

運用大數據分析台灣不動產市場 -實價登錄的應用

經濟研究處

俞欣榮

114 年 1 月

摘要

我國不動產實價登錄自 2012 年 8 月起實施至今已累積大量資訊。此資料庫除了提供成屋、預售屋與租賃的交易價格外，亦提供每一筆不動產交易之座落地點、主要用途、建物型態等詳細內容。有鑑於總體經濟及金融穩定和不動產市場高度相關，加以充實不動產統計資料是政策分析中最重要的基礎，本文嘗試以實價登錄之開放資料為基礎，詳述該資料庫檔案的命名規則、欄位資訊等相關格式，同時針對各期公布資料進行清理。藉由與運用實價登錄編製房價指數之政府(臺北市、內政部)與民間單位(信義房屋)的篩選標準比對，本文初步得出與上述單位相近的統計結果，有助於即時掌握住宅、預售屋交易與高價住宅交易的價量情勢。

本文主要貢獻及政策意涵如下：

- (一)本文建構之成屋與預售屋價量資料，可做為不動產統計分析的額外來源，以更精確掌握不動產市場的交易動態。
- (二)學者研究發現：基於媒體係市場傳遞訊息並影響市場情緒的重要管道，本文建立之實價登錄資料庫能有助檢視媒體報導訊息之正確性與合理性。
- (三)本文清理與彙整之實價登錄價量資料，未來可結合不動產相關放款統計及聯徵中心提供之相關數據，量化評估央行選擇性信用管制相關措施成效與妥適性。

目錄

壹、	研究動機.....	1
一、	實價登錄制度的源起與法源依據	1
二、	現行有關實價登錄資料庫的應用	2
貳、	內政部實價登錄資料庫簡介.....	6
一、	實價登錄揭露資料之範圍與檔案格式	6
二、	實價登錄揭露資料之檔案命名原則	7
參、	研究架構：資料庫篩選原則與資料清理流程	10
一、	實價登錄資料下載與清理工作流程	11
二、	本研究與其他單位資料篩選比對	13
伍、	結論與政策意涵	24
參考文獻.....		26
附錄、	本研究與其他單位資料篩選比對表	27

運用大數據分析台灣不動產市場-實價登錄的應用

壹、研究動機

地政為庶政之母，對於多數人民而言，購買住宅往往是一生中最昂貴的消費；從金融面觀之，房地產為金融市場中的主要擔保品與資產，而房貸利息亦為固定收益中的重要來源，因此不動產的交易資訊是影響人民權益與金融穩定重要的環節。基於及時且完整的統計資訊是研擬政策分析的基石，不動產相關交易資訊之量與質，應配合整個社會、經濟發展需求而因應調整，其品質良窳除了作為人民購屋置產的決策參考外，亦為政府部門擬定相關措施的之重要依據。

根據住宅法第 47 條之一指出：「為引導住宅市場發展，主管機關應定期蒐集、分析及公布住宅資訊」。在實價登錄制度實施前，僅有信義房價指數、國泰房地產指數等民間單位提供之房價資訊，惟其皆受限於編製範圍、市場類別與母體樣本，因此渠等房價指數僅反映部分不動產成交的相關資訊。以信義房價指數為例，儘管其資料蒐集自 1991 年第 1 季迄今，惟資料來源為該公司仲介成交資料，且樣本篩選標準僅限於公寓、華廈及電梯大樓，並排除預售與偏差產品如工業住宅等，故信義房價指數的編製只能呈現部分成屋的成交價格走勢，無法運用此房價指數推論預售屋或高價住宅的交易動態。因此，若能擁有完整的交易檔案及相關數據，則能運用當前資料探勘的方式建立不動產價量資料庫，以更清楚掌握當前住宅市場的動向。

一、實價登錄制度的源起與法源依據

全球金融危機後，臺灣房市量能自 2009 年起快速回溫，部分地區的房地短期交易頻繁，並帶動房價上漲。為因應高房價引發的民怨，避免房價泡沫引發的系統性風險，我國政府於 2010 年 4 月推出「健全房屋市場方案」，以促進房地產業發展、穩定國內金融與減輕民眾購屋負擔為主要目標。在此背景下，內政部為促進不動產交易資訊透明化、降低資訊不對稱，透過《平均地權條例》、《地政士法》、《不動

產經紀業管理條例》三法修正推動「不動產成交案件實際資訊申報登錄」(簡稱實價登錄)制度，於 2012 年 8 月 1 日開始實施，同時於該年 10 月 16 日起開放查詢，以「區段化」，「去識別化」方式揭露交易資訊，如主要用途、地點、建築完成日期、建物型態¹與交易面積等內容，涵蓋不動產(含房地或房地+車位、純土地、純建物、車位等)買賣、租賃與預售屋買賣資訊。2021 年 7 月進一步實施實價登錄 2.0，除將門牌、地號完整揭露外，並要求預售屋交易即時登錄，責令預售屋買賣之申報時間，由「委託代銷契約終止後 30 日內」提前至「簽訂預售屋買賣契約書後 30 日內申報」，同時明定預售屋的定型化契約報請地方政府備查、禁止紅單轉售等措施，以保障消費者之權益。在經歷十多年的資料累積，實價登錄施行至今的可查詢筆數已達數百萬筆，完整的不動產公開交易資訊不只成為人民購屋決策的重要依據，亦為政府機關在不動產政策分析上提供寶貴的資料。

二、現行有關實價登錄資料庫的應用

(一) 房價指數的編製

表 1 初步整理各類國內房價指數的編製的方法與資料來源。可以發現，自 2012 年實價登錄資料公布至今，已有不少政府單位及民間機構運用實價登錄資料根據其需求與目的編製房價指數，彌補部分民間企業運用其仲介成交資料產生的篩選偏誤。以政府機關為例，行政院 2011 年 10 月 26 日核定「民國 101 年至民國 104 年整體住宅政策實施方案」，內政部遂以實價登錄資料編製住宅價格指數並定期發布作為該資料庫的具體應用，成為衡量不動產市場價格變動的參考指標²之一；2013 年 7 月起臺北市政府亦根據實價登錄揭露之買賣案件，依據「特徵價格法」編製公寓、大樓與 15 坪以下住宅(小宅)的價格指數；2019 年 11 月起安富財經與國立清華大學安富金融工程研究中心

¹ 建物型態包含公寓(5 樓(含)以下無電梯)、華廈(10 樓(含)以下有電梯)、住宅大樓(11 樓(含)以上有電梯)、透天厝、套房與其他(即非住宅)。

² 內政部於 2015 年起運用實價登錄資料，以「特徵價格法」編製住宅價格指數，並於 2019 年起改以「類重複交易法」編製房價指數。

表 1 國內主要房價指數之相關資訊

	內政部 住宅價 格指數	臺北市住宅 價格指數	國泰 房價 指數	信義 房價 指數	政大 永慶 房價 指數	政大 永慶 豪宅 指數	清華安富 房價指數	全臺 房價 指數
頻率	季	月	季	季/月	季	月	季/月	季
發布單位	內政部 營建署	臺北市 政府	國泰 建設	信義 房屋	永慶 房屋		安富 財經	瑞普 萊坊
調查資料來源及對象	實價 登錄	實價 登錄	市調 資料	仲介 成交	仲介 成交 資料	實價 登錄	實價 登錄	實價 登錄
	成屋 價格	成屋價格	新推 案	成屋 價格	成屋 價格	成屋 價格	成屋 價格	成屋 價格

資料來源：本文整理

聯合編撰，參考實價登錄資訊並輔以計量經濟模型建構清華安富房價月指數，以提供市場清楚了解房價、交易量變動情形的趨勢。透過不同方法與樣本篩選的編製過程，民眾得以針對各地區、各類型的房價走勢能有更清楚的輪廓。

(二) 實價登錄於房價預測與不動產相關政策分析之應用

除了應用於編製房價指數外，實價登錄資料亦可用於資料分析與政策研擬，在大樣本資料下分析臺灣不動產交易各類變數的統計性質，進而對房價、交易量進行預測與政策分析。

在結合資料科學應用方面，黃士峰、廖育苹(2022)根據不動產實價登錄的成交資料，以廣義加成模型(generalized additive model, GAM)，同時涵蓋各類數值與非數值解釋變數，如：主要用途、建物型態、房屋位置、面積、建物型態等變數，用以預測 2014 到 2019

年間高雄市的不動產成交價格；曹家恩、王健安(2016)利用臺灣不動產實價登錄資料探討住宅、商辦、純土地的需求之價格彈性，同時分析三者之價量關係。

政策分析方面，李明軒、林祖嘉、鄭輝培(2022、2024)運用實價登錄與財政統計年報資料，探討不動產有效稅率對於房地產市場價格與交易量之影響。除此之外，該研究團隊亦運用該資料評估有效稅率與空屋率的關聯。由於實價登錄資料能篩選不同地區房價的分量資料，因此該研究團隊得以評估有效稅率對於不同價格分量的影響。結果顯示：不動產有效稅率對於不同分位之價格與交易量皆有顯著影響，且對於分位越高之市場抑制效果越大。

在應用評估臺灣財富分配方面，考量到不動產價格為多數民眾重要的資產組成，連賢明等(2021)採用財政部 2004 至 2014 年的財產登錄檔，建立涵蓋房屋、土地、股票、存款、債券、金融票券與房屋貸款的個人財富資料。為確保所持有資產貼近現實，該研究進一步使用房地產實價登錄資料來計算房屋與土地價值，並納入個人透過投資公司持有的公司股權(即間接持股)來估算股票價值，以此推估臺灣財富分配的變化。

(三)現行不動產相關統計政策分析的資料面向

在各項不動產統計面向中，除了不動產相關放款統計屬於金融面外，不動產市場的「生產面」與「需求面」等構面資料如交易價格、交易量等亦是分析不動產交易動態的重要資訊。其中當前「交易價格」構面的資料來源主要是以內政部住宅價格指數、民間公布的信義房價指數(中古屋)及國泰房價指數(預售屋)為主，並輔以內政部公布的都市地價指數及主計總處公布的營造工程物價指數捕捉建物與土地的成本面變化；而在「交易量能」構面中，資料的主要取得來源則是內政統計月報的「建物所有權買賣移轉登記」(含第一次登記)與「土地買賣所有權移轉登記」，以掌握不動產市場的交易動態。儘管上述資料已提供不動產之價量資訊，惟在當前實價登錄資料庫已累積大量成

交資料的背景下，運用實價登錄資料庫分析不動產交易動態，相較以現有的房價指數資料分析具備以下優勢：

1. 國泰房價指數及信義全國房價指數運用仲介成交與市調資料，以「特徵價格法」編製並公布每季房價指數。此類房價指數主要聚焦於全國與主要都會區，且更新頻率為每季更新一次，因此無法掌握其他地區的房價波動；實價登錄的資料係根據「不動產成交案件實際資訊申報登錄及預售屋銷售資訊備查辦法」³，固定每旬公布前 30 天登錄的不動產交易紀錄。不只能掌握到鄉鎮市區的行政區資料，也能夠初步掌握前一個月內的部分交易的價量動態⁴。
2. 承上，因實價登錄詳載各類不動產交易資料，透過有效清理並進行資料分群(data clustering)，即可進一步分析特定地區於特定期間之下，不動產交易的總價分量下如何受到政策等相關變數的影響。舉例而言，運用實價登錄資料，能統計不同成屋總價的分量走勢，並運用此類變數更有效評估如高價住宅對於政策變數(如高價住宅信用管制)的影響效果，作為不同分量下房價下行風險(Housing at Risk, HaR)估算的資料來源。
3. 2009 年起，財團法人金融聯合徵信中心(簡稱聯徵中心)定期提供銀行新承做購置住宅貸款、建築貸款相關統計(如地區別貸款金額、貸款筆數、平均成數等)，以掌握不動產貸款情勢之變化。藉由建構實價登錄資料庫，未來結合聯徵中心的信用資料進行交叉驗證，有助精進未來不動產相關的政策分析。

基於當前政府相關單位針對不動產統計與調查項目之內涵或定義不盡相同，須審慎解讀不同單位公布之不動產統計。有鑑於不動產市場攸關總體經濟與金融穩定，加以完善相關統計資料是各項政策措施的重要基礎，本研究嘗試以實價登錄資料為基礎，依照不同需求進行資料篩選，俾提升掌握不動產市場的趨勢與變化。本報告除前言外

³ 本辦法第 10 條規定：「第 10 條規定：經紀業經營仲介業務者，應於簽訂租賃契約書之日起三十日內，填具不動產成交案件實際資訊申報書，向直轄市、縣（市）主管機關或使用電子憑證以網際網路方式申報登錄」

⁴ 實務上，實價登錄因人為等因素導致資料缺漏情形，整體交易之資料仍須 1 年後較穩定。

圖 1 內政部實價登錄資料庫



資料來源：實價登錄系統

，第貳章簡述實價登錄資料庫揭露之資料格式與命名規則並說明資料整理方法；第參章介紹本文有關實價登錄資料庫中成屋、預售屋在不同單位下的資料清理篩選規則；第四章則運用本研究建構之資料庫，針對高價住宅進行簡要統計分析；第伍章則為結論與政策意涵。

貳、內政部實價登錄資料庫簡介

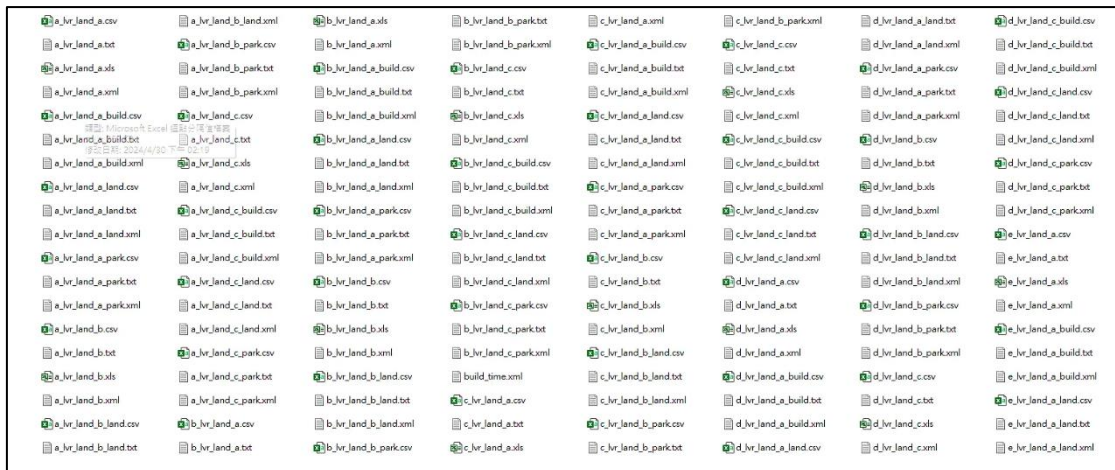
一、實價登錄揭露資料之範圍與檔案格式

不動產實價登錄提供民眾查詢的法源依據，是依據平均地權條例及不動產經紀業管理條例規定⁵，要求不動產買賣或租賃之權利人及義務人(即買賣雙方)於案件申請所有權買賣移轉登記、購買預售屋或租賃時，仲介經紀業於簽訂租賃契約書後 30 日內、代銷經紀業應於案件簽訂契約 30 日內，填具不動產成交案件實際資訊申報書，向直轄市、縣(市)主管機關或使用電子憑證以網際網路方式申報登錄⁶，再

⁵ 土地及成屋的實價登錄法源依據係根據「平均地權條例」第 47 條；租賃與預售屋的實價登錄法源依據則係依據「不動產成交案件實際資訊申報登錄及預售屋銷售資訊備查辦法」第 10 條與第 11 條。

⁶ 詳「不動產經紀業管理條例」第 24-2 條

圖 2 實價登錄網頁下載之檔案內容截圖



資料來源：實價登錄系統

經由地方政府有關單位確認申報資料之真實性與合理性後，整理分類其內容並定期揭露於內政部實價登錄的供應系統(圖 1)。為提升政府資料取用與分析上的便利性，該系統亦提供不同的原始檔檔案格式，如 CSV、TXT、XLS、XML 檔案，方便使用者在不同的程式環境下進行資料分析。圖 2 為實價登錄系統可供下載之檔案截圖。在上述資料格式中，因 CSV 檔對於 XML 檔是較簡單的檔案格式，同時較 TXT 檔更適合儲存與傳送大量結構化資料，且儲存容量低於 XLS 檔。因此，為使本研究開發之程式能在日後被更廣泛地應用並持續維護，同時保有易於操作的特性，以下即針對 CSV 格式實價登錄揭露資料之命名規則進行補充說明。

二、實價登錄揭露資料之檔案命名原則

在檔案命名規則方面，現行實價登錄系統於每月 1 日、11 日、21 日發布最新資料(發布目前 45 天內完成登錄並符合揭露原則者)，並且於每年之 1 月、4 月、7 月、10 月將月資料合併為季資料，同時依照不同的縣市與不動產成交類別進行以下分類，說明如下：

(一)縣市別命名規則

實價登錄揭露之交易資料，係依照縣市別方式彙整於 CSV 檔內，並使用英文字母 a-z 為檔案字首命名該 CSV 檔，命名對照表詳表

表 2 各縣市不動產交易檔案名稱之字首與縣市別對照表

檔案名稱 字首	縣市別	檔案名稱 字首	縣市別
a	臺北市	m	南投縣
b	臺中市	n	彰化縣
c	基隆市	o	新竹市
d	臺南市	p	雲林縣
e	高雄市	q	嘉義縣
f	新北市	t	屏東縣
g	宜蘭縣	u	花蓮縣
h	桃園市	v	臺東縣
i	嘉義市	w	金門縣
j	新竹縣	x	澎湖縣
k	苗栗縣	z	連江縣

資料來源：實價登錄系統、本文整理

表 3 實價登錄不同案件與交易類別命名規則

CSV 資料最末命名格式	內容描述
_a.csv	成屋交易資料(合計)
_a_land.csv	成屋交易之土地案件
_a_build.csv	成屋交易之建物案件
_a_park.csv	成屋交易之停車場案件
_b.csv	預售屋交易資料(合計)
_b_land.csv	預售屋交易之土地案件
_b_park.csv	預售屋交易之停車場案件
_c.csv	租賃交易資料(合計)
_c_land.csv	租賃交易之土地案件
_c_build.csv	租賃交易之建物案件
_c_park.csv	租賃交易之停車場案件

資料來源：實價登錄系統、本文整理

2。舉例來說，若自該系統下載的 CSV 檔字首為 a，則該 CSV 檔即為臺北市內的相關交易；若 CSV 檔案字首為 f，則為新北市成交案件；若 CSV 檔案字首為 c，則為基隆市成交案件，以此類推。

(二)檔案的命名規則

除了根據縣市別進行分類外，實價登錄系統揭露之檔案亦針對成屋買賣、預售屋買賣與租賃案件，以字元_a、_b、_c 於檔案末端命名進行區分，並依交易標的進一步區分為「合計」、「建築物」、「土地」、「停車位」之交易內容，詳見表 3。除此之外，由於表 3 之成屋交易資料(_a.csv)、預售屋交易資料(_b.csv)、租賃交易資料(_c.csv)之合計資料已涵蓋土地、建物與停車場案件，本文後續將以「合計」資料整理內政部定期公布之成屋交易、預售屋交易與租賃案件，作為日後不動產、預售屋與租賃交易之額外資料來源。

(三)實價登錄系統揭露之欄位資料內容

表 4 彙整了成屋、預售屋與租賃案件在實價登錄下記載的欄位資訊，除了交易標的之地址，交易日期以外，亦針對物件的內容進行詳細的紀錄。儘管上述案件的紀錄欄位多數相同，惟因成屋、預售屋與租賃案件交易本質的差異，因此實價登錄系統揭露資料亦進行部分調整(表 4)。舉例而言，有別於租賃資訊，實價登錄針對成屋與預售屋交易進一步詳細載明交易年月日、交易筆棟數、土地移轉總面積、建物之移轉總面積、移轉層次及總價元等資訊；除此之外，自 2020 年政府推動「健全房地產市場方案」實施後，2021 年 7 月後的預售屋交易必須於成交後 30 天內即時申報實際交易資訊，同時提供建案名稱與棟及號；2023 年 7 月起更增加解約情形以供查詢。

表 4 實價登錄不同交易類別之欄位差異

共同欄位內容	不同欄位內容		
	成屋	預售屋	租賃
交易標的	交易年月日		租賃年月日
鄉鎮市區	交易筆棟數		租賃筆棟數
土地位置建物門牌	土地移轉總面積 平方公尺		土地面積 平方公尺
都市土地使用分區	建物移轉總面積 平方公尺		建物總面積 平方公尺
非都市土地使用分區	移轉層次		租賃層次
非都市土地使用編定	總價(元)		租屋總額(元)
總樓層數	車位移轉總面積 平方公尺		車位面積 平方公尺
建物型態	車位總價元		車位總額元
建物主要用途	移轉編號	解約情形	租賃期間
主要建材	主建物面積	棟及號	出租型態
建築完成年月	附屬建物面積	建案名稱	附屬設備
建物現況格局 (廳、房、衛、隔間)	陽台面積		有無管理員
交易標的	電梯		有無電梯
有無管理組織			有無附傢俱
單價元/平方公尺			租賃住宅服務
車位類別			
編號			
備註			

資料來源：實價登錄系統、作者自行整理

參、研究架構：資料庫篩選原則與資料清理流程

本節將進一步詳細說明本文資料的清理流程，運用程式語言 Python 提供之相關套件針對實價登錄系統公布之檔案進行資料清理，剔除交易日期內含亂碼，資料不齊之問題資料。

一、實價登錄資料下載與清理工作流程

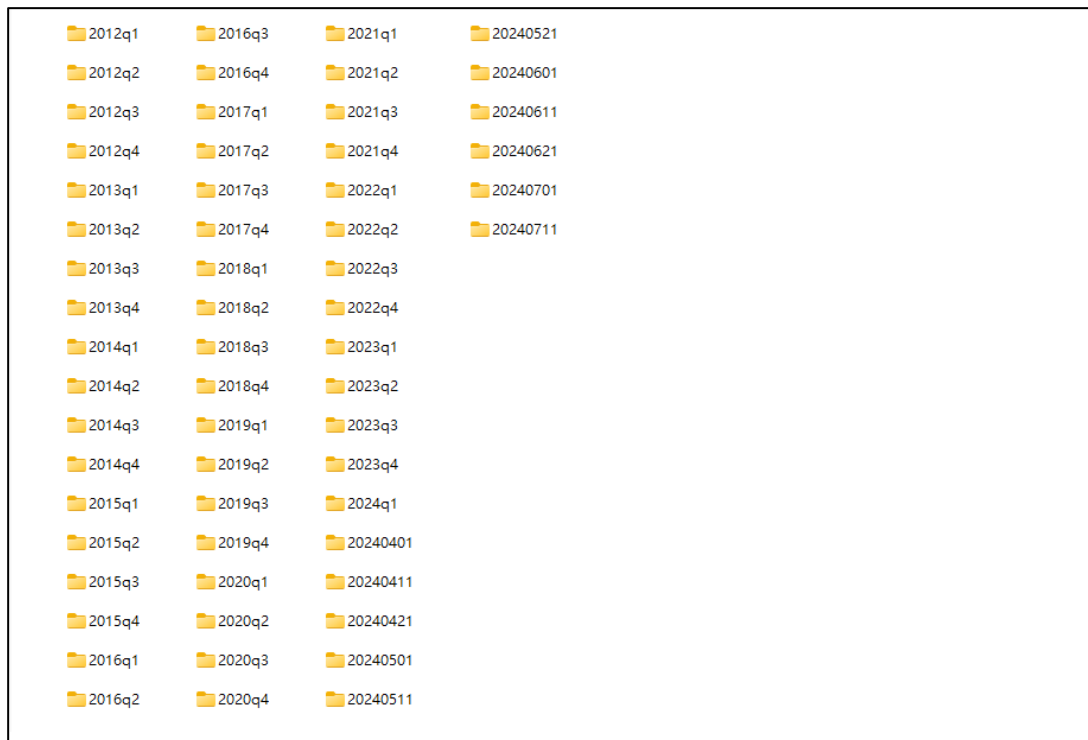
由於實價登錄系統在目前的尚未提供應用程式開發介面 (Application Programming Interface, API) 串接取值之服務⁷，因此本文無法直接運用 R、Python 或 Matlab 等程式逕行串接實價登錄的後端資料庫進行資料分析。換言之，目前該系統僅提供靜態開放資料供用戶下載。針對該系統提供之「非本期下載資料」，資料儲存的命名規則與整理步驟如下：

1. 若非本期資料為季資料，則該資料經下載解壓縮後，將該檔案儲存於“yyyyqn”命名之資料夾，其中 yyyy 為西元年，n 為該季之季度。舉例而言，若下載之資料為 2020 年第一季，則該期資料的下載資料將儲存於“2020q1”之資料夾中，以方便未來資料維護管理。
2. 若非本期資料為近期資料，則實價登錄系統將於每旬之 1 日、11 日、21 日公布其上個月以來之交易案件。針對當期之非季度資料，其儲存資料夾之命名規則為“yyyymmdd”，其中 yyyy 為西元年，mm 與 dd 分別為月份及日期。圖 3 為每筆從公布日期下載後儲存的資料夾命名規則說明，資料夾內含各地區之相關資料檔，檔案內容如圖 2。
3. 最後，經由下載各期資料並運用上揭資料夾的歸類原則，本研究進一步依據表 3 之交易類別命名規則，運用 Python 程式轉換為 pandas 套件之資料格式，並排除「交易年月日」欄位日期格式錯誤之交易資料，作為後續不動產成屋、預售屋與租賃交易之資料分析基礎。

由上述檔案分類與資料整理原則，表 5 初步彙整移除交易日期異常之資料清理結果。其中無法辨識交易日期的成屋資料共 4,709 筆，占整個資料期間成屋總樣本數的 0.11%；預售屋資料共有 463,822 筆，其中無法辨識交易日期的資料共有 65 筆。藉由上述的初步篩選，本研究組建之資料庫未來可根據不同的條件、針對不同交易類別進行篩

⁷ 詳政府資料互動專區之問答說明。

圖 3 實價登錄檔案下載歸類截圖



資料來源：作者自行整理

表 5 初步資料清理結果(刪除異常交易日期)

	成屋	預售屋	租賃
原始資料樣本數	4,202,920	463,822	463,989
剔除樣本數 (交易日期異常)	4,709	65	0
剩餘樣本數	4,198,211	463,757	493,989
剔除比例(%)	0.11%	0.01%	0%

資料來源：作者自行整理

備註：實價登錄資料公布起訖為 2012q3 至 2024/08/01。

選，除了可作為關注房市動態的額外資料來源外，亦能透過此資料進行反事實推論與情境分析，作為後續政策研擬的研究基礎。

二、本研究與其他單位資料篩選比對

從表 1 之彙整可知，目前實價登錄資料已成為不少政府或私人機構依據其需求編製房價指數的資料來源。為了驗證本研究資料篩選結果之妥適性並與政府機關或私人機構之公布資料一致，以下將分別與臺北市政府、內政部針對不同交易案件(成屋、預售屋)之公布結果進行比對(有關本研究與其他單位資料篩選比對表，詳見附錄)。

在進行資料比對前，實價登錄資料庫資料因以下因素，可能導致本研究篩選之結果無法完全與其他機構公布結果產生一致，其原因說明如下：

1. 資料登錄時效性較所有權移轉登記資料落後：如前所述，因地方政府須確認實價登錄資料正確性與合理性，以致內政部每期揭露之近期各月(季)交易資訊⁸，需經過一段期間更新後才會較完整，因此各單位公布之不動產交易等統計結果，可能因實價登錄資料公布期間的落差產生統計結果的差異。此外，有鑑於最新揭露資料可能涵蓋過去數月交易資料，為提升資料的完整性，須隨時回溯修正更新資料。
2. 處理人工申報輸入問題(如文字、數據寫法不一致、相關欄位資料缺失等)的差異：由於實價登錄資料欄位眾多(含交易標的、地址、土地移轉總面積、土地使用分區、交易日期、主要用途、建物型態、建材、車位總價等，詳表 4)。因該資料係由權利人或義務人申報登錄，常有人工申報輸入問題(如文字、數據寫法不一致、相關欄位資料缺失等)，故各界運用內政部實價登錄資料庫均需進行資料清理。惟因現行並無一致性資料清理標準，致不同機構或學者可能因清理程序的差異影響後續的統計結果。

儘管存在上述的誤差，文獻上針對不動產市場關注的資料多從大樣本與整體角度進行分析，由於部分些微的數據差異並不會改變整體不動產的趨勢，故本研究之篩選資料仍有助掌握當前不動產市場的相

⁸ 實價登錄系統揭露日前 45 天交易資料

關交易動態。

(一) 與臺北市政府地政雲實價登錄交易統計量統計之比較

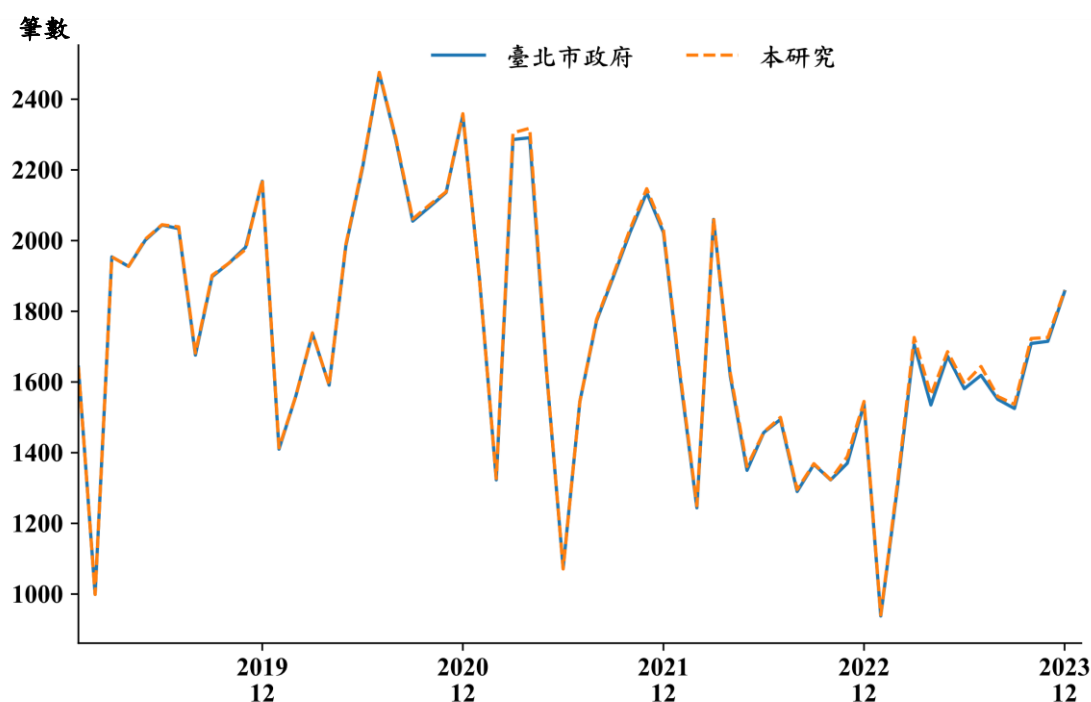
自 2017 年 4 月 20 日以來，「臺北地政雲」為臺北市政府的提供民眾查詢臺北市不動產買賣資訊的主要平台。該系統內的「大數據進階統計」功能將北市的買賣案件區分為「全部」、「房地」及「房地＋車位」三大項，並針對上述類別進行交易量及均價進行簡單統計分析。為了驗證本研究之資料庫與臺北地政雲提供的資料一致，以下針對台北市「房地＋車位」為例，進行以下檢定：

1. 於「臺北市地政雲」的實價登錄系統中，下載 2019 年 1 月至 2023 年 12 月之臺北市「房地＋車位」交易，並選擇其主要用途為「住宅」與「非住宅」交易。
2. 運用相同條件於本研究建構之資料庫進行篩選，並以月為單位計算各月份之交易量。篩選條件如下：
 - (1) 在「交易類別」欄位中篩選滿足「房地」之樣本，排除「土地」及「建物」之不動產交易，並將交易地區聚焦於「臺北市」。
 - (2) 運用以下公式計算每坪單價，並移除異常值(每坪單價小於或等於 0)：

$$\text{每坪單價} = \frac{\text{不動產總價元} - \text{車位總價元}}{\text{建物移轉面積}(m^2) - \text{車位移轉面積}(m^2)} \times 3.305$$

- (3) 篩選期間之起點與終點分別為 2019 年 1 月 1 日至 2023 年 12 月 31 日，共篩選出 104,469 筆資料，篩選程序詳表 6。

圖 4 資料篩選比較：臺北地政雲 vs. 本研究之交易量統計



資料來源：作者自行整理

表 6 本研究參考臺北地政雲資料清理流程表

清理項目		篩選條件	保留筆數	清除數	剔除率
原始樣本數：					
1.交易地區篩選「臺北市」			123,435		
2.交易年月日介於 2019.01.12 至 2023.12.31					
減：					
排除資料	(1)交易標的	非屬房地		18,959	15.36%
異常值	(2)總價元	缺失值、 總價等於 0		7	0.006%
			104,469		15.366%

資料來源：作者自行整理

表 7 臺北地政雲與本研究各月絕對值誤差統計

單位：筆數

誤差 平均	誤差 中間值	誤差 眾數	誤差 標準差	誤差 最小值	誤差 最大值
6.85	5	6	6.70	0	28

資料來源：作者自行整理

圖 4 描繪臺北地政雲以及運用本研究資料庫建構篩選之臺北市交易量走勢，可以發現：本文篩選的結果與臺北地政雲的篩選結果幾乎吻合。表 7 進一步以統計方式呈現兩者的誤差結果，可以發現兩筆資料的每月平均誤差僅 6.85 筆，占該段期間臺北地政雲樣本的平均誤差 0.39%。兩筆資料每月誤差筆數的最大值為 28 筆，落在 2023 年 4 月的交易期間；中位數則僅 5 筆。換言之，本資料庫的建構能夠確切掌握臺北市的不動產交易資料，藉由設定妥適的篩選條件即能產生與臺北地政雲幾乎一致的數據結果。

(二) 與內政部不動產資訊平台統計之比較

除了臺北市地政雲提供交易資料外，內政部「不動產資訊平台」內的「住宅統計」亦會定期提供各縣市每季的住宅成交量。然而，如果仔細比較臺北市地政雲的交易結果，可以發現兩者之間的數據呈現顯著的差異。以 2020 年 Q2 為例，臺北市地政雲公布之交易筆數為 5,792 筆⁹，然而內政部公布之「臺北市住宅買賣移轉筆數」僅 4,012 筆，兩者資料差距接近 1,800 筆。因此，若研究者不能清楚理解兩者資料之篩選差異，逕自以「房價指數」與「住宅成交量」等指標分析臺北市不動產價量走勢，將容易產生統計結果上的錯誤判讀，可能影響實證分析的穩健性。

為了與不動產資訊平台的交易結果進行比較，本節運用以下篩選條件針對實價登錄資料庫進行篩選並進行以下處理。首先根據不動產資訊平台「住宅買賣移轉筆數」的說明，資料來源與篩選方法如下：

⁹ 2020 年 4、5、6 月臺北市地政雲公布之交易資料分別為 1591、1989、2212 筆。

1. 資料來源為內政部地政司提供。
2. 資料之項目定義為辦理移轉登記，登記原因為買賣，主要用途登記為住家、住商、住工、國民住宅以及農舍等。

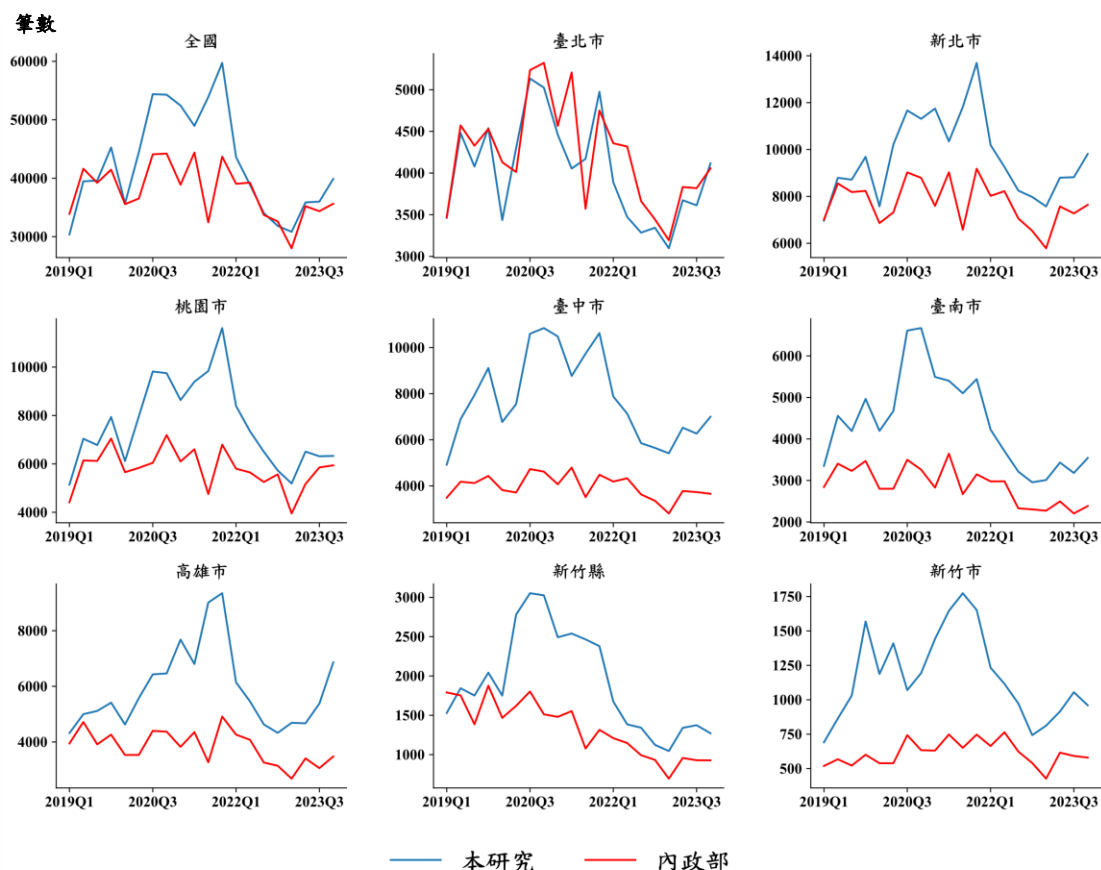
本研究建構之實價登錄資料之篩選程序如下：

1. 交易期間之起點與終點分別選為 2019 年 1 月 1 日至 2023 年 12 月 31 日，原始資料共 1,829,531 筆資料。
2. 在「交易類別」欄位中篩選「房地」之不動產交易，共清除 504,065 筆。
3. 在主要用途欄位中，篩選「住家」、「住商」、「住工」、「國民住宅」以及「農舍」之交易樣本，剔除 282,753 筆資料。
4. 運用先前每坪單價之公式，刪除每坪單價及總價等於 0 的異常值資料，共移除 1,430 筆資料，篩選程序詳表 8。

表 8 本研究參考不動產資訊平台資料清理流程表

清理項目		篩選條件	保留筆數	清除數	剔除率
原始樣本數： 交易年月日介於 2019.01.12 至 2023.12.31			1,829,531		
減：					
排除資料	(1)交易標的	非屬房地		504,065	27.55%
	(2)主要用途	非屬以下： 住家、住商 住工、農舍 國民住宅		282,753	15.45%
異常值	(3)總價元	缺失值、 等於 0		29	0.08%
	(4)單價元/ 平方公尺	等於 0		1,401	
			1,041,283		43.08%

圖 5 資料篩選比較：不動產資訊平台 vs. 本研究之交易量統計



資料來源：不動產資訊平台，作者自行整理

表 9 本研究實價登錄與不動產資訊平台住宅交易量資料相關係數

全國	臺北市	新北市	桃園市	臺中市	臺南市	高雄市	新竹市	新竹縣
0.7	0.79	0.59	0.60	0.72	0.75	0.36	0.64	0.53

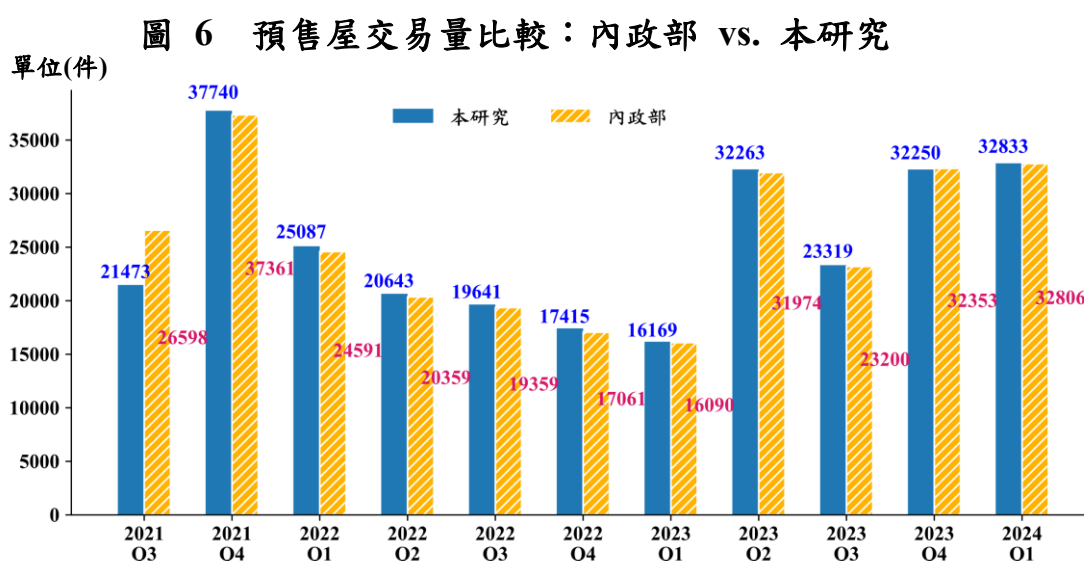
資料來源：不動產資訊平台，作者自行整理

圖 5 與表 9 彙整了不動產資訊平台與本研究實價登錄篩選的結果與兩者資料的相關係數。可以發現，儘管本研究盡可能按照不動產資訊平台之標準篩選，惟受資料來源不同的影響，除了臺北市之交易量較為吻合外，其餘地區的交易量件數普遍明顯高於不動產資訊平台公布之交易件數，顯示不動產交易在不同的資料來源、交易內容與篩選條件下，將產生截然不同的統計結果，連帶影響後續相關的實證分析。

(三)預售屋交易：與內政部統計之比較

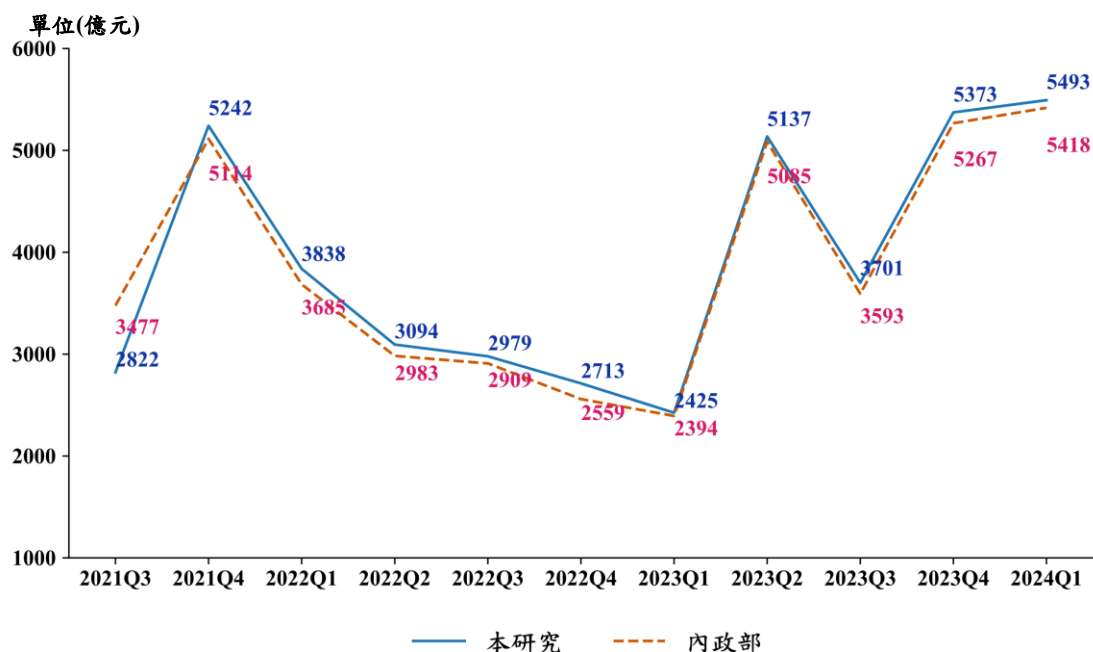
預售屋係指交易日期早於完工日期的房屋。在預售屋制度下，開發商得透過較低門檻條件吸引民眾購買；消費者則能透過較低之頭期款與工程款減輕購屋負擔，因此預售屋交易在臺灣的不動產市場亦佔有相當程度的比例，同時能反映整體不動產市場的前景與資金熱絡的程度。

2020 年初隨著主要經濟體採行擴張性貨幣政策因應疫情，加以美中貿易戰以來臺商資金回流，國內房市交易熱絡，預售屋建案熱銷情況屢見不鮮，房市價量非理性的上揚使預售屋易成為炒作的投機商品。為促使預售屋不動產交易資訊更加透明並減少投機炒作，立法院於 2020 年 12 月通過「平均地權條例」、「地政士法」及「不動產經紀業管理條例」等三大法修正草案，即實價登錄 2.0 制度，並自 2021 年 7 月 1 日施行，要求銷售預售屋賣方(預售屋買賣契約出賣人，如不動產開發業或自然人、委託經紀業代銷成交者或受託之經紀業者)為申報義務人，應在與買受人簽定預售屋買賣契約之日起 30 日內，完成成交資訊申報登錄；2023 年 7 月進一步要針對預售屋針對解約情形進行揭露，若預售屋買賣契約若有解約情形，應於解約預售屋買賣契約起 30 日內辦理登錄申報。



資料來源：實價登錄系統，作者自行整理

圖 7 預售屋銷售總額比較：內政部 vs. 本研究



資料來源：實價登錄系統，作者自行整理

表 10 本研究實價登錄與內政部交易量與銷售總額比較

項目 資料來源	平均 交易量(件數)	平均 總銷金額(億元)
本研究	25,348	3893
內政部	25,614	3862
平均每月誤差	265	31
誤差百分比	1.05%	0.80%

資料來源：內政部，作者自行整理

圖 6 與圖 7 分別比較 2021 年 Q3 至 2024 年 Q1 期間交易量與銷售總額之結果，兩者資料之平均件數、銷售總額與相對誤差值整理於表 10。由於內政部並未說明其篩選條件，因此本節直接呈現原始資料的統計。由上述圖表可知：除了 2021 年 Q3 的資料誤差較大以外，其餘各期間的誤差均相當小¹⁰，平均誤差約在 1% 左右。藉由與內政部公布資料比對確認資料的一致性，未來可運用本研究資料篩選各地區(如鄉鎮市區)於預售屋的熱絡程度，並運用此資料庫的提供之資訊作為篩選的條件線索，檢測該地區的不動產的放款狀態，針對該地區的

¹⁰ 針對 2023 年 Q3 誤差較大之問題，未來將詢問內政部相關單位的資料來源，以提升資料的可信度與完整性。

表 11 2023 年、2024 年上半年預售屋交易量變化

單位：筆數

地區	2023 上半年 (1)	2024 上半年 (2)	相差 (2)-(1)
臺北市	1,615	2,988	1,373(85.0%)
新北市	9,777	11,015	1,238(12.8%)
桃園市	9,215	13,504	4,289(46.5%)
臺中市	8,653	14,740	6,087(70.3%)
臺南市	5,222	5,857	635(12.2%)
高雄市	4,392	9,322	4,930(112.2%)
新竹市	807	1,184	377(46.7%)
新竹縣	2,019	3,198	1,179(58.4%)
其他	6,732	10,454	3,722(55.3%)

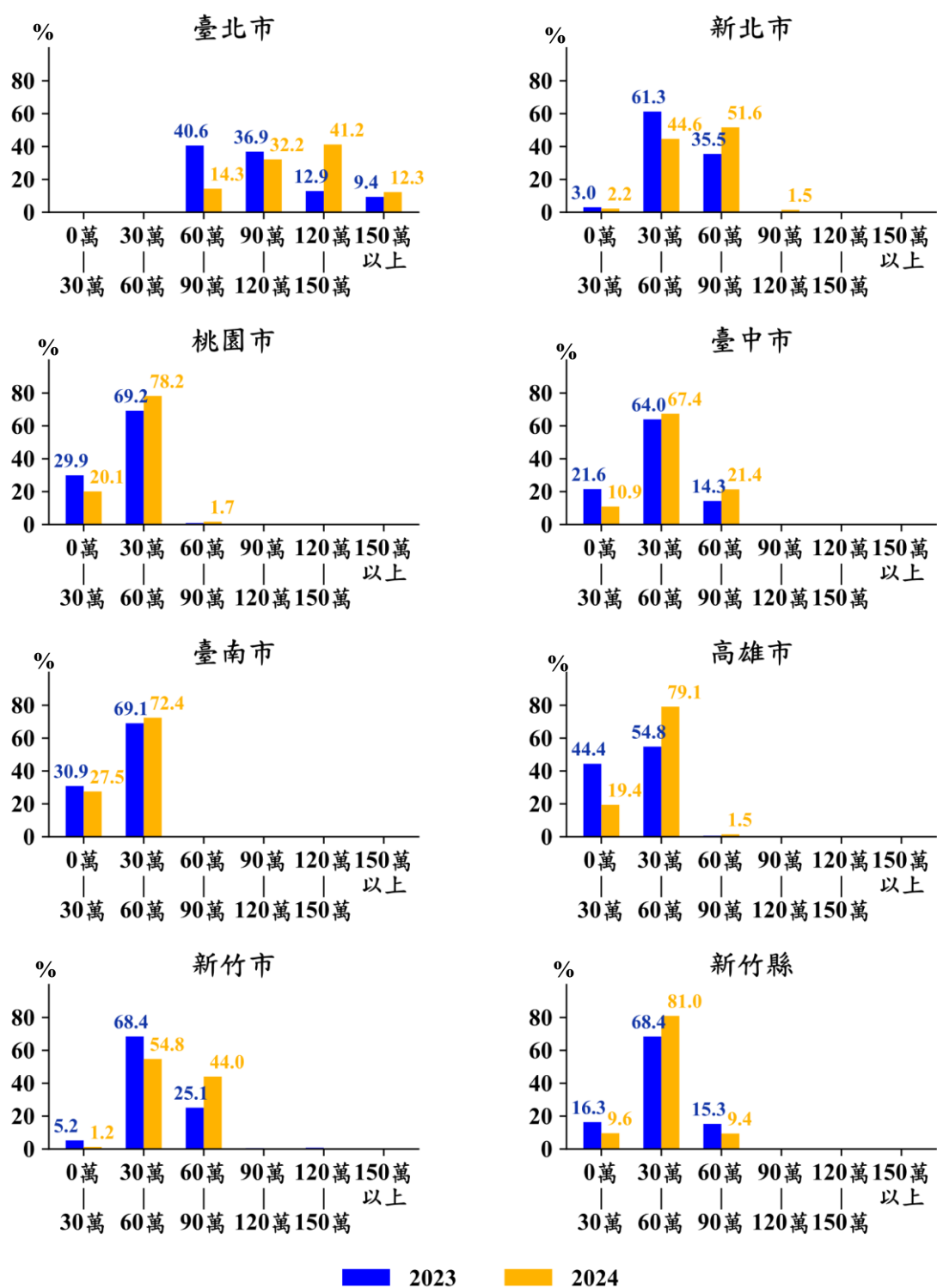
註：括弧內為 2024 年上半年交易量年增率

資料來源：作者自行整理

授信風險進行壓力測試。舉例而言，2023 年在新青安措施的推動下，成屋市場交易熱絡，亦帶動了預售屋市場的行情。表 11 彙整 2023 年與 2024 年上半年主要都會區預售屋的交易量變化。從該表可知，2024 年上半年高雄市、臺北市及臺中市預售屋交易分別較去年同期增加 112.2%、85%、70.3% 較多，而其他縣市均較去年同期增加。

除了能分析各縣市的預售屋交易量的變化以外，由於預售屋的價格能反映當前購屋者對於未來房價的預估走勢以及當前營建成本的變化，因此運用實價資料能彙整主要都會區的預售屋成交單價的區間分布。圖 8 分別整理主要都會區 2023 年上半年與 2024 年上半年預售屋每坪單價成交區間變化。可以發現各都會區的成交單價區間分布明顯右移：如臺北市的預售屋中每坪價格落於 120-150 萬的中古屋交易占比由 12.9% 大幅上升至 41.2%；60-90 萬的預售屋交易占比則 40.6% 降為 14.3%，顯示當前預售屋在近期需求推升以及營建成本居高下，建商逐漸反映在預售屋交易市場中。

圖 8 主要都會區預售屋交易單價區間占比：
2023H1 vs. 2024H1



資料來源：實價登錄，作者自行整理

**表 12 預售屋成交總價平均數與中位數：
2023Q2 vs. 2024Q2**

單位：萬元

期間	2023Q2		2024Q2		年增率(%)	
成交總價	平均數	中位數	平均數	中位數	平均數	中位數
全國	1592	1347	1844	1530	15.80	13.59
臺北市	3709	3037	4397	3440	18.55	13.29
新北市	1824	1616	2246	2032	23.14	25.74
桃園市	1376	1249	1589	1462	15.53	17.01
臺中市	1730	1528	2068	1845	19.52	20.75
臺南市	1273	1168	1316	1206	3.38	3.25
高雄市	1202	1055	1443	1285	19.99	21.80
新竹市	2422	1963	2579	2605	6.50	32.74
新竹縣	1935	1660	1953	1728	0.94	4.10
其他	1059	960	1182	1100	11.66	14.58

資料來源：實價登錄，作者自行整理

預售屋每坪單價的顯著上揚亦反映在成交總價上。表 12 彙整 2023Q2 與 2024Q2 全國與主要都會區的預售屋成交總價的平均數與中位數，並計算兩種價格的年增率。從表 12 可知，近期除了新竹縣、臺南市以外，其餘地區的成交總價在平均數或是在中位數的年增率均逾 10%。由於預售屋的頭期款與工程款通常僅分別占成交總價的 10% 至 15%¹¹，因此儘管成交總價高，在取得成本較低與對未來房價上漲預期心理驅動下，導致近期預售屋的交易價量齊揚。

另一方面，預售屋的建造應符合建築法 54 條規定：起造人自領得建造執照或雜項執照之日起，應於六個月內開工，並得申請展期一次，期限為三個月。由於近年主要都會區新建案展延開工宅數屢創新高，顯示在近年營建工程成本上漲的背景下，建商的資金調度漸趨謹慎。透過本研究的初步篩選，未來可結合聯徵資料分析特定地區(如臺中市、桃園市)建商針對特定地區預售屋的建築貸款、土地融資等資金運用狀況進行分析，以保障購屋者的權益。

¹¹ 以 2024H1 臺北市預售屋中位數為例，若房屋總價為 3440 萬元，則預售屋頭期款僅需 344 萬至 516 萬左右。

伍、結論與政策意涵

2012 年以前，國內不動產相關統計長期存在制度面問題，國內不動產相關統計長期因資料涵蓋範圍、統計頻率與對象定義的不同，致使出現數據分歧的現象；2012 年後，政府在健全房市措施的推動下促成實價登錄資料庫的建立，除了提供民間購屋決策外，亦提供政府與有關單位相關政策的研擬。

實價登錄資料自成立至今已累計上百萬筆交易資料，再加上制定公共政策、檢視施政效果的過程中需仰賴大量資料進行佐證，本文參酌過去不動產相關統計經驗與實務，詳細說明實價登錄資料庫的命名規則、欄位資訊等相關檔案格式，並運用程式針對各期公布資料進行儲存與篩選。藉由與政府與民間單位(臺北市、內政部、信義房屋)的篩選標準進行比較後，本研究之實價登錄資料庫能初步得與上述單位相近的結果，有助於即時掌握不動產市場中的住宅交易、預售屋交易與高價住宅交易的價量情勢。以近期預售屋市場為例：2024 年上半年以來，全台主要都會區之預售屋交易相較 2023 年上半年呈現價量齊揚的走勢，不但預售屋交易量大幅成長，每坪單價分布亦顯著右移，顯示當前預售屋市場在近期出現顯著的結構轉變。

本文具有下列意涵與未來發展方向，分述如下：

- (一) 藉由建構之成屋與預售屋之價量資料，可做為觀測不動產統計資料的額外來源，以更精確且即時掌握不動產的交易動態。
- (二) 朱芳妮、楊茜文等(2020)發現：新聞媒體對於不動產市場的報會顯著影響下一期的市場預期與情緒，進而影響不動產市場的價量走勢。基於媒體係市場傳遞訊息的重要管道，因此針對媒體的報導內容，可由本研究建立之實價登錄資料庫針進行查核驗證，客觀檢視事件訊息之正確性與合理性。
- (三) 本研究清理與彙整之實價登錄價量資料，未來可結合相關不動產放款統計及聯徵中心提供之相關數據，進一步延伸蔡釗旻

(2022)¹²、鄭漢亮(2023)¹³房市下行風險之研究，針對不動產成屋與預售屋的金融脆弱性(或風險)進行評估；同時亦以本研究之資料來源進一步延伸俞欣榮(2020)衡量總體審慎政策工具的研究成果，透過更完善的資料檢視各項措施的施政成效。

¹² 蔡釗旻(2022)針對房市風險之研究係以國泰房價指數為被解釋變數。

¹³ 鄭漢亮(2023)運用信義房價指數與分量隨機森林機器學習篩選有用的特徵變數，預測臺灣房價未來的下行風險。

參考文獻

- 朱芳妮、楊茜文、黃御維、陳明吉(2020)，「媒體傳播效應與房市變化關聯性之驗證」，管理學報，第 37 卷 3 期，頁 225—257。
- 連賢明、曾中信、楊子霆、韓幸紋、羅光達(2021)，「臺灣財富分配 2004-2014：以個人財產登錄資料推估」，第 49 卷 1 期，頁 77—130。
- 黃士峰、廖育莘(2022)，「基於廣義加成模型之房屋價格預測方法」，中國統計學報，第 60 卷 2 期，頁 95—124。
- 李明軒、林祖嘉、鄭輝培(2022)，「不動產有效稅率對於房地產市場價量影響之探討：臺灣縣市級資料之應用」，第 31 卷第 2 期，頁 1—27。
- 鄭輝培、林祖嘉、李明軒(2024)，「不動產有效稅率對不動產市場空屋率之影響：臺灣鄉鎮級資料之應用」，都市與計劃，第 51 卷 1 期，頁 29—51。
- 許怡君、黃信閔(2023)，「不動產貸款統計資料揭露之探討(業務改進)」，中央銀行內部報告。
- 鄭漢亮(2023)，「房價風險與危機預警：隨機森林之應用」，央行內部研究
- 蔡釗旻(2022)，「評估臺灣房價風險值 (HaR)」，央行內部報告。
- 曹家恩、王健安(2016)，「不動產需求的價格彈性分析：來自實價登錄大數據的新證據」，暨南大學財務金融學系碩士論文。
- 蔡曜如(2023)，「國際精進不動產統計經驗對我國之啟示」，中央銀行內部報告。
- 俞欣榮(2020)，「租稅措施與房貸成數管制對房市效果評估：台灣的個案研究」，央行內部研究。

附錄、本研究與其他單位資料篩選比對表

	臺北市府 地政雲	內政部 不動產資訊平台	信義房屋 高價住宅
篩選期間	2019/1/1~2023/12/31	2019/1/1~2023/12/31	2014/7/1~2023/12/31
交易標的	「房地+車位」、 「房地」	「房地」	「房地+車位」、 「房地」
主要用途		「住家」、「住商」、 「住工」、「農舍」、 「國民住宅」	「住家」、「住商」、 「住工」、「農舍」、 等與「住」字相關用途
建物型態			「住宅大樓」、「華廈」、 「公寓」、「套房」、 「透天厝」、「農舍」
屋齡			大於 1 年
排除項目			
總價元	非缺失值、總價大於 0	非缺失值、總價大於 0	非缺失值、總價大於 0
單價元 每平方公尺		大於 0	大於 0
			非正常交易(親友、特殊 相關交易)與地上權

資料來源：本文整理