CAPI) 系統進行訪問, 由訪員攜帶配備 CAPI 系統的平板電腦, 與受訪者進行訪談, 並在平板電腦上進行記錄。CAPI 系統的採用, 使得蒐集完成的資料可快速回傳, 便於進行資料檢核, 及訪員行為與訪問進度的管理; 對檢核發現問題的資料, 亦可即時與受訪者聯繫, 以補問或確認資訊。此外, 藉由 CAPI 系統問卷編輯功能, 可將跳答、問項適用年齡範圍等條件預設於系統中, 方便訪員詢問、填寫, 並減少答案不一致的情況; 對過去蒐集到的某些資訊 (如受訪者回答父親已過世), 亦可預先帶到該受訪者的系統資訊中, 便於訪員詢問或確認, 以減少對受訪者的侵擾。2018年採用的電腦輔助網路調查 (computer-assisted web interviewing, CAWI) 系統, 亦是由調查研究專題中心自行開發, 除具有與前述 CAPI 系統相同的問卷編輯功能外, 可因應網路調查的特性, 以系統寄發訪函、催收函。

為確保訪問品質, 在訪問後會抽取部分樣本, 利用電話或電子郵件進行複查 (retest); 對於難以自前述管道進行複查的樣本, 則會請原訪員以外的人員進行實地複查。複查的目的, 除了確認訪員的訪問態度是否良好、訪問是否依照規定的方式進行外, 也重複詢問問卷中的部分題目, 確認答案記錄的正確性。

除在調查執行過程中即進行資料檢核外, 在田野調查結束後, 會進一步進行資料清理, 並製作過錄編碼簿 (codebook) 等後設文件 (metadata)。資料檢核的項目包含不合理值 (illegal values) 及邏輯檢核等。除了確認每個變項的數值皆在合理範圍內, 也確保前後題項的答案符合邏輯的一致性。另會就相同樣本不同波次的資料進行比對, 以儘量減少前後波次答案不一致的可能性。

為控制追蹤樣本的流失率, 需要盡可能地在每次追蹤調查中接觸到受訪者, 並讓受訪者願意持續接受訪問。為提高受訪者的訪問意願, 在每次進行田野調查時, 均會對完訪的受訪者致贈謝禮, 以答謝受訪者的協助。在調查計畫開展的前幾年, 是對完訪的受訪者致贈禮品; 之後, 考量訪員攜帶禮品的負荷, 及多數受訪者偏好實質禮金, 改為提供一定金額的超商禮券, 作為對完訪者的謝禮。為讓接觸到受訪者的機會盡可能提升, 在每次進行田野調查時, 會請受訪者確認聯絡資訊, 並盡可能提供地址、電話、電子郵件信箱等多種聯繫資訊, 便於進行後續接觸與訪問。於兩波田野調查之間, 計畫團隊會在農曆新年及中秋節寄送賀卡向受訪者表達祝福, 並於賀卡中檢附回函, 希望聯絡資訊異動的受訪者能利用回函告知新的資訊; 在回函中, 偶爾會附加簡單的題項, 以探詢受訪者的近況。對於主樣本、子女樣本, 計畫小組建有樣本名單資料庫, 完整記錄所有樣本的聯絡資訊、歷年訪問情形、特殊需求 (如希望以電話事先聯繫) 等資訊。在保障受

訪者個人資訊、遵守研究倫理的前提下，盡可能掌握每位受訪者的聯絡資訊、過去的訪問情況等訊息，以盡量提升後續完成追訪的可能性。

2.1.4 臺灣調查的樣本完訪情況

臺灣家庭動態調查的五群主樣本，分別於 1999、2000、2003、2009、2016 年進行首波調查，完訪的樣本數分別為 999、1,959、1,152、2,182、1,973 筆。在五群主樣本完成首波調查後，後續即進行追蹤訪問。五群樣本歷次調查的完訪數，可參考圖 1。在追蹤調查中，部分樣本可能隨著訪問波次而流失 (attrition)。樣本流失的原因，包括拒訪、聯繫不到、因病無法受訪、過世等因素。1999、2000、2003 年進行首波調查的主樣本，至 2018 年調查為止，相較首波調查的完訪數，約各有四成五左右的樣本持續完訪。2009、2016 兩群主樣本因起始調查的年份較晚，流失的樣本比例分別為 35%、22% 左右。整體而言，各群主樣本在第一至二波調查間有較高的樣本流失比例 (約 20%)，其後的樣本流失情況則趨於緩和。與國外類似追蹤調查的樣本流失趨勢相似。³

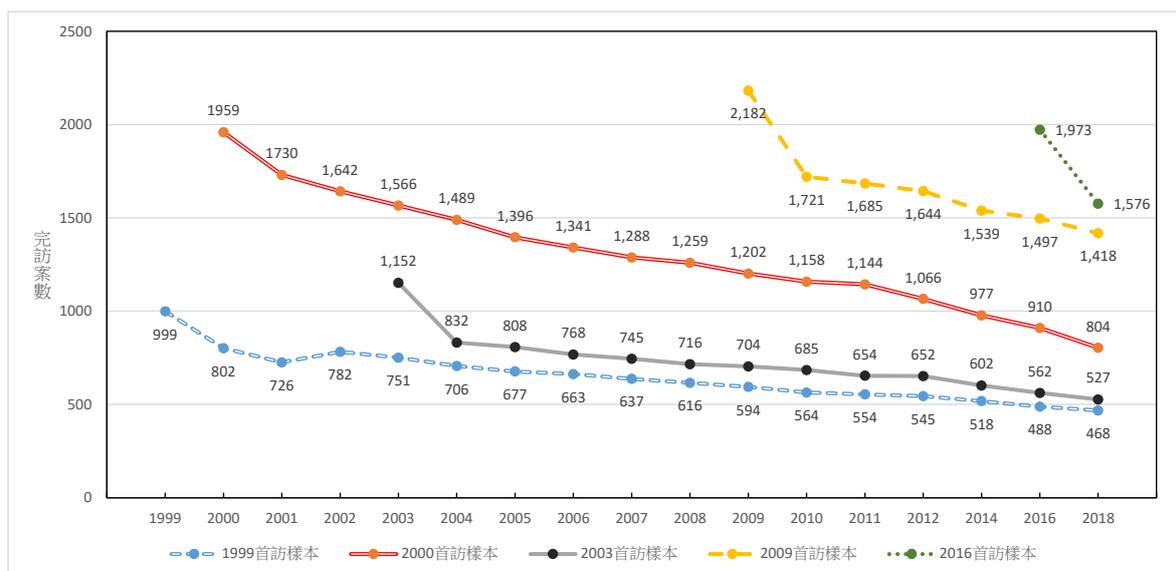


圖 1: 臺灣調查主樣本各波次調查的完訪數。

子女樣本部分，自 2000 年起，凡主樣本子女滿足 1977 年以後出生且年滿 16 歲的條件，均納入訪問對象。在子女樣本 16 歲到 24 歲階段，以子女樣本問卷訪問；滿 25 歲時以主樣本首次訪問問卷訪問；之後便併同主樣本，以主樣本追蹤問卷進行訪問。前述三類問卷的子女樣本

³國外追蹤調查的樣本流失情況，可參考相關文獻 (如 Lillard and Panis 1998) 的討論。

完訪數，可參考圖 2。子女樣本一旦納入訪問對象，便會持續進行追訪，不論其對應的主樣本是否仍然受訪。由於並非所有符合條件的主樣本子女皆能於滿 16 歲時順利完訪，對未完訪的個案，計畫仍會嘗試接觸。因此，每位子女開始接受訪問的年齡，有可能高於 16 歲。由於每次調查都有新完訪的子女樣本，同時也有流失的子女樣本，自圖 2 無法看出子女樣本流失情形。然整體而言，隨著時間累積，轉為主樣本的子女樣本有增加的趨勢，至 2018 年已達 1,470 案。

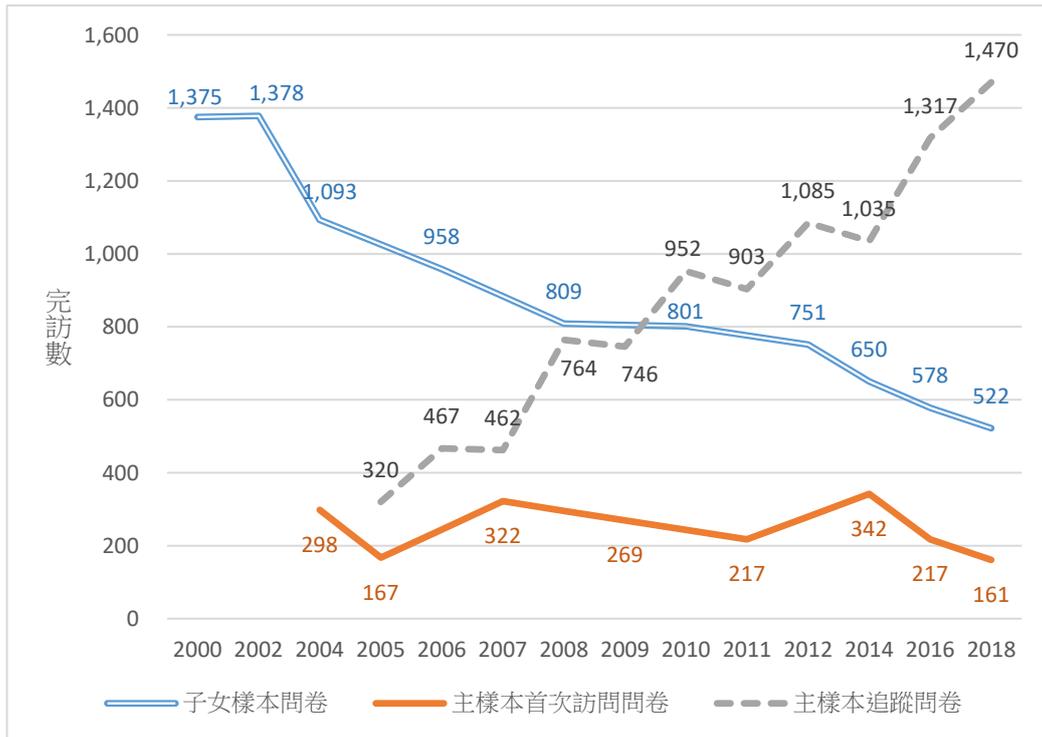


圖 2: 臺灣調查子女樣本歷次調查的完訪數: 依問卷類型區分。

2.2 中國大陸調查

華人家庭研究計畫於 2003 年，與中國社會科學院人口所合作，選定上海、浙江、福建三個省市，進行與臺灣調查相近的追蹤調查。首次調查以 2003 年滿 25-68 歲 (1935-76 年出生) 的成年人口為對象，分別於三個省市進行抽樣。首次調查原訂於 2003 年執行，因受 SARS 影響，延至 2004 年完成調查。

在問卷內容方面，中國大陸 2004 年首次調查的問卷，是以臺灣 2003 年主樣本首次調查問卷 (問卷代號為 RI2003) 為藍本，依據民情、用語不同酌予修訂，並增加中國大陸合作單位有興趣的題組。主要題組包含：個人基本資料、教育經驗與最高教育狀況、工作經歷與現況、婚姻及

配偶資料、家庭價值與態度、親屬（含子女、父母、配偶父母、兄弟姊妹）資料與互動情況、住宅與居住安排、家庭決策與收支、子女教育與養育、家庭關係與和諧、家庭經濟管理等。後續的追蹤調查問卷，則以臺灣調查的主樣本追蹤問卷為藍本進行修訂，主要題組包含：生活及工作狀況、婚姻及配偶資料、住宅與居住安排、親屬（含子女、父母、配偶父母）資料與互動關係、家庭決策與收支、網絡與社會態度、子女生育與養育、家庭價值與態度等。

首次調查的抽樣規劃，是以預定完成 5,000 案為目標，以福建、浙江各完成 2,000 案，上海市完成 1,000 案為分配原則。⁴採分層三階段隨機抽樣方法，在三個省市分別進行抽樣。依中國大陸 2000 年第五次人口普查資料，以各縣（區）及單位為第一抽樣單位（primary selection unit, PSU），計算各縣（區）級單位人口數，並以非農人口比例進行分層，⁵採等機率抽樣法抽取第一抽樣單位。其中，在福建、浙江分別抽取 22 個縣（區）級行政區；在上海市抽取 11 個縣（區）級行政區。⁶並在每個中選的縣（區）抽取 3 個鄉鎮（街道）。對每個中選的鄉鎮（街道），再抽取 3 個村委會（居委會）作為調查行政區。最後，在中選的村委會（居委會）中，以等距抽樣的方式選定受訪家戶。在每個村委會（居委會）中抽出 12 個受訪家戶，每戶以戶中抽樣方式抽出一位受訪者進行訪問，預計完訪 10 位受訪者。唯若未能達成訪問目標，則不再加抽樣本，也不以換戶的方式進行替代樣本抽樣。戶中抽樣是採 Kish 方法，在符合條件的家戶成員中，以事先設計的 Kish 表抽取中選的受訪者。

完成首次調查的樣本，原預計以每兩年訪問一次的頻率進行追蹤，唯部分年度的訪問因行政作業、在地合作單位等因素，並未完全按照規劃進行。後續的追蹤調查，分別於 2006、2011、2013、2017 年執行。在中國大陸進行調查，問卷需先通過中國國家統計局的審核。至於三個省市的田野調查，則是由人口所分別選取合適的在地單位，與在地單位合作進行調查。由在地單位負責訪員招募，人口所的研究人員負責訪員培訓；在訪員培訓完成後，隨即進行在地的面訪調查，並由人口所與在地單位合作進行訪員與調查進度管理。在每次田野調查執行期間，華人家庭研究計畫的成員會分赴三個省市進行實地視察，瞭解各省市的調查概況；並於完成在地視察後會合，討論各地的調查情況，進而與人口所的研究團隊進行會商。

截至 2017 年的調查為止，三個省市的調查都是採紙本問卷的方式進行面訪調查。完訪的紙本問卷，會匯集於人口所進行資料鍵入，及初步的資料整理與檢核，再將資料檔交付給臺灣的研

⁴完整的抽樣說明，可參考張翼（2005）。

⁵各地區因都市化程度不同，設定劃分城鄉的都市化的標準亦不同，浙江為 19.90%、福建為 17.99%、上海市則為 64.32%。

⁶浙江、福建兩地在縣（區）層級抽樣時，考慮到城鄉間的人口分配差異，將預定抽出樣本數在城鄉之間平均分配，分別抽取 11 個單位。上海市則依據人口總數比例分配樣本量，於城市地區抽取 7 個縣級單位，農村地區抽取 4 個縣級單位。

究團隊，進而由臺灣研究團隊的工作人員進行進一步的資料檢核。在對所發現的問題完成彙整後，由臺灣研究團隊人員赴人口所，就所發現的問題與問卷掃描檔進行比對、查核，以確認受訪者的答案，藉以減少問卷資料可能存在的問題。

2004 年首次調查完訪的 4,684 筆樣本，隨後續追蹤調查的進行緩步流失（各波次的完訪情況可參考表 1）。至 2016 年調查，三個省市的總完訪樣本數為 3,535 案，約為首次調查完訪數的 75.47%。在三個省市中，以上海的流失率較高，與臺灣調查在都市地區流失率較高的情形類似。

表 1: 中國大陸調查的歷年完訪情況。

調查年份	完訪樣本數				追蹤完成率 (%)			
	上海	浙江	福建	總計	上海	浙江	福建	總計
2004	903	1,856	1,925	4,684	100.00	100.00	100.00	100.00
2006	795	1,770	1,771	4,336	88.04	95.37	92.00	92.57
2011	667	1,483	1,636	3,786	73.86	79.90	84.99	80.83
2013	595	1,471	1,544	3,610	65.89	79.26	80.21	77.07
2016	574	1,445	1,516	3,535	63.57	77.86	78.75	75.47

3. 釋出的調查資料與取得方式

臺灣的家庭動態調查，在每次調查資料完成資料檢核、後設文件製作後，均透過學術調查研究資料庫 (Survey Research Data Archive, SRDA, <https://srda.sinica.edu.tw/>) 公開釋出。有意使用資料的人士，只要符合 SRDA 的會員資格，即可申請成為 SRDA 的會員，並自 SRDA 網站下載家庭動態調查資料。家庭動態調查的外釋資料，以單期間卷資料檔的方式提供下載，所提供的資料相關檔案除 SAS、SPSS、Stata 三種統計軟體格式的資料檔外，也包含問卷、過錄編碼簿檔案，及行職業代碼說明檔案等文件。至 2018 年 8 月為止，已釋出 1999 至 2014 年各期資料，2016 年調查的主樣本追蹤訪問資料亦已釋出。

中國大陸調查部分，2004 及 2006 年的調查資料，同樣可於 SRDA 免費下載。至於 2011 及 2013 年的調查資料，基於中國國家統計局的要求，無法將資料對外公開提供使用；目前是採個案申請的方式，由有意使用的研究者將研究大綱以電子郵件寄給計畫小組，經研究團隊同意後提供相關資料。

對於華人家庭研究計畫所蒐集的調查資料，使用者除可應用單期間卷資料進行分析外，更可

將相同樣本不同期次的完訪資料進行跨波次的合併，以進行追蹤調查資料分析。在資料檔中，每位受訪者都有對應的受訪者編號變項，此編號是固定不變的，不會隨調查波次改變。資料使用者可利用受訪者編號變項，將相同樣本不同波次的問卷資料予以合併。此外，對臺灣調查所蒐集的主樣本、子女樣本問卷資料，可自資料檔中的樣本戶號編碼瞭解樣本是否來自同一家戶，用以合併同一家戶的主樣本、子女樣本資料，進行親子兩代的資料分析。

對問卷中的個別題項，若受訪者「拒答」或回答「不清楚、不知道」，在資料檔中均有對應的特殊代碼。對於「不適用」於受訪者的題項（例如，受訪當時沒有工作者無須回答目前工作的行職業、收入等題項；父親已過世的受訪者，無須回答父親目前的居住安排題項），資料檔中亦有「不適用」的特殊代碼。研究者在使用資料時，需要留意問卷與過錄編碼簿的說明，以便在篩選、整理變項資料或合併跨期資料時，不致因為未能掌握資料的性質，而造成分析上的問題或分析結果的錯誤解讀。

另外，為減少受訪者的負擔，某些個人基本資訊題項（如性別、教育程度、居住地區），不見得在每波調查中都納入問卷；對於前一次訪問至當期調查未變動的題項，也可能不會追問訪問當時所處的狀態。在進行跨波資料合併時，需留意問卷內容，及問卷中的跳題說明。對於沒有變動的變項，資料使用者可能需要追溯前期的問卷資料，以獲取所需的變項資訊。⁷

為方便使用者進行跨期資料分析，降低資料使用的門檻，華人家庭研究計畫曾嘗試由研究團隊進行跨期資料合併，再將合併完成的資料檔提供研究者使用。然問卷結構複雜且題項變項數目龐大，跨期資料檔會隨著調查波次的累積而愈益龐雜，未必方便研究者使用。因此，現正積極建立資料庫資料檢索及合併系統，讓使用者可透過網站介面選取所需的變項、樣本群、調查年度，再下載系統合併完成的資料檔。目前系統已開放試用，所包含的題組受訪者基本資料、教育歷程與最高教育程度、工作經驗與現況、婚姻與配偶資訊、家庭決策與收支等題組，調查年份涵蓋 1999 至 2007 年。⁸使用者以 SRDA 帳號登入即可使用。至於其他題組及後續年度的資料，則仍在持續清理中，將視處理進度逐步開放使用。

4. 調查資料的應用與方法研究課題

為方便學者瞭解資料的應用情況，及研究者應用資料做了哪些研究，對於應用家庭動態調查資料所發表的研究，研究計畫與 SRDA 的工作人員持續蒐集、整理衍生著作清單。根據 2018 年

⁷為減少資料使用者在整理資料上的負擔，未來會在釋出的資料檔中，補入受訪者基本變項的資訊；對已釋出的資料，也會在增補相關資訊之後重新釋出。

⁸開放試用的系統網址為 <https://psfd.sinica.edu.tw/datacenter/>。

7 月的統計, 所蒐集到的著作清單計有 74 篇期刊論文、7 篇專書論文、124 篇研討會論文、180 篇碩博士論文, 並有 37 項科技部研究計畫報告。相關著作列表, 可詳華人家庭研究計畫網頁及 SRDA 網頁。⁹

為拓展資料應用, 華人家庭研究計畫每年都舉辦資料使用工作坊或學術研討會。在資料使用工作坊中, 除了介紹資料結構、問卷內容、資料合併技巧、資料取得方式外, 也邀請曾經使用家庭動態調查資料進行研究的學者分享資料使用經驗。學術研討會則透過徵稿與邀稿等方式, 透過資料應用成果的分享、交流, 讓更多的研究者瞭解家庭動態調查資料可能的研究議題。

除了用於探究家庭、勞動、人口等方面的研究議題外, 家庭動態調查資料亦可作為教學素材, 或方法研究方面的研究資源。由於臺灣的家庭動態調查資料自 1999 至 2016 年的問卷資料, 均已透過 SRDA 對外開放, 開授研究方法或資料分析課程的老師們, 可將家庭動態調查資料作為教學素材, 引領學生應用一般統計方法或追蹤資料研究方法 (如存活分析模型、random-effects 模型) 進行分析。

家庭動態調查作為追蹤調查, 有許多追蹤調查方法研究的相關議題可進行研究。例如, 樣本流失問題為追蹤調查方法與實務領域最重要的議題之一。¹⁰ 樣本流失是指, 伴隨著追蹤調查的進行, 部分樣本可能因接觸不到、拒訪等因素而無法完成訪問。在進行追蹤調查時, 為控制樣本流失率, 調查單位會多方採行各種措施, 以提高接觸到樣本的可能性, 並降低樣本拒訪的可能性。國外某些追蹤調查計畫, 在法規的許可下, 會透過與公務登記資料的連結, 以瞭解追蹤樣本是否仍然健在, 或查詢樣本是否曾更新註冊地址。另外, 調查單位會透過定期或不定期的聯繫, 以掌握樣本近期是否可能搬遷或異動聯絡資訊。為降低樣本拒訪的可能性, 除提供完訪誘因 (如對完訪者提供謝禮) 外, 也會藉由提高參與感、對拒訪可能性較高的受訪者提供不同的誘因 (如不同的訪函訴求) 等方式, 設法降低樣本流失的可能性。在方法研究方面, 樣本流失的影響原因是重要的研究課題之一。過去的研究多發現, 較年輕、工作不穩定、未婚的樣本較可能在追訪中流失。若受訪者處於中年、居住於都市地區, 拒訪的可能性較高。年紀較大的受訪者, 則較可能因健康因素而無法追訪成功。¹¹

在合併多波追蹤調查資料進行研究時, 研究者需要留意所研究的課題與樣本流失機率間的關聯。以個人在失業、就業狀態間的轉換為例, 如果失業樣本在追訪中流失的可能性較高, 即需

⁹家庭動態調查資料的衍生著作清單網址為 https://psfd.sinica.edu.tw/web/plan_04.htm。另外, 也可自「學術調查研究資料庫」的著作目錄查詢網 (<https://srda.sinica.edu.tw/booksearch.php>), 於左側功能欄的「引用資料來源」中選擇「華人家庭動態資料庫」, 進行衍生著作查詢、瀏覽。

¹⁰對追蹤調查的樣本流失議題, 有興趣的研究者, 可進一步參考相關文獻, 如 Behr and Rendtel (2005)、Hill and Willis (2001)、Lillard and Panis (1998) 等。

¹¹相關討論可參考于若蓉 (2005) 等研究。

要評估在未控制樣本流失的自我選擇 (self-selection) 機制下，對於失業、就業狀態轉換的估計結果是否會產生偏誤。因此，對任一研究課題而言，樣本流失是否構成重大問題，不見得與樣本流失率的絕對高低有關，而跟研究課題與樣本流失的關聯程度有關。

在樣本流失問題之外，追蹤調查方法研究另一頗受關注的問題是追蹤相依 (panel conditioning) 的問題。所謂的追蹤相依效應，是指受訪者在追訪中提供的答案受到先前訪問或接觸的影響。例如，對健康方面的題項，在先前調查中提供的資訊(如嚼檳榔者得到口腔癌的比例)，可能影響到受訪者的相關健康行為 (如嚼檳榔的行為)；對社會議題的態度題，可能會讓受訪者更關切相關社會議題，而影響到受訪者在後續追訪時的回應。¹²

自前述樣本流失、追蹤相依的討論可知，追蹤調查規劃、執行層面所需注意的問題較一般的橫斷面調查複雜，可進行的方法研究課題也更為多元、豐富。由於追蹤調查資料的品質與調查規劃、執行密切相關，大型追蹤調查計畫往往會投入相當多的人力、資源，進行方法方面的研究，以改善調查設計、執行，並維繫資料品質。華人家庭研究計畫自 1999 年啟動調查以來，已累積了豐富的問卷資料，及伴隨調查蒐集而來的周邊資料 (paradata)，如每次調查與受訪者接觸結果的紀錄、訪員在問卷完訪後所做的訪問評估紀錄、複查問卷資料等。為此，華人家庭研究計畫自 2017 年起正式成立方法小組，期能更有系統地探究追蹤調查方法的相關議題。未來期待能有更多有興趣的研究者，共同參與相關方法研究。

5. 結論

追蹤調查的資料蒐集與建置，相較一般橫斷面調查，更為艱辛，所需投入的人力、物力也更多。然相較一般橫斷面調查資料，追蹤調查資料有其研究上的優點。例如，藉由相同樣本在不同調查時間點所蒐集到的問卷資料，可分析樣本持續某種狀態的時間 (如失業) 或自某種狀態轉換為他種狀態 (如自失業轉換為就業) 的轉換機率。另外，由於每筆觀察資料均有時間、個體 (如個人、家戶、廠商) 兩個維度，相較橫斷面資料，追蹤調查資料更能釐清變項間的因果關係，或分析變項的動態變化。¹³

華人家庭研究計畫自 1999 年推動以來，以家庭為核心議題，在臺灣訪問最久的樣本群，完成了十七個波次的資料蒐集；訪問對象除 25 歲以上的成年人口外，也包含其家戶中滿 16 歲的子女。在臺灣所進行的家庭動態調查，堪稱臺灣規模最大、完成波次最多的追蹤調查。由於在歷

¹²對於追蹤相依的問題，有興趣者可參考相關研究，如 Das et al. (2011)、Toepoel et al. (2009)、Sturgis et al. (2009)、Warren and Halpern-Manners (2012) 等。

¹³相關討論，可參考于若蓉 (2015)。

次調查中，問卷的核心題項內容幾乎沒有變動，可就所累積的追蹤調查資料，應用存活分析、隨機效果模型等分析方法，探討個人在某種狀態（如失業）的存續期間、分析所關注的變項（如收入）的動態變化等。另外，由於在問卷中對配偶、父母、配偶父母、子女、兄弟姊妹所詢問的資訊相當豐富，可將問卷資料進一步整理成夫妻配對資料、兩代（或更多代）資料、兄弟姊妹同儕（siblings）資料，分析研究者所關心的課題（如夫妻間的家務分工、代間資源移轉、兄弟姊妹間的教育資源分配）。藉由主樣本、子女樣本的合併資料，可以深入分析親子兩代在教育歷程、婚配、工作、居住安排、家庭收支等面向的情況。

除運用臺灣的調查資料進行分析外，可與中國大陸上海、福建、浙江三個省市蒐集的追蹤調查資料，進行跨地域的比較研究。對於家庭、勞動、人口等領域有興趣的研究者，家庭動態調查資料可說是非常難能可貴的研究、教學素材。此外，家庭動態調查所蒐集的豐富周邊資料，配合問卷資料，可進行方法研究相關課題的研究。

對家庭動態調查蒐集到的珍貴資料，華人家庭研究計畫在完成每期調查資料清理後，即開放供研究者使用，期望有更多學者專家應用建置不易的資料進行研究。對這項調查有興趣的研究者，可透過華人家庭研究計畫網頁或學術調查研究資料庫的家庭動態調查資料專區，瀏覽問卷內容。對擬使用的問卷資料，在申請加入學術調查研究資料庫的會員後，即可自網站下載資料相關檔案。而計畫小組目前也正積極進行各波期資料的清理，期待所建置的資料合併系統能更完善，以降低研究者的資料使用門檻，能更方便地選取所需的資料進行研究。

展望未來，除提供更好的客製化資料下載功能外，也將持續進行追蹤調查資料的蒐集，並致力於研析追蹤調查方法，以期將研究成果回饋到調查實務，用於改善調查規劃與執行流程。除研究家庭、勞動、人口等特定領域的研究者外，非常期盼統計學者能將家庭動態調查資料用於學術研究，或作為教學素材。在資料的使用上，如發現任何問題，也歡迎隨時與計畫小組聯繫。

致謝

作者分別為中央研究院人文社會科學研究中心調查研究專題中心研究員兼華人家庭研究計畫召集人，以及華人家庭研究計畫的資深專案經理。感謝成功大學統計學系趙昌泰老師的邀稿。另感謝洪琮瑋先生在製作圖形、文件編排上提供協助。

參考文獻

[1] 于若蓉(2005)。樣本流失與勞動參與：華人家庭動態資料庫的分析。調查研究，第18期，頁

- 45-72。
- [2] 于若蓉(2015)。固定樣本追蹤資料分析。社會及行為科學研究法：資料分析，瞿海源、畢恆達、劉長萱、楊國樞主編，第七章，269-314 頁，臺北：東華書局。
- [3] 李唯君(1996)。收支動態長期追蹤研究的起源與發展。調查研究，第2期，頁159-178。
- [4] 侯佩君、杜素豪、廖培珊、洪永泰、章英華 (2008)。臺灣鄉鎮市區類型之研究。調查研究—方法與應用，第23期，頁7-32。
- [5] 章英華、朱敬一 (2005)。家庭動態資料庫簡介。「華人家庭動態資料庫」學術研討會，臺北：中央研究院經濟研究所，2005年12月2-3日。
- [6] 張翼(2005)。2004 年家庭變動調查的方案設計、組織實施與數據初評。「華人家庭動態資料庫」學術研討會，臺北：中央研究院經濟研究所，2005年12月2-3日。
- [7] 羅啟宏(1992)。臺灣省鄉鎮發展類型之研究。臺灣經濟月刊，第190期，頁41-68。
- [8] Behr, A., E. Bellgardt, and U. Rendtel (2005). Extent and Determinants of Panel Attrition in the European Community Household Panel. *European Sociological Review*, 21, 489-512.
- [9] Das, M., V. Toepoel, and A. van Soest (2011). Nonparametric Tests of Panel Conditioning and Attrition Bias in Panel Surveys. *Sociological Methods & Research*, 40, 32-56.
- [10] Hill, D. H., and R. J. Willis (2001). Reducing Panel Attrition: A Search for Effective Policy Instruments. *Journal of Human Resources*, 36, 416-438.
- [11] Lillard, L. A., and C. W. Panis (1998). Panel Attrition from the Panel Study of Income Dynamics: Household Income, Marital Status, and Mortality. *Journal of Human Resources*, 33, 437-457.
- [12] Sturgis, P., N. Allum, and I. Brunton-Smith (2009). Attitudes over Time: The Psychology of Panel Conditioning. in *Methodology of Longitudinal Surveys*, edited by Peter Lynn, pp. 113-126, New Jersey: Wiley.

- [13] Toepoel, V., M. Das, and A. van Soest (2009). Relating Question Type to Panel Conditioning: Comparing Trained and Fresh Respondents. *Survey Research Methods*, 3, 73-80.
- [14] Warren, J. R., and A. Halpern-Manners (2012). Panel Conditioning in Longitudinal Social Science Surveys. *Sociological Methods & Research*, 41, 491-534.

[Received September 2018; accepted September 2018.]

The Panel Study of Family Dynamics: Samples, Questionnaires, Data Dissemination and Applications

Ruoh-Rong Yu[†] and I-Chia Huang

Research Center for Humanities and Social Sciences, Academia Sinica

ABSTRACT

The purpose of this study is to introduce the Panel Study of Family Dynamics (PSFD) survey conducted in Taiwan and the southeastern coast of China (Shanghai, Fujian, and Zhejiang). The major theme of the PSFD project is family issues. The main respondents of the Taiwan PSFD survey are adults aged 25 and older. Once the children of the main respondents reach age 16, they are included in the respondents. Since 1999, seventeen waves of survey data have been collected for the initial group of main respondents of Taiwan. As to the PSFD survey conducted in China, five waves of data collection have been completed since 2004. The survey data collected in Taiwan and China can be jointly analyzed for cross-regional comparison purposes. The panel data collected are highly valuable in research fields such as population, family, and labor studies. In addition, the paradata collected along with the survey data can be used to explore research topics on survey methodology. In addition to introducing the sample structure, questionnaires, and data collection practice, this study also introduces how to get access to data and search for publications based on the data. Potential applications of the data are also discussed. Through this study, the authors expect that the PSFD data can be more widely applied by statisticians and other researchers.

Key words and phrases: Panel Study of Family Dynamics, panel study, survival analysis, random-effects model, sample attrition.

JEL classification: C81.

[†]Corresponding to: Ruoh-Rong Yu
E-mail: yurr@gate.sinica.edu.tw