

計畫編號：109 保發-11.1-保-01-06-001(15)

**水土保持工程巨量資料探勘與計量實證分
析研究**

**Study of Soil and Water Conservation
Engineering Using Big Data Mining and
Econometric Empirical Analysis
成果報告書**

執行單位：國立臺灣大學

執行期間：109 年 2 月 6 日至 109 年 12 月 31 日

計畫主持人：荷世平 教授

協同主持人：蔡宛珊 教授

行政院農業委員會水土保持局 編印

中華民國 109 年 12 月 31 日

(本報告書內容及建議純屬執行單位意見，僅供本局施政參考)

水土保持工程巨量資料探勘與計量實證分析研究

摘要

近年巨量資料之應用研究成為趨勢，數據管理已成為政府單位關注的議題，如何將過去累積資料轉換為有效益之資訊十分重要。水土保持局歷年執行相關集水區整體治理及治山防災等各計畫，並建立水土保持管考系統，其內部資料量體繁多，應具大數據分析之可行性，若加以萃取、分析，有助於提供有效及快速決策，以提昇反應速度及管理績效。本計劃參照 104 年委辦計畫中水保局相關單位專訪之訪談結果，以「工程管考系統」資料庫為基礎，針對工程專案發包之流標情況及工程逾期狀況進行探索性分析。經過專家判斷從資料庫欄位中選擇出具有意義的自變數及控制變數後，建立大數據計量分析模型之架構，進行變數間顯著性之實證，以發掘變數之間的關聯性，進而深入討論執行機關對於流標及工程逾期可能原因及其改進策略意涵。我們亦透過大數據計量分析之結果，建立巨量資料處理流程及大數據計量模型分析範本，作為未來大數據資料處理基礎，提供未來工作成果評估分析及未來政策研擬等相關參考依據。

關鍵詞：數據管理、管考系統、大數據分析、決策、實證分析

Study of Soil and Water Conservation Engineering Using Big Data Mining and Econometric Empirical Analysis

Abstract

Research on big data applications has become a trend in recent years, and how to convert big data into useful information has become the topic issue for government agencies. The Bureau of Soil and Water Conservation has implemented various projects related to the overall management of watersheds and disaster prevention and established a solid water and conservation project management system. The data collected by the system are becoming large in quantity and are feasible for big data analysis. If properly extracted and analyzed, the data will help provide effective and rapid decision-making and improve response and management performance. Based on the interview results of research commissioned by the Bureau in 2014, we use the Engineering Management Examination System's database and conduct exploratory analyses on the issues of procurement bidding failure and project schedule delay. By carefully selecting model variables, we build the big data econometric model and perform the significance tests between the variables. Then, we investigate the possible reason for the procurement failure and schedule delay by interviewing the managers in the Bureau. We also establish a big data analysis framework model template, which will serve as the basis for the Bureau's big data analysis in the future. We hope that this research can provide the relevant reference used for performance analysis and derive strategic implications for improving future performance. Also, this result can improve the execution strategies and extend to policy research in the future.

Keywords: data management, solid water and conservation project management system, big data analysis, empirical study, project performance

目次

摘要.....	I
Abstract.....	II
目次.....	III
表次.....	V
圖次.....	VII
第一章 緒論.....	1-1
1.1 計畫緣起與背景	1-1
1.2 計畫目的.....	1-1
1.3 計畫範圍.....	1-2
1.4 計畫項目與內容	1-4
1.5 研究架構.....	1-5
1.6 計畫執行進度	1-6
1.7 預期成果.....	1-6
第二章 大數據探索性實證分析	2-1
2.1 研究執行構想	2-1
2.2 研究方法.....	2-2
2.3 研究流程.....	2-9
第三章 文獻回顧	3-1
3.1 過往計畫探討	3-1
3.2 專家訪談.....	3-6
3.3 相關研究討論	3-8
第四章 探索性實證模型假說	4-1
4.1 假說推導.....	4-1
4.2 影響發包遲延之假說	4-1

4.3 影響逾期情況之假說	4-3
4.4 小結.....	4-5
第五章 探索性實證模型研究設計	5-1
5.1 實證研究概念	5-1
5.2 實證模型相關檢定	5-2
5.3 變數之選定.....	5-3
5.4 模型介紹.....	5-10
5.5 資料蒐集.....	5-13
第六章 大數據分析結果與討論	6- 1
6.1 敘述性統計.....	6-1
6.2 相關係數矩陣	6-7
6.3 回歸模型檢定	6-8
6.4 回歸分析結果與討論	6-9
6.5 策略擬定.....	6-32
第七章 結論與建議	7-1
7.1 結論.....	7-1
7.2 建議.....	7-3
參考文獻.....	參-1
附錄.....	附錄-1
附錄一、期初審查會議紀錄暨回覆辦理情形	附錄-1
附錄二、期中審查會議紀錄暨回覆辦理情形	附錄-4
附錄三、期末審查會議紀錄暨回覆辦理情形	附錄-8

表次

表 1-1 工作預定進度.....	1-6
表 3-1 顯著山崩潛感因子值域統計量.....	3-4
表 5-1 山地原住民及偏遠地區一覽表	5-6
表 5-2 模型一假說表.....	5-11
表 5-3 模型二假說表.....	5-12
表 6-1 應變數發包遲延(60 天)敘述性統計結果	6-1
表 6-2 應變數發包遲延(90 天)敘述性統計結果.....	6-2
表 6-3 應變數工程逾期敘述性統計結果.....	6-3
表 6-4 自變數工程區域敘述性統計結果	6-3
表 6-5 自變數季節敘述性統計結果	6-4
表 6-6 自變數工程規模之金額敘述性統計結果	6-5
表 6-7 自變數執行彈性敘述性統計結果	6-5
表 6-8 自變數汛期與工期敘述性統計結果	6-6
表 6-9 相關係數解釋參照表	6-7
表 6-10 模型一之變數相關係數矩陣(無控制變數)	6-7
表 6-11 模型二之變數相關係數矩陣(無控制變數).....	6-7
表 6-12 回歸模型之 BREUSCH-PAGAN-GODFREY 與 WHITE 檢定表 ..	6-9
表 6-13 模型一回歸結果整理.....	6-10
表 6-14 MODEL1 假說驗證表.....	6-11
表 6-15 模型二回歸結果整理.....	6-12
表 6-16 MODEL2 假說驗證表.....	6-13
表 6-17 模型一各區回歸結果整理(北區)	6-15
表 6-18 模型一各區回歸結果整理(中區)	6-16
表 6-19 模型一各區回歸結果整理(南區)	6-17

表 6-20 模型一各區回歸結果整理(東區)	6-18
表 6-21 模型二各區回歸結果整理(北區)	6-20
表 6-22 模型二各區回歸結果整理(中區)	6-21
表 6-23 模型二各區回歸結果整理(南區)	6-23
表 6-24 模型二各區回歸結果整理(東區)	6-24
表 6-25 新增資訊欄位建議表	6-34

圖次

圖 2-1 大數據實證研究流程圖.....	2-9
圖 3-1 廠商、工程品質關聯性模型-歷年工程進度資料.....	3-2
圖 3-2 廠商、工程品質關聯性模型-歷年的工程預算資料.....	3-2
圖 3-3 廠商、工程品質關聯性模型-歷年的抽驗與督導分數資料..	3-3
圖 5-1 計量經濟學的基礎理論.....	5-1
圖 5-2 計量模型設置流程圖.....	5-2
圖 5-3 水土保持局工程管考系統資料.....	5-14

第一章 緒論

1.1 計畫緣起與背景

近年巨量資料之應用進而研究成為趨勢，在資料驅動（data driven）的世界裡，整理、分析資料，並從中建立預測模型，成為學界與業界的新潮流。巨量資料的資料特質和傳統資料最大的不同是：(1)資料來源多元、(2)種類繁多、(3)大多是非結構化資料，而且(4)更新速度非常快，導致資料量大增，若要使用巨量資料創造價值，不得不注意數據的真實性。然而討論與應用巨量資料，其資料量的多寡並非唯一特點，而是探討有無能力在合理的時間內利用各種類型的資料，整理、分析出資料之間的脈絡與關聯性，進而產生具有影響力的決策。

水土保持工程管考系統自民國 90 年建置迄今，已累積超過 3 萬 5 千餘件工程執行相關資訊，然而過去管考系統多僅透過傳統基礎統計方式呈現，如某一計畫工程發包件數、發包率、總執行進度、進度差異值、預算執行情形、達成率及支用比等訊息。工程管考系統雖累積巨量資料，執行單位卻未了解其中蘊含有價值的知識，傳統上的統計值，無法傳遞管考系統資料之間所隱藏的關聯性，亦無法使執行單位了解。隨著巨量資料議題發酵及技術的進步，過去較難發現的執行累積規則與現象，有機會透過資訊匯集進一步分析、挖掘與探勘，並將原先無法預期之關係加以呈現，並且在工程專案上需要更有效率的決策與執行面之方針。

1.2 計畫目的

為使大數據及相關統計分析方法能妥善使用在水土保持局累積近 20 年、超過 3 萬 5 千餘件工程執行之資料，參考 104 年委辦計畫

-集水區保育治理巨量資料分析先期示範計畫及 104 年委辦計畫-水土保持工程管考大數據資料集分析暨智慧應用先期計畫，本研究

利用水土保持局工程管考系統資料庫為基礎，以大數據實證方法驗證目前水土保持局發包策略，並檢討目前策略之有效性，策略的目標在於透過發包策略之擬定，來控制工程專案流標及逾期情況。本研究以(1)發包遲延及(2) 工程逾期之因素探討作為大數據計量回歸模型之主題進行設計，避免以無關聯的資料產出較無相關之結果，並作為後續管理等決策應用之參考。

大數據計量回歸模型涵蓋大數據之特色，有別於傳統實證建立在已知或者有大概脈絡的假說上，大數據計量著重在探索性的性質上，即依經驗、直觀的想法，探索變數間可能之因果；亦具有傳統實證模型的要素，有別於許多不同大數據探索模型，大數據計量特別重視控制變數之納入與選取，以確保不會落入“偽因果”的陷阱，而在一般大數據探索上，並沒有辦法明確的知道變數間的因果關係與實際上的關聯性，因而較無法從其結果去進行後續策略制定等工作。

1.3 計畫範圍

本計畫目標係以提昇水土保持局資料分析之創新作為，透過計量回歸分析與巨量資料資料分析建置研究方法，針對本局長期執行且累積的巨量資料進行計畫目標擬定、方法論架構之建置及問題分析，並確認資料挖掘試辦範圍，首先以小範圍並有目標性之局部資料進行整理分析以建立相關流程與研究方法，後續增加分析資料的範圍，且針對本計畫所分析的目標，鎖定時間區間，收集全數資料進行整理分析，以局內相關類型巨量數據進行相關研討以作為後續工作推展之參考

依據。

本計畫第一部份為針對水土保持相關工程歷年管考資料(103-108)進行巨量分析之先期研究，擬將研究範圍設定於台灣北部地區收集歷年包含流域綜合治理計畫、整體性治山防災計畫、重劃區外緊急農路設施改善與農村再生社區發及環境改善計畫等各項不同類型之水土保持工程管考資料進行(1)發包遲延及(2)工程逾期之巨量資料探勘與實證分析。

第二部份擴大資料範圍及規模，將研究範圍設定為工程歷年管考資料(103-108 年)全台灣地區，包含管考系統所分之四類(流域綜合治理、重劃區外農路工程、農村再生社區發及環境改善、整體性治山防災計畫工程)計畫資料，進行上述之模型實證。

而原先期中報告所述之第二部份，為針對大尺度全台灣崩塌規模及中尺度流域土砂量為分析目標，進行崩塌率及流域土砂量等重要因素探討與實證分析，並對比於原本委辦計畫之差異，以此萃取其最重要之可能性，以此為基礎建立巨量資料之分析。在與 104 年委辦計畫-集水區保育治理巨量資料分析先期示範計畫的執行單位接洽計畫資料後，本研究團隊認為大尺度崩塌資料之領域與本次計畫的專業與類型差異較大，原先資料類型與本研究模型所採用之類型有異，需更充裕的時間了解資料的分類、意涵與使用方法，因此將原先預定進行大尺度全台灣崩塌規模及中尺度流域土砂量分析之研究內容從本計畫中移除，專注於管考系統資料的巨量資料處理分析與回歸模型之實證。

1.4 計畫項目與內容

一、水土保持工程管考系統資料分析

1. 文獻回顧：本研究基於 104 年委辦計畫-水土保持工程管考大數據資料集分析暨智慧應用先期計畫之執行過程，採用專家應用需求訪談之回顧，透過其計畫彙整之各單位意見與建議，作為本研究回歸實證與大數據分析應用之參考方向。
2. 變數及模型確立：經文獻回顧萃取研究方向及主題後，本研究首要進行假說設立，並以此設計應變數、自變數、控制變數，建置大數據計量實證模型。
3. 工程管考系統資料整理：針對水土保持工程管考系統內之相關資料進行蒐集、整理、歸類，以確立本計畫工程資料關聯性分析之主題與方向，並確保資料完整性與正確性。本研究欲使用工程管考系統內四個種類之資料（流域綜合治理、重劃區外農路工程、農村再生社區發及環境改善、整體性治山防災計畫工程）之歷年資料（103-108 年），並採用北區、中區、南區、東區作為區域分類之依據，進行大規模分析。
4. 資料合理性評估：經由資料整理過後，篩選工程管考系統中因誤值、未更新或漏值之資料進行剔除，以符合資料處理之需求、確保後續資料分析之準確性。
5. 模型分析與討論：本研究設置之模型進行分析、並針對分析結果之數據(包含回歸係數、p-value、adjusted R-square)進行解釋與討論，並透過其於檢定，確認結果正確性。
6. 策略擬定與系統建議：本研究於模型實證討論完後，將針對管考系統現況與本研究討論之觀點進行綜合比較，並針對水

保局管理層面欲了解之現象，進行新增建議與專案發包、執行狀況改善之策略建議。

1.5 研究架構

本研究結案報告書分為七個章節，章節內容如下：

- (一) 緒論，介紹本研究之研究緣起與背景、目的、範圍，同時簡述本研究之執行項目與內容，最後介紹本研究之整體架構、執行進度安排及預期結果。
- (二) 大數據探索性實證分析：詳述本研究之工作執行構想，及所有研究方法和流程，包含專家訪談及大數據計量實證分析。
- (三) 文獻回顧：此章節分為三大主題，分別針對過往計畫探討、專家訪談之整理、及相關研究討論，作為後續實證模型之基礎。
- (四) 探索性實證模型假說：基於本研究欲探討水土保持局內部對於工程執行狀況之看法，本研究建立相關之假說，於以實證局內看法之真實性。其中包含影響發包遲延和影響逾期情況之假說。
- (五) 探索性實證模型研究設計：本研究之計量模型進行變數、模型設置，及簡述蒐集之資料類型。
- (六) 大數據分析結果與討論：針對數據之分析結果，進行深入討論本研究之實證結果與真實看法之差異，並點出本研究之貢獻。
- (七) 結論與建議：對本研究之期末階段進行整體性之結論整理，並提出後續建議。

1.6 計畫執行進度

表 1-1 工作預定進度

工作項目 \ 月次	2月	3月	4月	5月	6月	7月	8月	9月	10月	11月	12月
計畫執行期間											
文獻回顧											
工程管考系統資料分析											
-變數、模型確立											
-北區資料蒐集											
-分析結果討論											
-全區資料蒐集、整理											
-分析結果討論											
-策略擬定與方針建議											
期中報告階段						✓					
期末報告階段										✓	
繳交成果報告											✓

1.7 預期成果

在擬定計畫目的以及預期作業關聯下，本計畫期望：

- (一)以巨量資料應用為基礎，引領水土保持局得知過去於工程管理與治理上難以掌握的確證之模糊操作，未來可由大數據或回歸分析模型之結果，協助判斷，減少資源投入之不確定性並提升資源效益。
- (二)針對過去為人詬病之發包與計畫執行進度期程問題，探討其影響之重要因素，協助決策者了解有關變數間之關聯性，於未來決策、管理、及相關業務上能有參考之依據。

第二章 大數據探索性實證分析

2.1 研究執行構想

本研究以水土保持工程管考系統資料庫為基礎，選定先期試辦區域，為研究標的，運用水土保持歷年目錄、工程資料及全生命週期執行結果，以各執行工程為分析單元，透過(1)資料探勘方法剖析創建相關議題之假說；(2)透過大數據計量分析技術實證前述假說之正確性，其結果將可提供辦理如：管考資料關聯性探討、工程資源調配、發包策略擬定、專案執行計畫制定及成效分析等相關之決策應用參考。

資料探勘技術乃透過巨量資料尋找資料中所暗示或存在之可能因果關係或規則，再針對複雜度高之議題上，其結果常可解決傳統物理性或理論性探討諸多簡化假設下之限制。然而巨量資料之探勘所解析出之結果卻也常落入統計巧合之謬誤，因此，我們提出以計量分析實證研究方法，透過假設檢定與統計推論之原理，針對巨量資料探勘之結果進行實證分析，以確認探勘所得結論之正確性與實證證據，避免資料探勘上可能之統計巧合與謬誤。

「工程管考系統」係將水土保持局所有自辦或委辦工程進行電腦化管理，舉凡經水土保持局核定之工程或計畫均集中儲存於水土保持局之資料庫伺服器，各執行單位可透過網際網路方式回報各項工程之執行狀況、每月進度、預算經費、品質抽驗及施工照片等資料，每月

月底，各執行單位可直接透過網路技術，上傳會計報告至水土保持局主計室，系統自動將工程明細資料彙整為會計報告，提供主計人員核對工程帳務。

本計畫擬建議分析方向以大數據資料之「變數關聯探索」為主，「關係實證」為輔，圍繞於此兩大重點。關聯分析是指如果兩個或多個事物之間存在一定的關聯關係，那麼就能透過之間的關聯性發掘變數間之深層意涵，它的目的是為了挖掘隱藏在模型變數間的相互關係，以此制定策略。經由資料清查動作後，萃取可能會影響發包順利及工程進度因子欄位，並透過適合的分析方法建立模型，以此探索並實證關鍵影響因素。

2.2 研究方法

2.2.1 專家訪談法

本研究欲了解影響局內部管理及決策面問題之主要因素，包含發包遲延及工程逾期情況，故訪談對象以水土保持局的長官及具有討論決策之人員，探討目前實際工程專案執行上之疑慮，及有興趣的關聯性。因本研究基於 104 年委辦計畫-水土保持工程管考大數據資料集分析暨智慧應用先期計畫已完成此專家訪談，故本研究參考此份計畫之訪談結果，作為本研究專家訪談之參考文獻。

2.2.2 大數據分析方法

大數據 (Big Data) 又被稱為巨量資料，概念是過去 10 年廣泛用於企業內部的資料分析、商業智慧 (Business Intelligence) 和統計應用之大成。因為資料量急速成長、儲存設備成本下降、軟體技術進化和雲端環境成熟等種種客觀條件就位，方才讓資料分析從過去的洞悉歷史進化到預測未來。一般而言，大數據的定義是 Volume (容量)、Velocity (速度) 和 Variety (多樣性)，但也有人另外加上 Veracity (真實性) 和 Value (價值) 兩個 V (Anuradha, 2015)。但其實不論是幾 V，大數據的資料特質和傳統資料最大的不同是，資料來源多元、種類繁多，大多是非結構化資料，而且更新速度非常快，導致資料量大增。而要用大數據創造價值，不得不注意數據的真實性。然而大數據分析裡包含許多種類，包含 Text mining, Web mining, Graphic mining, Natural Language 等等，所使用的方法亦不全然相同。類神經網路 (Neural Networks) 是整個大數據分析、機器學習、深度學習的根本。本計畫尚未決定採用何種類神經網路，包含 K-NN, SVM, LSTM 等等，而最常被使用的 CNN (Convolutional Neural Networks) 亦是處理較複雜的資料所採用的方法，透過擷取特徵值，進行 Model Training，並達到預測跟自動辨斷。不同資料類型採用不同的分析方法則可發揮不同的功能與效益，如結構式資料搭配回歸分析適合用於探索資料中設

計變數之關聯性與潛在現象、非結構式資料搭配類神經網路，適合用於辨識(Identification)、生成(Generation)、預測(Prediction)模型等。

大數據優點在於透過大量數據來改善決策，能對業務的分析與預測越來越準確；而缺點包含數據質量的真實性與完整性、個人私有數據的洩露風險等，另有文獻提出反對資料的探勘取代專家所提出的理論假說(Oguntimilehin & Ademola, 2015, Bottles et al., 2014)。

2.2.3實證研究法

計量實證由計量經濟學而來，是建立在統計方法的發展上估計經濟的關係、檢定經濟理論，以及評估與執行政府及企業之政策。計量經濟學一般的應用是預測一些重要的總體變數，諸如利率、通貨膨脹率，以及國內生產毛額(GDP)。除了常見的總體經濟指標預測之外，計量經濟方法亦可用於其他的領域(ex.在教育領域中探討學校支出對學生表現的效果)。此方法與數理統計有所不同，前者集中焦點在收集與分析非實驗性之經濟資料，非實驗性資料(觀察性資料或回溯性資料)之收集並非透過對個人、公司或經濟部門所做之可控制的實驗而來；實驗性資料在自然科學中常在實驗室環境中被收集，但在社會科學中非常難以取得。通常要從事這種可控制的實驗是不可能、成本很高或是道德上不可行的。本質上，計量經濟學是盡可能利用數理統計方法，複回歸分析在此領域中是主幹，當想對某些企業決策或政策分

析很重要之相關因素時，即可使用此方法。實際分析(empirical analysis)即是利用資料來檢定理論或是估計某個關係。而實證分析的步驟大略為：定義題目，確立估計模型與參數，訂定假說，對這些參數進行假設檢定，並進行在其他條件不變的觀念下之因果推論。

透過專家訪談法建立基礎認知，以決策者及關心之疑慮之角度出發，使用工程管考系統之資料建立假說及計量模型，探討(1)流標(工程發包順利性)與(2)逾期(工程執行情況)之議題，影響其狀況之主要因素為何。透過結果的分析，了解不同因素影響決策的權重為何，讓未來決策單位對於工程管理、決策制定能有更進一步的參考。

2.2.4大數據計量分析

大數據經濟學(Big Data Economics)是在經濟學研究和採用大數據，並且採用大數據思想對傳統經濟行深化的新興交叉學科。大數據經濟學不僅要研究如何建模、管理和應用大數據，而且要深入研究傳統經濟學如何應對大數據帶來的挑戰並進行改良，大數據經濟學需要經濟學家、領域專家和信息技術專家等密切合作，對人文社科與自然科學的跨學科研究提出了更高的要求，並且對整個經濟學、社會學、公共管理等將帶來革命性變革。大數據經濟學的研究內容包含：(1)大數據計量經濟學，這是和傳統計量經濟學對應的一個學科，也是大數據經濟學下的子學科，在大數據背景下，經濟學建模與分析方法與傳

統計量經濟學完全不同，迫切需要採用全新的思路和方法進行研究。

對信息技術專家們而言，大數據經濟學僅僅是演算法和建模問題，但是如果沒有經濟學理論指導，沒有經濟學家的思維，必然會導致研究方向的迷失。一些大數據領域的學者認為“要相關，不要因果”，這是非常要不得的，傳統經濟學理論至今仍然到處閃爍著智慧的光芒，對經濟現象的深入見解時刻發揮著重要的作用，所以大數據背景下的經濟學分析不能主要靠信息技術的建模專家來進行，必須繼續依靠大數據計量經濟學家；(2)大數據統計學(Big Data Statistics)，如前所述，大數據給統計學帶來的挑戰是革命性的，在某些領域，傳統統計學所採用的抽樣調查方式必將徹底淘汰。此外，傳統統計學所要求的精確數據與數據加工方式可能是畫蛇添足甚至敗筆之舉，人們更加重視對一手數據而不是經過加工過的二手統計數據進行分析。大數據時代，人們更加關注原始數據、關注半結構化甚至非結構化數據，瀏覽記錄、查詢關鍵詞、微薄文字、照片等等都是寶貴的數據資源。在大數據時代，傳統統計學也必須進行變革，對數據儲存手段、處理設備、處理方法都提出了新的要求；(3)大數據領域經濟學，包括大數據生態經濟學、大數據環境經濟學、大數據金融學、大數據城市經濟學、大數據工業經濟學、大數據農業經濟學、大數據交通經濟學、大數據建築經濟學、大數據商業經濟學、大數據信息經濟學、大數據人口經濟學等

學科，借用大數據的思想和技術來進行各應用經濟領域的研究。

在以上大數據經濟學的各學科中，大數據統計學是基礎，大數據計量經濟學是研究方法，而大數據領域經濟學是具體的運用，他們之間存在著密切的共生關係。大數據由於是基於總體的，很大程度上解決了傳統巨集觀經濟學與微觀經濟學缺乏較強邏輯聯繫的問題，此外大數據對傳統計量經濟學帶來的一個有益之處就是，結構化的大數據更加接近常態分佈，這樣就降低了小樣本假設檢驗失效問題。

2.2.5 小結

經由上述採用方法之討論，本研究在研究方法中採用大數據經濟學的理論、透過大數據統計學作為基礎，並以大數據計量經濟學作為研究方法。與大數據分析方法不同點在於，大數據計量經濟學採用迴歸模型、過程中需要專家去協助判斷，由專家先判斷有用的欄位，而並非將所有的資料全部放進分析模型中，以此作出變數間關係的釐清，增加模型的精準度，判斷出可能的真實解釋。

如同 1.2 計畫目的所述，大數據計量迴歸模型包含大數據之探索性，有別於傳統實證建立在已知或者有大概脈絡的假說上，即依經驗、直觀的想法，探索變數間可能存在之關聯性及因果；亦保有實證模型的基本要素，著重於應變數、自變數的設計與控制變數的選擇，大數據計量特別重視控制變數之納入與選取，以確保不會落入“偽因果”

的陷阱，進而做出錯誤的判斷。在一般大數據探索上，並沒有辦法明確的知道模型變數間之因果關係，亦沒有辦法與實際狀況進行比對，因而較無法從其結果去進行後續策略制定等工作，亦可透過大數據計量的探索與實證，確立模型變數的正確，以增加分析結果的準確性。

本研究經資料整理後，建立的計量模型，透過 Eview8 軟體跑回歸分析結果，得出因子的顯著性來驗證假說。本研究採用大數據計量模型來做為研究主題的原因為：大數據計量模型可以進行大範圍的探索，並透過將原生資料量化的方式觀察因素之間的重要程度與之間的顯著性，以評估問題、優勢及關鍵因素對探討議題的影響程度。

2.3 研究流程

本研究之流程請參閱圖 2-1。

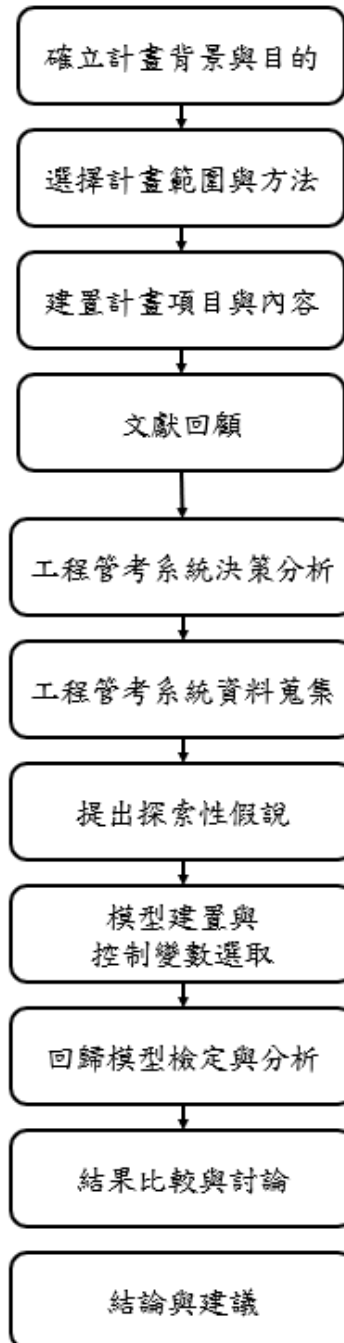


圖 2-1 大數據實證研究流程圖

第三章 文獻回顧

3.1 過往計畫探討

一、104 年委辦計畫-水土保持工程管考大數據資料集分析暨智慧應用先期計畫

此計畫以「工程管考系統」資料庫為基礎，透過大數據分析技術解析歷年工程資訊，用以決策支援參考，此計畫包含實作面與規畫面的部份，本研究僅針對實作面之計畫成果進行探討。

實作面部份，此計畫將工程生命週期施工階段的進度及品質相關資料，導入其所建置之工程進度預警模型、廠商與工程品質關聯性分析模型，期望對於進度不符之計畫或工程預先提出警訊，可提供補救措施或策略調整，並由歷史資料推估各監造與施工廠商的施作品質，作為工程品質管控參考。其中進度預測模型運用到 s 曲線多項式與倒傳遞類神經網路演算法，其中變數欄位包含地點、工期、經費、工程類型，皆為工程管考系統內之資料欄位；其計畫模型已經過 10000 件歷史工程之驗證，達 90% 以上的準確率。而廠商與工程品質關聯性分析模型依水保局歷年的抽驗與督導資料，推估出各廠商未來品質抽驗與督導結果，以約 6000 件資料驗證，達 95% 以上的準確率，此計畫期望透過此分析結果協助水保局及各分局即早掌握各計畫及工程期限內未來各月份進度趨勢，能精準的管理各項工程，有效管理風險與

計畫執行率。

工程編號	年度	月份	預定進度	實際進度		工程編號	年度	月份	預定進度	實際進度	
eng_id	Pyear1	Pmonth1	preprg	prg	progress_lastmod2	eng_id	Pyear1	Pmonth1	preprg	prg	progress_lastmod2
100-ADE-01-4-051	100	4	7	20	2011/4/28 上午 11:00:59	100-ADE-01-4-051	100	4	7	20	2011/4/28 上午 11:00:59
100-ADE-01-4-051	100	5	31	54	2011/5/26 下午 01:48:38	100-ADE-01-4-051	100	5	31	54	2011/5/26 下午 01:48:38
100-ADE-01-4-051	100	6	53	90	2011/6/30 下午 04:33:49	100-ADE-01-4-051	100	6	53	90	2011/6/30 下午 04:33:49
100-ADE-01-4-051	100	7	77	95	2011/7/29 下午 04:41:22	100-ADE-01-4-051	100	7	77	95	2011/7/29 下午 04:41:22
100-ADE-01-4-051	100	8	100	100	2011/8/18 下午 04:35:59	100-ADE-01-4-051	100	8	100	95	2011/8/18 下午 04:35:59

進度符合: 當工程最後一月份, 預定進度=實際進度 進度落後: 當工程最後一月份, 預定進度>實際進度

工程編號	年度	月份	預定進度	實際進度	
eng_id	Pyear1	Pmonth1	preprg	prg	progress_lastmod2
100-ADE-01-4-051	100	4	7	20	2011/4/28 上午 11:00:59
100-ADE-01-4-051	100	5	31	54	2011/5/26 下午 01:48:38
100-ADE-01-4-051	100	6	53	90	2011/6/30 下午 04:33:49
100-ADE-01-4-051	100	7	77	100	2011/7/29 下午 04:41:22
100-ADE-01-4-051	100	8	100		2011/8/18 下午 04:35:59

進度超前: 工程實際進度100的月份, 早於預定進度100的月份

圖 3-1 廠商、工程品質關聯性模型-歷年工程進度資料

資料來源：陳美心、周天穎、林峰正、林伯勳(2015)

工程編號		工程款_發包				工程款_變更設計					
Eng_id	run_type	Des1	Tot_des	Send1	Tot_Send	Cost	Changel	Tot_chang	Acc1	Tot_acc	Tot_pay
100-ADE-E	7228618	7937617	5778000	6385749	6550000	922000	1177896	922000	1177536	1177536	
100-ADE-E	112720	117812	107000	113194	112000	0	0	107000	113194	113194	
100-ADE-E	2253000	2498826	1793000	2000515	2127800	0	0	1758000	1961569	1948021	
100-ADE-E	3610000	3929985	2893000	3167319	3426000	2370300	2598138	2370300	2598138	2598138	
100-ADE-E	2505023	2735685	2050000	2241875	2368000	2049847	2241640	2049847	2241640	2226684	
100-ADE-E	2271427	2482851	1721000	1885609	2156000	1692652	1855659	1692652	1855659	1832300	
100-ADE-E	2465000	2684027	1968000	2152085	2318000	1965000	2148814	1965000	2148814	2121697	
100CR3-2-E	4268000	4601988	4059129	4400000	4200000	0	0	4059062	4400000	4390278	
100CR4-0-E	3181000	3415326	2540000	2787247	2978000	0	0	2540000	2787247	2787247	
100CR4-1-E	1089353	1186406	769000	850097	1002000	0	0	769000	850097	850097	
100CR4-2-E	2477956	2708590	2030000	2237177	2278000	2023028	2229511	2019898	2226108	2226108	
100FR300-E	6533000	7167129	5376000	5957871	5980000	0	0	5376000	5957869	5957869	
100FR302-E	7794300	8653457	7680000	8565127	7790000	0	0	7680000	8565127	8565127	
100FR311-E	4521612	4623658	4023500	4023500	3750000	0	0	4023500	4023500	4023500	

工程款追加: 工程款_變更設計>工程款_發包
 工程款不變: 工程款_變更設計=工程款_發包
 工程款減少: 工程款_變更設計<工程款_發包

圖 3-2 廠商、工程品質關聯性模型-歷年的工程預算資料

資料來源：陳美心、周天穎、林峰正、林伯勳(2015)

資料時間: 民國90年~103年

抽驗: 6007件

督導: 4980件

工程編號	施工廠商	監造廠商	時間	結果	督導或是抽驗
A	B	C	D	E	F
100-AD8-15-2-123	雲龍營造有限公司	綠環工程顧問股份有限公司	2012/1/3 00:00	93	抽驗
100-AD8-24-3-157	德勝營造有限公司	清隆工程顧問有限公司	2012/1/4 00:00	84	抽驗
100-AD8-29-3-181	政達營造工程股份有限公司	亞隆工程技術顧問有限公司	2012/3/21 00:00	82	抽驗
100-AD8-30-3-183	康誠土木包工業有限公司	綠環工程顧問股份有限公司	2012/5/17 00:00	81	抽驗
100FR3005-016	理昇營造有限公司	建昌工程顧問有限公司	2011/1/22 00:00	89	抽驗
100FR3005-026	開弘工程有限公司	山科技術顧問有限公司	2011/9/25 00:00	80	抽驗
100FR3164-001	上什營造有限公司	寶全工程顧問有限公司	2012/3/22 00:00	87	抽驗
100FR3215-009	弘盛土木包工業	興隆工程顧問股份有限公司	2011/1/5 00:00	80	抽驗
100FR3245-012	輝豐營造股份有限公司	山河工程顧問有限公司	2012/5/29 00:00	85	抽驗
100FR3330-011	宏興營造廠	源隆技術顧問有限公司	2012/7/10 00:00	83	抽驗
100FR3436-019	仕林營造有限公司	鴻宇工程顧問有限公司	2012/5/22 00:00	79	抽驗
100FR3523-001	匯凱土木包工業	名義工程顧問有限公司	2012/6/1 00:00	80	抽驗
100MFM001-009	重仁營造有限公司	楊其弘土木技師事務所	2011/8/3 00:00	81	抽驗
100MFM002-003	萬興土木包工業	錦峰工程技術顧問有限公司	2011/4/12 00:00	80	抽驗
100-RADE-06-3-002	致立營造有限公司	權成測量有限公司	2011/1/26 00:00	77	抽驗
100S-9C-1-461-006	品盛營造有限公司	森德工程技術顧問有限公司	2011/6/16 00:00	83	抽驗
100S-9C-1-467-001	路宜營造有限公司	四海工程顧問有限公司	2012/1/11 00:00	84	抽驗
100S-9C-1-403-003	同家營造有限公司	水土保持防治室	2011/7/22 00:00	89	抽驗
100S-9C-2-404-002	鑫發營造有限公司	山河工程顧問有限公司	2011/9/28 00:00	82	抽驗
100S-9C-4-404-005	古那營造有限公司	楊其弘土木技師事務所	2012/5/30 00:00	75	抽驗
100S-9C-4-401-001	平谷營造股份有限公司	乾利技術顧問股份有限公司	2011/9/21 00:00	82	抽驗
100S-SME-1-32-452	東島營造有限公司	陳新南水土保持技師事務所	2012/5/24 00:00	89	抽驗
100S-WF-1-001-003	百豐營造股份有限公司	勇輝工程顧問有限公司	2011/9/9 00:00	89	抽驗
100S-WF-1-301-002	振源營造有限公司	東和工程顧問有限公司	2011/7/26 00:00	80	抽驗
100S-WF-1-116-001	偉洋營造有限公司	智全工程顧問有限公司	2012/5/16 00:00	82	抽驗
100S-WF-3-M01-011	至力營造有限公司	山河工程顧問有限公司	2011/6/3 00:00	89	抽驗

抽驗結果: 合格/不合格
合格: ≥ 60
不合格: < 60

資料轉換

督導結果: 甲/乙/丙
甲: > 85
乙: $75 \sim 85$
丙: < 75

圖 3-3 廠商、工程品質關聯性模型-歷年的抽驗與督導分數資料

資料來源：陳美心、周天穎、林峰正、林伯勳(2015)

二、104 年委辦計畫-集水區保育治理巨量資料分析先期示範計畫

本研究針對計畫之關聯性分析相關模型進行探討。此計畫運用歷年山崩目錄、地文資料、數值地形與降雨資料，透過資料探勘技術，建立全國尺度集水區潛在崩塌發生區位評估模式，以了解全國集水區隨降雨變化崩塌率變動情形、崩塌熱區及易致災集水區分布，輔以既有工程與投資經費關聯性分析，以流域單元歸納災害、工程與投資效益關聯性，預先掌握治山防災先機，其進行之相關分析包含：(1)全國尺度集水區潛在崩塌發生區位評估模式與結果(2)工程分布與經費分布分析(3)災害損失與投資經費關聯性分析(4)工程經費與災害分布關聯性分析。其中對於山崩潛感因子之顯著性檢定，歸納出地形因子(高程、坡度、坡向、曲率等)、地質因子(地層、斷層、順向坡等)、水文

因子(水系、地形濕度指數、雨量)及人文因子(道路開發)等項次，而經以崩塌地作為應變數，山崩潛感因子作自變數，並以 $p\text{-value}<0.05$ 為顯著與否之判斷依據，此研究結果顯示(1)高程、(2)坡度、(3)正弦坡向及餘弦坡向、(4)地形曲率、(5)地形(濕度)指數、(6)地形粗糙度、(7)坡度粗糙度、(8)地形崎嶇指數、(9)斷層距離、(10)水系距離、(11)稜線距離、(12)道路距離、(13)地層、(14)順向坡及(15)最大 24 小時雨量共 15 項為顯著變數，因子值域統計結果如表 3-1 所示，以此作為後續預測模型建置應用。

表 3-1 顯著山崩潛感因子值域統計量

資料來源：國立臺灣大學 104 年委辦計畫(2015)

變數	單位	範圍	平均值
(1)高程	m	22~3,872.7	1,947.4
(2)坡度	%	0~3,245.6	1,622.8
(3)正弦坡向及餘弦坡向	rad	-1~1	0.0
(4)地形曲率	1/m	-1,057.2~696.0	-180.6
(5)地形(濕度)指標	m	1.9~16.5	9.2
(6)地形粗糙度	m	0~130.5	65.3
(7)坡度粗糙度	%	0~559.4	279.7
(8)地形崎嶇指數	m	0~123.8	61.9
(9)斷層距離	m	0~22,931.9	11,466.0
(10)水系距離	m	0~2,352.9	1,176.5
(11)稜線距離	m	0~2,251.1	1,125.6
(12)道路距離	m	0~10,043.0	5,021.5

(13)地層		類別	共 26 類地層	-
(14)順向坡		binary	0(非順向坡) 1(順向坡)	-
(15)最大 24 小時雨量	莫拉克颱風	mm	548~1,448	998
	100 年重現期	mm	554~1,108	831

然而其中僅針對 p-value 進行討論，而尚提及迴歸模型、迴歸係數、模型解釋程度，並且以因子卡方檢定進行因子辨定，是否控制變數間潛在相互影響之效果不得而知，嚴謹程度稍嫌不足。在本研究進行完文獻回顧後，與 104 年計畫執行團隊接洽，欲取得其計畫執行資料，以供本研究在同樣基礎下進行分析。然後在本團隊了解資料後，其資料類型與本計畫研究方向有異，又其研究領域非本計畫專業，需更深入探討，在時間規劃上較無法克服，因此移除此部份後續研究。

三、水土保持工程管考系統擴充及維運管理計畫

此計畫為 109 年水土保持局執行計畫，本研究因探討到工程管考系統之資料，在與水土保持局人員討論，由局處提供此執行計畫書供本研究參考，並針對其計畫提出之工程管考系統現有功能與資料盤點、檢討，與資料議題探勘及研究進行討論。

針對此計畫所提出之“工程管考資料議題探勘及研討”，其提出工程管考系統資料結合相關外部系統開放資料，藉以優化水土保持工程施政決策品質，並透過訪談、問卷方法，擬定數據分析方法、架構、目標。其所提出的資料分析議題研擬包含：工程核定與發包（工程核

定、工程發包、廠商優選）、工程預算（核定底價、工程經費追加、影響預算執行率）、工程執行（停工因素、用地取得與工程進度、工程構造物延壽）、工程品質（優良廠商、督導比例、廠商抽驗比例、工程抽驗結果與空間環境）。

3.2 專家訪談

本研究探討104 年委辦計畫-水土保持工程管考大數據資料集分析暨智慧應用先期計畫第二章第二節應用需求訪談之內容，問卷設計主針對工程管考系統現有欄位資訊，延伸至工程發包、工程預算、工程執行與工程品質等議題，此問卷結果包含可量化舉不可量化兩類資料，可量化資訊供後續相關統計分析及分析主題發想之參考；不可量化資訊供分析主題發想之參考。針對回收問卷意見，本研究在探討所有意見後彙整，可量化類型資訊可分為以下幾類(陳美心等，2015)：

1. 工程發包

以大數據分析方法，有(1)各分局(或各鄉鎮)之工程合理單價預測及(2)依據工程所在區位，輔助概估所需合理經費，以利發包作業。對於計量迴歸分析方法，有(1)了解過去無法發包工程之原因(如：標比偏低、工程施作區域偏遠、鄰近缺乏承攬廠商等)；(2)不同區域發包是否影響完工難易度，由歷史發包紀錄探索不同區域(或各分局)不同預算級距之流廢標比例。

小結：透過原有工程管考系統內之系統化資訊，得知過去無法發包工程之原因。

2. 預算

了解辦理工程經費變更追加工程之影響因素。

3. 執行

以計量迴歸方法有(1)哪種類型工程發包較容易完工(工程發包類型與完工之關聯性)；(2)哪種類型工程進度易受天候影響而有工期延宕；(3)無法如期完成之工程，是否受地形地勢複雜所影響？是否有包含變更之現象？；(4)用地取得與工程期程延遲、取消施作、變更設計之關聯性；(5)不同經費級距之工程是否受歷史氣象資料及強降雨情況造成工期延遲。

小結：經專家訪談回饋，本研究認為可針對過去延期進行因素探討，涵蓋上述之變因。

4. 品質

以計量迴歸方法有(1)工程抽驗結果不合格在空間分布具有某種關聯性；(2)標比 65% 以下者之工程承包廠商，與工程抽驗結果不合格之關聯性；(3)依據不同區域(或各分局)預算級距之工程抽驗結果不合格之關聯性；(4)曾工程抽驗結果不合格、遭全民督工通報、或曾發生重大工安事件之承攬廠商，是否影響工程品質；(5)全民督工通報是否

皆位處鄰近省道、縣道或鄉道等交通易達區域；(6)工程管考系統之竣工資料建置較完整之承攬廠商，與其工程經費、期程、工程抽驗結果之關聯性。

小結：綜觀上述，由 104 年委託計畫中之受訪者期望由分析工程抽驗結果不合格之承攬廠商是否具有前例，以及工程抽驗結果不合格在空間分布具有某種關聯性或重複發生，然而因類型之資料於工程管考系統內較不易蒐集，因此品質相關的議題發想僅作為參考。

3.3 相關研究討論

針對招標發包之問題，戴貞德，林信宗於其研究「政府採購雙贏決策之研究-以台電高雄區營業處為例」中探討不同人員對專業能力、採購決策、招標成效及非專業考量的認知差異，其中人員考量不同擔任職務、不同年資、不同服務單位與不同往來機構之因素，專業能力包含進階訓練者、受基礎訓練及未曾受訓者；採購決策包含公平性、公開度、採購效率、採購品質、採購成本；招標成效包含機關績效、機關形象。其分析結果表示，不同擔任職務人員對(1)採購決策(2)招標成效(3)非專業考量各變項的認知有差異性存在，亦指機關主計人員、採購機關業務單位人員、工程承包商、財物供應商、勞務承攬廠商對的認知有所差異；而不同年資對於採購決策各變項的認知在公開招標時有顯著差異。有別於本研究尋求工程專案相關資料與發包順利與否，

此研究著實探討承辦人員之條件與發包順利與否之關係(戴貞德、林信宗，2009)。

針對工程逾期之問題，李振榮於其研究「工程逾期造成工程爭議問題之研究中」提及，由於營建工程涉及的範圍廣闊，在工程進行中難免會遇到不可預知的情況發生，而導致工程被迫展延甚至停工之可能，因此工程無法順利完工，難免業主與承包商雙方會有爭議發生。利用【司法院資料檢索系統】取得有關工程展延訴訟之判決書，藉由判決書內文得到關於工程展延之相關因素及展延費用請求之舉證，再由失誤樹分析（Fault Tree Analysis）收斂之最小切集合之結果探討其主要造成展延之原因，其最終分析結果表示，展延以變更設計、天氣因素、未履行契約、無法如期交付工地、施工品質上的瑕疵、施工進度遲延、管線遷移、位址搬遷、關聯承商之延誤、當地居民抗議等因素，其中以變更設計因素發生機率為高(李振榮，2014)。王華弘，謝彥安在「道路鋪面工程中展延工期問題之研究」提及道路鋪面工程因牽扯介面繁多，更容易遭遇工程延誤的問題，如於施工初期可能遭逢：業主計畫用地徵收延宕、地下管路調查作業或協調遷移曠日廢時、地上物拆遷作業申請許可執照繁瑣等因素；施工中則可能有：天候因素、業主變更設計、業主要求不同檢驗標準、開挖後發現額外廢棄物清運處理等狀況，加上外在環境、物價變動、勞工缺乏等風險(王華弘、謝

彥安，2016)。

本研究參考以上資料，包含水保局過往委託計畫、整理計畫內之專家訪談、及部份學術論文確立本計畫之研究方向，本計畫將著重於(1)是否順利發包之因素探討及(2)工程執行狀況(逾期與否)。而在 109 年水土保持管考系統計畫中提到類似的主題，本研究首要依據 104 年專家訪談整理之結果，確認此兩個研究方向，並且所選擇的研究方法理論與變數的選擇與上述 109 年執行計畫有所不同，因而將其 109 年之執行計畫書認定為參考文獻，並完善本研究發包及工程執行狀況因素探討之實證模型。

第四章 探索性實證模型假說

4.1 假說推導

本研究旨在於實證局內對發包策略之看法與原先認為之影響因素，以此看法建立假說以完成證實及因素討論，期望提出原先實務看法之重要程度與相關解釋，基於水土保持局內執行人員、長官對於發包策略議題之看法，佐以第二章文獻回顧專家訪談之回覆統整結果，建置實證模型假說。方向為二：(1)影響發包遲延之因素探討及(2)影響逾期情況之影響因素，並參照工程管考系統之資料欄位，確立探討之因果關係，進而推導本研究之假說，以下針對兩實證模型進行假說之說明。

4.2 影響發包遲延之假說

1. 工程區域：綜合專家訪談之回饋，受訪者對於執行工程之區域對於發包作業之關係存在著一定的疑慮，亦可能因為工程的地區偏遠、當地資源、環境潛在風險而影響投標廠商之意願，故本研究認為工程區域與發包遲延之情況有一定的相關性。

H1a：工程區域對於發包遲延之情況有顯著正相關，亦指工程區域越偏遠，發包遲延情況上升

2. 季節：不同的季節對於公家機關而言，可能存在著不同的專案招標的習慣，亦或是機關有著不一樣的規定，以致於在某個季節存在著較大量的執行專案或容易頻繁的進行工程施作，也就是所謂的“淡”、“旺”季的差異。本研究認為，發包遲延與何時發包可能存在著一定的相關性。

H2a：季節對於發包遲延之情況有顯著相關，亦指不同的季節間

存在著順利發包與否之差異

3. 工程屬性：不同的工程類型與屬性，存在著不一樣的工程內容、設計與要求，就水土保持局近幾年的計畫類型來說，包含流域綜合治理計畫、重劃區外緊急農路設施改善計畫、農村再生社區發展及環境改善及整體性治山防災計畫，些許計畫的類型相似，但依舊存在著差異，故本研究認為不同的工程屬性應與發包遲延之情況有一定的相關性。

H3a：工程屬性對於發包遲延之情況有顯著相關，亦指不同的工程屬性間存在著順利發包與否之差異

4. 工程規模之金額：發包對於專案的工程規模而言，有著非常直觀的關聯性，工程的大小時常需考慮招標廠商的品質、財務狀況、執行能力，以致影響招標廠商對於承攬此專案的意願，另外，基於專家訪談的回饋，受訪者對於欲招標的工程存在著以大規模規劃執行的想法，以過往經驗認為規模越大的專案，廠商承攬的意願越高。故本研究認為工程規模之金額對於發包遲延之情況有顯著負相關。

H4a：工程規模之金額對於發包遲延之情況有顯著負相關，亦指工程規模之金額越小，發包遲延情況上升。

5. 設計預算彈性：預算的金額常倚賴當年度的經費劃分、工程執行內容來衡量每個工程的執行預算，對於主辦機關而言，以最有效率的發包金額完成預算，不僅能使局處的經費使用有更多的發揮空間，亦可以達到更大的效益，因此在核定的預算範圍內，若可以適度的減少設計工程款，減少不必要的設計及開銷，將是主辦機關非常在意的要點。然而，對於廠商而言，合

理的投標底價是執行工程專案最重要的條件之一，若主辦機關在設計工程款上及壓縮的不合理、或者太低，即可能造成廠商在投標時，需以更低的價格去執行，不僅增加執行難度，亦可能造成投標意願下降。本研究以工程管考系統內設計工程款之欄位，與預算額度之欄位計算比值，以此代表主辦機關給定的執行工程款是否有足夠發揮的空間，另一方面即代表廠商在施作時能發揮的空間是否足夠(執行彈性大不大)。故本研究認為若比值越大，即代表給定設計的工程款項較多，後續發包的工程造價對於廠商而言有較大的執行彈性。

H5a：設計預算彈性對於發包遲延之情況有顯著負相關，亦指設計預算彈性越小，發包遲延情況上升

4.3 影響逾期情況之假說

1. 工程區域：同第二節所述，可能因為工程的地區偏遠、當地資源、環境潛在風險而影響廠商執行狀況，故本研究認為工程區域與逾期情況應呈現正向關係，越是偏遠的工程範圍應執行狀況越不樂觀。

H1b：工程區域對於逾期情況有顯著正相關，亦指越偏遠的工程區域，逾期狀況上升

2. 季節：同第二節所述，不同的季節對於施作廠商而言，可能也有承接專案的數量、施作人力、資源分配的影響，與後續施作廠商投案的規劃亦有所關聯，甚至是季節的風俗習慣。本研究採用專案開工時間作為探討標準，認為何時開工與逾期狀況可能存在著一定的相關性。

H2b：季節對於逾期情況有顯著相關，亦指不同的季節間存在著

工程執行狀況之差異

3. 工程屬性：同第二節所述，不同的工程類型與屬性，存在著不一樣的工程內容、設計與要求，直觀來說不同的工程類型其所需的資源、人力、風險、難易度有一定的差異，故本研究認為不同的工程屬性應與逾期情況有一定的相關性。

H3b：工程屬性對於逾期情況有顯著相關，亦指不同的工程屬性間存在著工程執行狀況之差異

4. 工程規模之金額：工程規模的大小潛在著非常多的風險與不確定性，包含人力資源的分配、材料的足夠與否、工程安全的風險高低、財務資金的穩定，甚至是環境等非人為因素，皆是因規模大小而影響，越大的工程規模所必須承擔的施工風險越大，無論是人力、材料的資料供應是否穩定，或是進度、環安衛、財務的管理，其中的一個環節若是沒有控管得宜，即有可能影響整個工程的執行狀況，本研究工程規模以金額衡量，故本研究認為工程規模中金額對於逾期情況有顯著正相關。

H4b：工程規模之金額對於工逾期狀況有顯著正相關，亦指工程規模之金額越大，工程逾期狀況上升

5. 設計預算彈性：同第二節所述，對於廠商而言，合理的工程執行金額非常重要，若主辦機關在設計工程款上及壓縮的不合理、或者太低，即可能會增加廠商執行難度，因而造成工程延宕，增加工程執行上的風險，亦可能影響營運上的困難，廠商在施作時能發揮的空間是否足夠(執行彈性大不大)十分重要。故本研究認為設計預算彈性的比值，影響後續的工程造價，以及直接影響廠商執行上的風險高低。

H5b：設計預算彈性對於逾期情況有顯著負相關，亦指彈性越小，工程逾期狀況上升

6. 汛期：對於水土保持局主要負責的案件類型，多半與河川、山坡地、農地、道路等與環境息息相關，尤其水與土的成份更是重要，因此降雨所帶來的影響更是執行人員與廠商擔心的一大隱憂，降雨量的多寡固然重要，但降雨天數的多寡相較於降雨量存在著更大的逾期風險。因此本研究將多數關心的變數“降雨量”改為“汛期”，認為汛期應與逾期狀況有非常大的正向相關，施工期間包含汛期的天數越多，潛在的風險越大。

H6b：汛期對於逾期狀況有顯著正相關，亦指汛期佔施工日期的比例越高，工程逾期狀況上升

4.4 小結

基於本研究於此章節提出之兩個模型假說，模型一針對發包遲延之情況提出，模型二針對逾期之情況提出，並詳述(1)假說如何建立、(2)假說內容為何及(3)假說與模型之相關性為何，後續第五章探索性實證模型研究設計將針對此兩個模型之假說進行變數設計和模型設置，以達相互呼應。

第五章 探索性實證模型研究設計

5.1 實證研究概念

本研究使用計量經濟學實證分析，計量經濟學是一整套結合經濟理論與数理統計理論的方法(如圖 5-1 所示)，並用定量的方法研究經濟活動規律及其應用。其中「計量」就是：以統計方法做定量研究。本研究之回歸模型以複回歸分析（Multiple Regression）方法為出發點，實證分析即是利用資料來檢定理論或是估計某個關係。

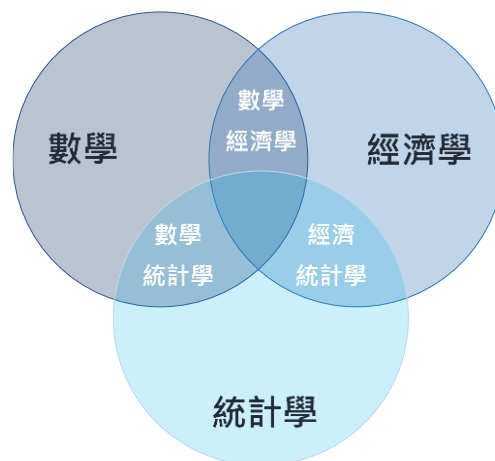


圖 5-1 計量經濟學的基礎理論

使用計量研究通常有兩個目的：

- 1.用於理論檢驗：估計經濟的關係、檢定經濟理論，以及評估與執行政府與企業之政策。
- 2.用於預測應用：經濟學家最一般的應用是預測一些重要的總體變數，諸如利率、通貨膨脹，以及國內生產毛額(GDP)。(Wooldridge, 2017)

依蒐集的數據型態之不同可分為兩種資料結構：(1)橫斷面資料

Cross-sectional Data 與(2)縱斷面資料 Time-series Data。

設置計量模型之流程如下，如 錯誤! 找不到參照來源。所示：

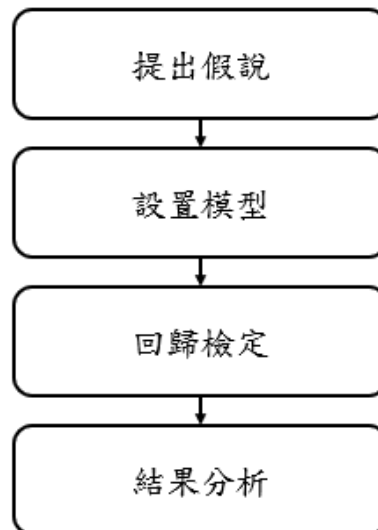


圖 5-2 計量模型設置流程圖

5.2 實證模型相關檢定

本研究模型使用普通最小平方法(Ordinary Least Squares, OLS)多元線性回歸分析法來檢驗假設。其估計方法是選取參數，使樣本誤差項的平方和極小化。根據高斯馬可夫定理(Gauss-Markov theorem)，若殘差符合分布為 i.i.d.的假設，則 OLS 所估計出來的係數，是「最佳線性不偏估計量」(Best Linear Unbiased Estimator, BLUE)，具有不偏、有效的性質(Wooldridge, 2017)。

本研究使用 OLS 進行參數估計，並且進行以下檢定，以確保模型參數之必要性及正確性。

1. 相關係數矩陣檢定

以相關係數矩陣檢定來看變數之間有沒有多重共線性，檢查模型中有無太過相似的因子。

2. 異質性檢定 Breusch-Pagan-Godfrey 檢定

判斷模型自變數之未觀測誤差 u 是否符合同質性假設，若為異質性，則需再作 Hetero Robust 檢定，做異質性之調整。

3. White 檢定 (White, 1980)

用以檢驗模型是否有因異質變異而造成模型估計的誤差。

4. Hetero Robust test：經過異質性調整後，使用調整過後之模型運算進行參數(自變數)和(1)順利發包與否及(2)工程執行狀況兩個應變數的相關性之判斷

5. 聯合檢定

是以 F 統計量之 F 分配來檢定條件變異是否有不對稱的現象。

5.3 變數之選定

5.3.1 應變數之選定

基於文獻回顧之專家訪談，本研究設定之探討主題為(1)發包遲延及(2)逾期情況之因素探討，因此在應變數的設定上即圍繞在此兩個部份。

本研究以“發包遲延”(Bid_delay)作為第一個模型之應變數，變

數衡量上則以工程管考系統內之資料欄位“核定日期”與“決標日期”之差先進行日期差異計算，後續以發包花費天數“60”天為基礎，判定此發包遲延與否，以虛擬變數(dummy variable)進行二元分類，0 為天數小於 60 天，1 則為天數大於 60 天。期中採用之發包花費天數基礎為 90 天，其依據為過去水土保持局累積發包的經驗數字，而近幾年已將數字調降為 60 天，並且參照期中審查意見更改為 60 天，為本次判定條件之依據。而採用二元分類之虛擬變數，未直接採用發包天數原因為：此表示方法有百分比的意涵存在，並且更能表達發包遲延情況之精神(不順利 1、有順利 0)，而非以發包延誤時間(天數)，作最後結果的解釋。本研究於此變數討論，認為在發包遲延的計算上亦有修正的空間，核定日期到決標日期中間包含了專案計畫之準備期等行政程序，較複雜之專案準備期越長，在應變數計算的數值上佔的比例越高，有偏差的可能。因此本研究認為若有開標與決標等較細節之資料作為衡量的依據，回歸結果將越準確。

第二個模型之應變數則以“逾期情況”(Project_Delay)代表，衡量上以“開工日期”加上“規定施工期限”作為“預期完工日期”，並與竣工日期之差進行日期差異計算，若天數為負值或 0，或正值並且與合約工作日之比例小於 10%，則定義為 0；若為正值

並且與合約工作日之比例大於 10%，則代表此工程有逾期之情況，以 1 表示，此表示法亦有百分比含意，可代表自變數與應變數間逾期情況之程度。本研究於此變數討論，認為此變數之計算亦有修正的空間，疑於資料庫內資料欄位，在預期工期計算上未把合理的展延工期納入，因此本研究認為若將合理展延天數納入分析當中，其結果才能更為精準的代表工程實際逾期之情況。

5.3.2 自變數之選定

本研究之自變數選定，基於 4.2、4.3 內容提到之所有假說設定變數，此部份敘述自變數之衡量標準。

1. 工程區域(Rural)

工程區域於本研究之期中報告以較簡略之方法先行衡量，透過工程管考系統中“鄉鎮”之欄位進行虛擬變數之分類，以鄉鎮是否為原住民鄉為依據，以 0 代表非原住民鄉，1 代表原住民鄉，以下為原住民鄉之選項：南澳鄉、大同鄉、烏來區、復興區、尖石鄉、五峰鄉。本研究於期末更改衡量方式，以內政部公告之“山地、平地原住民及離島等偏遠地區一覽表”之偏遠地區為依據，因部份縣別之山地原住民鄉並非歸類於偏遠地區，因此後續改以偏遠地區作為衡量的標準。若為當地縣別之偏遠地區，則以 1 代表；反之以 0 代表，一覽表 5-1 如下所示：

表 5-1 山地原住民及偏遠地區一覽表

縣別	山地原住民鄉	偏遠地區
台北縣	烏來鄉	石碇鄉、坪林鄉、平溪鄉、雙溪鄉、烏來鄉
桃園縣	復興鄉	復興鄉
新竹縣	五峰鄉、尖石鄉	五峰鄉、尖石鄉
苗栗縣	泰安鄉	泰安鄉、南庄鄉、獅潭鄉
台中縣	和平鄉	和平鄉
南投縣	仁愛鄉、信義鄉	中寮鄉、仁愛鄉、信義鄉
嘉義縣	阿里山鄉	番路鄉、大埔鄉、阿里山鄉
台南縣		楠西鄉、南化鄉、左鎮鄉、龍崎鄉
高雄縣	那瑪夏鄉、桃源鄉、茂林鄉	田寮鄉、六龜鄉、甲仙鄉、那瑪夏鄉、桃源鄉、茂林鄉
屏東縣	三地門鄉、霧台鄉、瑪家鄉、泰武鄉、來義鄉、春日鄉、獅子鄉、牡丹鄉	滿州鄉、泰武鄉、春日鄉、獅子鄉、牡丹鄉
宜蘭縣	大同鄉、南澳鄉	大同鄉、南澳鄉
花蓮縣	秀林鄉、萬榮鄉、卓溪鄉	鳳林鄉、壽豐鄉、光復鄉、瑞穗鄉、富里鄉、秀林鄉、卓溪鄉
台東縣	海端鄉、延平鄉、金峰鄉、達仁鄉、蘭嶼鄉	海端鄉、延平鄉、金峰鄉、達仁鄉、蘭嶼鄉、鹿野鄉、卑南鄉、大武鄉、東河鄉、長濱鄉
澎湖縣		湖西鄉、白沙鄉、西嶼鄉、望安鄉、七美鄉
金門縣		金城鎮、金寧鄉、金沙鎮、烈嶼鄉、烏坵
連江縣		北竿鄉、莒光鄉、東引鄉

* 偏遠地區定義（內政部）：人口密度低於全國平均人口密度五分之一之鄉鎮市；或距離直轄市、縣市政府所在地七點五公里以上之離島，計六十五鄉鎮。

* 此表參照行政院衛生署照護處界定山地原住民鄉、離島鄉、平地原住民鄉及偏遠地區之標準辦理。

2. 季節(Season)

季節因應變數之不同，衡量方式有所不同。模型一之應變數以發包

遲延情況為核心，討論後以工程管考系統中各專案“核定日期”之

欄位進行四季之分類；模型二之應變數以工程逾期情況為核心，討論後以工程管考系統中各專案“開工日期”之欄位進行四季之分類，並皆以第三季作為比較基礎進行比較。

3. 工程屬性(Attr₁)

工程屬性依據本研究範圍設計，以工程管考系統內大分類方向：流域綜合治理計畫、重劃區外緊急農路設施改善計畫、農村再生社區發展及環境改善計畫及整體性治山防災計畫作為本研究歸類方式，共四個類別，本研究討論若以計畫內的細部分類來做為變數的設計，大計畫內的子計畫亦有些微的差異，本研究礙於資料欄位的處理，本次以最明顯的四大類別進行討論，並以整體性治山防災為比較基礎。

本研究針對四個類別進行定義概述：

流域綜合治理與整體性治山防災計畫-為妥善經營與管理各種不同功能山坡地、促進國土資源永續利用發展、調節集水區產砂量及增進集水區涵養水源能力，並落實黃金十年願景五永續環境、國土空間發展策略計畫及水庫集水區保育綱要中有關推動整體性治山防災計畫政策需要，爰擬具相關計畫，其中山坡地保育條例之山坡地範圍依現行各計畫性質分為三個區塊，包含整體性治山防災計畫、氣候變遷下大規模崩塌防減災計畫、流域綜合治理計畫。而為考量天然環境之特性及地方政府資源平均分配之目標，當中保育治理工作包含土砂災害防治、水庫集水區保育及國有非公有山坡地水土保持理與維護，又流域綜合治理與治山防災所定義的治理區域不同，前者為高潛勢易淹水地區，後者為其他地區，及預算經費來源不同，因此本研究將其二歸納為不同屬性。

重劃區外緊急農路設施改善計畫-其計畫擬解決之問題在於農路邊坡崩塌地處理、路基穩定、排水及路面改善，有效維持農路之暢通 和降低損壞之風險，改善農路上邊坡土體穩定及危險路段，確保行車平安。並針對農路上、下邊坡辦理緩衝綠帶建置、植生復育等措施，以營造優質綠色環境，增加節能減碳效

益。

農村再生基金計畫-農村基礎生產條件與生活機能改善-其計畫目標在於為促進農村永續發展及農村活化再生，改善基礎生產條件，維護農村生態及文化，提升生活品質，建設富麗新農村，辦理農村文化維護 20 個社區，生活機能改善 180 個社區，農村基礎生產條件改善 180 個社區，彎腰掃社區並耕種希望 1,000 個社區，僱工購料 50 個社區，社區生態保育 5 個社區等。

4. 工程規模之金額(Price)

依據工程管考系統各專案資料之“發包工程款”欄位，作為實際工程核定之規模，模型一、二皆有此自變數。

5. 汛期(Rainday_percent)

汛期為模型二逾期狀況之自變數，依據工程管考系統各專案之“開工日期”及“竣工日期”欄位進行汛期天數計算，本研究設定之汛期期間為水利署第十河川局所公開之資訊，“汛期為每年五月初至十一月底，在這七個月的期間內，是台灣雨量最豐沛的梅雨及颱風季節；梅雨主要來自大陸長江流域的雲雨系，而颱風則是太平洋低氣壓所形成，這兩者皆會帶來連續性豪雨，也是最易發生水患的時刻。”然而因各區的氣候狀況不全然相同，因此本研究就以經驗值之資料作為基準，並以汛期天數與工程實際天數計算出汛期佔工程實際天數之百分比(園汛期天數/工程實際天數)，以此作為呈現變數進行分析。

6. 設計預算彈性(BD_flex)

設計預算彈性意指原本預定此工程之預算，與實際設計工程款之比例，依據工程管考系統“設計工程款”及“預算額度”欄位進行比例計算(設計工程款/預算額度)，以此作為設計預算彈性之分析數據。

5.3.3 控制變數之選定

控制變數之定義為：此變數與應變數有直觀影響，並且與自變數有關聯性存在，本研究除假說設計之變數位，有諸多其他關係會影響到應變數與自變數之顯著性，礙於資料不足及資料量化不易，因此以控制變數之形式設置模型，以確保其他關係不影響研究主軸，增加模型分析結果之準確性。

1. 年度(Year)

本研究認為，年度作為很重要的控制變數，基於本研究之假說，本研究亦有許多沒有探討的因素與關係，機關內未探討的包含年與年間的預算差異、機關內部異動；自然條件未探討的包含自然環境的極端事件(包含天災、氣候、地震、颱風等)、水文水理變化等；人文條件未探討的包含景氣狀況，直接影響人力、資源供應等，上述之未探討因素皆會因年度之特性不同而所有改變，因此本研究認為年度為必要控制之變數。此資料以工程管考系統中“年度”之欄位

進行分析，模型一、模型二皆以中間年度 105 年為基礎。

2. 地區(Area_)

地區：不同的地區存在著不一樣的內外文化、習慣、環境，內文化為局處的處事文化，在一定的規則下，不同的發包、執行習慣勢必會影響整個專案過程和結果。外文化為區域文化，地域的經濟狀況、發展方向、文化民情、自然環境狀況種類分佈，會導致不同的執行方針，因此本研究認為在進行全體資料回歸分析時，需將地區變數控制，才達到其餘自變數比較之標準。有別於期中報告僅採用北區案子，本次將全台 103-108 年度案子皆納入分析資料中，並設計以“中部區域”為比較基礎。

依照國家發展委員會所出版之“都市及區域發展統計彙編”，將台灣依照十二項目，即人口、規劃及建設、公用事業、交通與觀光、公共衛生、住宅、環境品質、大眾秩序與安全、教育文化、地方財政、家庭收支與設備及就業等進行分區，而離島除澎湖縣外，因各案少，因此不納入分析當中。區域之依序為：

北部區域(包括臺北市、新北市、基隆市、新竹市、桃園市、新竹縣及宜蘭縣)

中部區域(包括臺中市、苗栗縣、彰化縣、南投縣及雲林縣)

南部區域(包括高雄市、臺南市、嘉義市、嘉義縣、屏東縣及澎湖縣)

東部區域(包括花蓮縣及臺東縣)

5.4 模型介紹

本研究於 5.3 節已針對模型中的各參數進行介紹，同時也針對其假說之概念加以說明。在 5.4 節，會呈現欲探討之兩個計量模型及

圖示與列表假說概念。

5.4.1 模型一概念圖及大數據計量模型

模型一主要探討發包遲延之因素，其中包含五個自變數、二個控制變數，分別是工程區域、季節、工程屬性、工程規模之金額、設計預期彈性，與年度、地區，並且針對收集的資料分類進行各區回歸模型以及全區回歸模型。本研究提出以下對模型影響之假說及回歸模型方程式，示意如下表 5-2 所示。

表 5-2 模型一假說表

假說	對應變數	假說內容
H1a	工程區域 (rural)	工程區域與發包遲延有顯著正相關
H2a	季節 (season)	季節與發包遲延有顯著相關
H3a	工程屬性 (attr_)	工程屬性與發包遲延有顯著相關
H4a	工程規模之金額(price)	工程規模之金額與發包遲延有顯著負相關
H5a	設計預算彈性 (execute_flex)	設計預算彈性與發包遲延有顯著負相關
	控制變數	年度(Year)、地區(Area_)

$$\begin{aligned}
 Bid_delay = & \beta_0 + \delta_1 Rural + \delta_2 Season1 + \delta_3 Season2 + \delta_4 Season4 + \delta_5 Attr_flood \\
 & + \delta_6 Attr_roadimprove + \delta_7 Attr_villagerenew + \delta_8 year103 + \delta_9 year104 \\
 & + \delta_{10} year106 + \delta_{11} year107 + \delta_{12} year108 + \delta_{13} Area_north \\
 & + \delta_{14} Area_south + \delta_{15} Area_east + \beta_{16} Price + \beta_{17} BD_flex + u
 \end{aligned}$$

5.4.2 模型二概念圖及大數據計量模型

模型二主要探討逾期之狀況，其中包含六個自變數、二個控制變數，分別是工程區域、季節、工程屬性、工程規模之金額、汛期、設計預算彈性及年度、地區，本研究提出以下對模型影響之假說及回歸模型方程式，示意如下表 5-3 所示。

表 5-3 模型二假說表

假說	對應變數	假說內容
H1b	工程區域(rural)	工程區域與逾期狀況有顯著正相關
H2b	季節(season)	季節與逾期狀況有顯著相關
H3b	工程屬性(attr_)	工程屬性與逾期狀況有顯著相關
H4-1b	工程規模之金額 (price)	工程規模之金額與逾期狀況有顯著正相關
H5b	設計預算彈性 (execute_flex)	設計預算彈性與逾期狀況有顯著負相關
H6b	汛期 (Rainday_Percent)	汛期與逾期狀況有顯著正相關
	控制變數	年度(Year)、地區(Area_)

Project_delay

$$\begin{aligned}
&= \beta_0 + \delta_1 Rural + \delta_2 Season1 + \delta_3 Season2 + \delta_4 Season4 + \delta_5 Attr_flood \\
&+ \delta_6 Attr_roadimprove + \delta_7 Attr_villagerenew + \delta_8 year103 + \delta_9 year104 \\
&+ \delta_{10} year106 + \delta_{11} year107 + \delta_{12} year108 + \beta_{13} Scale + \beta_{13} Price \\
&+ \beta_{14} BD_flex + \beta_{15} Rainday_percent + \delta_{16} Area_north + \delta_{17} Area_south \\
&+ \delta_{18} Area_east + u
\end{aligned}$$

等式之左端為應變數，等式右側為自變數和控制變數， u 為假設相關性為零的隨機誤差， β 為連續變數之係數，其中 β_0 為各回歸公式之截距， δ 為虛擬變數(Dummy Variable)之係數。

5.5 資料蒐集

本研究採用水土保持局工程管考系統於各工程計畫之紀錄資訊，年度區間為 103-108 年度，於模型一剔除缺漏資料共得 10832 筆資料，其中北部區域共 1035 筆、中部區域共 5103 筆、南部區域共 3415 筆、東部區域共 1279 筆；模型二剔除缺漏資料共得 10953 筆資料，其中北部區域共 1065 筆、中部區域 4138 筆、南部區域 4443 筆、東部區域 1307 筆，其蒐集資料之呈現如圖 5-3 所示

水土保持工程巨量資料探勘與計量實證分析研究

流域綜合治理計畫103-108.xlsx - Excel																			
流域綜合治理計畫103-108																			
序	工作項目	工程名稱	工程內容	預算總額(千元)	執行機關	執行單位	承辦人	縣市	鄉鎮	村莊	計畫名稱	核定日期	核定完工日期	開始動工日期	實際完成日期	實際完成日期	實際完成日期	實施方式	備註
1	土壤工程	103-207P-Q1-1001	上游崩塌地防沖擊工程	400	臺北分署	林口區	林口區	林口區	林口區	林口區	林口區	103/05/04	103/09/27	103/09/27	103/09/27	103/09/27	103/09/27	公開招標	防沖擊工程
2	土壤工程	103-207P-Q1-1002	崩塌地防沖擊工程	400	臺北分署	林口區	林口區	林口區	林口區	林口區	林口區	103/05/04	103/09/27	103/09/27	103/09/27	103/09/27	103/09/27	公開招標	防沖擊工程
3	土壤工程	103-207P-Q1-1003	崩塌地防沖擊工程	400	臺北分署	林口區	林口區	林口區	林口區	林口區	林口區	103/05/04	103/09/27	103/09/27	103/09/27	103/09/27	103/09/27	公開招標	防沖擊工程
4	土壤工程	103-207P-Q1-1004	崩塌地防沖擊工程	400	臺北分署	林口區	林口區	林口區	林口區	林口區	林口區	103/05/04	103/09/27	103/09/27	103/09/27	103/09/27	103/09/27	公開招標	防沖擊工程
5	土壤工程	103-207P-Q1-1005	崩塌地防沖擊工程	400	臺北分署	林口區	林口區	林口區	林口區	林口區	林口區	103/05/04	103/09/27	103/09/27	103/09/27	103/09/27	103/09/27	公開招標	防沖擊工程
6	土壤工程	103-207P-Q1-1006	崩塌地防沖擊工程	400	臺北分署	林口區	林口區	林口區	林口區	林口區	林口區	103/05/04	103/09/27	103/09/27	103/09/27	103/09/27	103/09/27	公開招標	防沖擊工程
7	土壤工程	103-207P-Q1-1007	崩塌地防沖擊工程	400	臺北分署	林口區	林口區	林口區	林口區	林口區	林口區	103/05/04	103/09/27	103/09/27	103/09/27	103/09/27	103/09/27	公開招標	防沖擊工程
8	土壤工程	103-207P-Q1-1008	崩塌地防沖擊工程	400	臺北分署	林口區	林口區	林口區	林口區	林口區	林口區	103/05/04	103/09/27	103/09/27	103/09/27	103/09/27	103/09/27	公開招標	防沖擊工程
9	土壤工程	103-207P-Q1-1009	崩塌地防沖擊工程	400	臺北分署	林口區	林口區	林口區	林口區	林口區	林口區	103/05/04	103/09/27	103/09/27	103/09/27	103/09/27	103/09/27	公開招標	防沖擊工程
10	土壤工程	103-207P-Q1-1010	崩塌地防沖擊工程	400	臺北分署	林口區	林口區	林口區	林口區	林口區	林口區	103/05/04	103/09/27	103/09/27	103/09/27	103/09/27	103/09/27	公開招標	防沖擊工程
11	土壤工程	103-207P-Q1-1011	崩塌地防沖擊工程	400	臺北分署	林口區	林口區	林口區	林口區	林口區	林口區	103/05/04	103/09/27	103/09/27	103/09/27	103/09/27	103/09/27	公開招標	防沖擊工程
12	土壤工程	103-207P-Q1-1012	崩塌地防沖擊工程	400	臺北分署	林口區	林口區	林口區	林口區	林口區	林口區	103/05/04	103/09/27	103/09/27	103/09/27	103/09/27	103/09/27	公開招標	防沖擊工程
13	土壤工程	103-207P-Q1-1013	崩塌地防沖擊工程	400	臺北分署	林口區	林口區	林口區	林口區	林口區	林口區	103/05/04	103/09/27	103/09/27	103/09/27	103/09/27	103/09/27	公開招標	防沖擊工程
14	土壤工程	103-207P-Q1-1014	崩塌地防沖擊工程	400	臺北分署	林口區	林口區	林口區	林口區	林口區	林口區	103/05/04	103/09/27	103/09/27	103/09/27	103/09/27	103/09/27	公開招標	防沖擊工程
15	土壤工程	103-207P-Q1-1015	崩塌地防沖擊工程	400	臺北分署	林口區	林口區	林口區	林口區	林口區	林口區	103/05/04	103/09/27	103/09/27	103/09/27	103/09/27	103/09/27	公開招標	防沖擊工程
16	土壤工程	103-207P-Q1-1016	崩塌地防沖擊工程	400	臺北分署	林口區	林口區	林口區	林口區	林口區	林口區	103/05/04	103/09/27	103/09/27	103/09/27	103/09/27	103/09/27	公開招標	防沖擊工程
17	土壤工程	103-207P-Q1-1017	崩塌地防沖擊工程	400	臺北分署	林口區	林口區	林口區	林口區	林口區	林口區	103/05/04	103/09/27	103/09/27	103/09/27	103/09/27	103/09/27	公開招標	防沖擊工程
18	土壤工程	103-207P-Q1-1018	崩塌地防沖擊工程	400	臺北分署	林口區	林口區	林口區	林口區	林口區	林口區	103/05/04	103/09/27	103/09/27	103/09/27	103/09/27	103/09/27	公開招標	防沖擊工程

圖 5-3 水土保持局工程管考系統資料

第六章 大數據分析結果與討論

本章節乃計量研究結果分析，分為四個部分，敘述如下。

6.1 針對整理之欄位資料，進行敘述性統計。

6.2 介紹相關係數矩陣，內容包含各變數之間之相關性分析。

6.3 進行回歸模型之計量相關檢定結果。

6.4 針對回歸分析結果進行討論，整理實證分析之數據所得之結論及其貢獻。

6.1 敘述性統計

本研究首先針對幾個重要變數之各區及全區資料進行基本統計分析，包含應變數發包遲延情況、逾期狀況，自變數工程區域、季節(核定季節及開工季節)、工程規模之金額、設計預算彈性與汛期。如下表 6-1-6-6 所示：

表 6-1 應變數發包遲延(60 天)敘述性統計結果

應變數	平均數(天)	標準差	次數(1)	次數(0)
北區 Bid_Delay	100.13	56.50	758	277
中區 Bid_Delay	105.64	50.78	4282	821
南區 Bid_Delay	104.25	55.98	2799	616
東區 Bid_Delay	102.50	47.16	1088	191
全區 Bid_Delay	105.25	52.67	8927	1905

*Bid_Delay：1 為發包遲延、0 為無發包遲延

第一部份由表 6-1 所得知，各區在發包順利與否的狀況，平均花費天數約為 100-105 天，與原先定義的超過 60 天即算發包不順利相比之下，各區的發包狀況不盡理想，若假設此資料為常態分佈，在一個標準差內，各區的平均發包天數有 68%機率落在 50-155，亦有涵蓋到未流標的範圍，但範圍不大，而從判斷的次數上亦可了解有無流標

的次數差異十分明顯。

表 6-2 應變數發包遲延(90 天)敘述性統計結果

應變數	次數(1)	次數(0)
北區 Bid_Delay	559	476
中區 Bid_Delay	3205	1898
南區 Bid_Delay	2067	1348
東區 Bid_Delay	824	455
全區 Bid_Delay	6655	4177

於表 6-2 將發包遲延標準改為 90 天，統計各區及全區發包遲延之次數，本研究考量採用資料之年份區間為 103-108 年度，與水保局執行單位討論，60 天的發包遲延標準從 108 年開始修正推行，應採用 90 天作為判斷依據，而若以 60 天為標準，可於分析結果辨斷標準轉變與執行策略的執行是否有明顯的差異，並且於期中審查建議以 60 天作為標準，故本研究保持原先設定，如圖 6-1、6-2 所示。

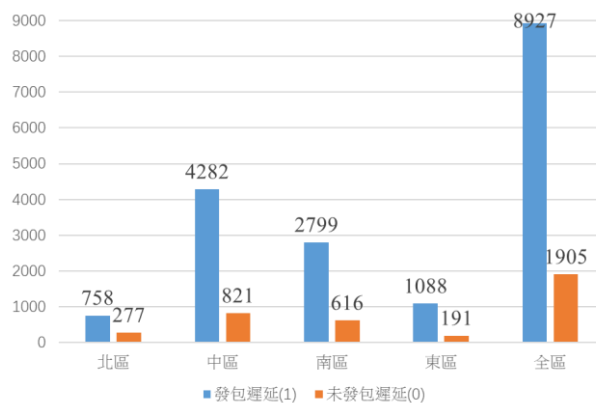


圖 6-1 全區與各區發包遲延次數(60 天)

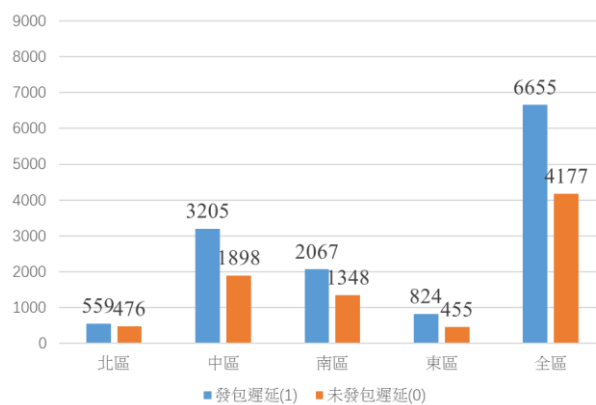


圖 6-2 全區與各區發遲延次數(90 天)

表 6-3 應變數工程逾期敘述性統計結果

應變數	平均數(天)	標準差	次數(1)	次數(0)
北區 Project_Delay	61.84	48.88	559	506
中區 Project_Delay	50.92	47.80	1536	2602
南區 Project_Delay	50.90	45.41	1531	2912
東區 Project_Delay	44.03	36.96	455	852
全區 Project_Delay	51.64	46.18	4081	6872

*Delay：1 為逾期、0 為無逾期

第二部份本研究將各區逾期天數大於工期 10%的資料萃取出計算平均數及標準差，由表 6-1 中得知，除北區逾期天數約為 60 天外，其餘皆落在 40-50 間，而北區也是唯一逾期案件次數大於非逾期的區域，同樣在假設資料為常態分佈的條件下，工程執行狀況的逾期天數有 68%的落在 3-108 天，如表 6-3 及圖 6-3 所示。

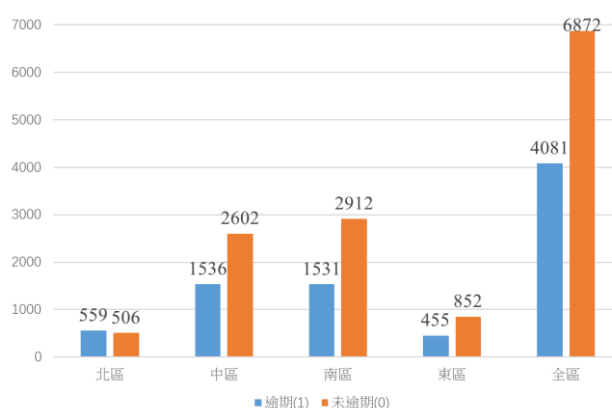


圖 6-3 全區與各區工程逾期次數

表 6-4 自變數工程區域敘述性統計結果

工程區域-模型一	偏遠地區(1)	非偏遠地區(0)
北區	347	688
中區	1006	4097
南區	969	2446
東區	878	401
全區	3200	7632
工程區域-模型二	偏遠地區(1)	非偏遠地區(0)
北區	344	721
中區	832	3306

南區	1159	3284
東區	900	407
全區	3235	7718

自表 6-4 中得知，在模型一的資料中，北、中、南區偏遠地區的案件與非偏遠的比例約略為 1:2，而東區則相反，約略為 2:1；在模型二的資料中，一樣偏遠地區的案件除東區外，皆是非偏遠數量大於偏遠，因此可了解此在本研究分析的所有資料中，案件偏遠地區的比重在各區中大約狀況。

表 6-5 自變數季節敘述性統計結果

核定季節-模型一	第一季	第二季	第三季	第四季
北區	369	252	211	203
中區	1551	1242	1496	814
南區	952	808	998	657
東區	449	297	244	289
全區	3321	2599	2949	1963
開工季節-模型二				
北區	261	259	306	239
中區	719	932	1032	1455
南區	880	863	1135	1565
東區	266	439	315	287
全區	2126	2493	2788	3546

由表 6-5 中得知，以在模型一的資料中，各區在第一季所核定的案件最多，其次第三季，再者第二季，累計各差約 400 件，而第四季的核定案件最少；而模型二的資料中，以第四季開工的案件最多，其次第三季、再者第二季，而第一季的開工案件最少，又各區各季的開工案件些微不同。

表 6-6 自變數工程規模之金額敘述性統計結果

工程金額-模型一	平均數(千)	標準差	中位數
北區	4842	3871.03	3780
中區	2725	2407.64	2220
南區	2578	2440.15	1938
東區	3110	2872.69	2480
全區	2926	2724.65	2326
工程金額-模型二	平均數	標準差	中位數
北區	4825	3829.69	3780
中區	2686	2627.25	2210
南區	2657	2602.82	2038
東區	3111	2900.84	2480
全區	2933	2861.71	2330

自表 6-6 中得知，模型一與模型二的資料差異不大，北區的工程金額平均約略為 484 萬，為各區最高，假設常態分佈下在兩個標準差中，95%的案件金額區間落在 1258 萬內，而中區、南區、東區工程金額平均約略為 300 萬上下，假設常態分佈下在兩個標準差中，95%的案件金額區間落在 890 萬以下，可了解北區的案件相比較下金額較大。

表 6-7 自變數執行彈性敘述性統計結果

設計預算彈性-模型一	平均數	標準差	中位數
北區	0.84	0.12	0.89
中區	0.85	0.11	0.89
南區	0.86	0.13	0.90
東區	0.85	0.10	0.89
全區	0.85	0.12	0.89
設計預算彈性-模型二	平均數	標準差	中位數
北區	0.84	0.12	0.89
中區	0.85	0.11	0.89
南區	0.86	0.19	0.90

東區	0.85	0.09	0.89
全區	0.85	0.15	0.89

自表 6-7 中得知模型一與模型二的資料，在執行彈性的平均數上均接近 0.84-0.86，假設常態分佈下在兩個標準差中，95%的案件執行彈性約略為 0.6-1，亦表示在工程設計完後，給予工程費用的折數為一個有固定的經驗值，劇烈跳動的幅度不大，較無給予廠商不夠的施作彈性的疑慮。

表 6-8 自變數汛期與工期敘述性統計結果

汛期-模型二	平均數	標準差	中位數
北區	0.62	0.34	0.68
中區	0.56	0.39	0.62
南區	0.54	0.41	0.59
東區	0.64	0.38	0.78
全區	0.47	0.39	0.46
工程工期(非變數)	平均數	標準差	中位數
北區	121.76	49.20	120
中區	106.45	50.11	100
南區	91.61	48.73	85
東區	95.28	46.32	90
全區	100.58	49.96	90

自表 6-8 中得知，在模型二的汛期的資料中，東區與北區所遇到的汛期佔工期的比例約為 0.62-0.64，中區跟南區約為 0.54-0.56。而在非變數的工程工期資料中，北區的平均工期天數較長，約為 122 天，而南區的天數最短，約為 95 天，其他約為 100 天上下，假設常態分佈下在兩個標準差內，95%的案件工期約在 219 天以內，案件工期大於一年的數量非常少量。

6.2 相關係數矩陣

各個變數之相關性由相關係數矩陣來表示，相關係數介於+1 到-1 之間，其係數意義可參照表 6-9。在回歸模型下，若相關係數達到 0.8 以上，及代表相關程度極高，要考慮可能有共線性問題，會使模型準確性受到影響，若遇到自變數有兩者皆不顯著，且彼此間之相關係數極高，則可利用聯合檢定(Joint hypotheses test)，來考量是否對應變數有聯合效果。

表 6-9 相關係數解釋參照表

相關係數	0.8 以上	0.8-0.6	0.4-0.6	0.2-0.4	0.2 以下	+	-
相關程度	極高	高	中	低	極低	正相關	負相關

表 6-10 模型一之變數相關係數矩陣(無控制變數)

	R	S1	S2	S4	AT_F	AT_R	AT_V	Price	E_F
R	1.00								
S1	-0.02	1.00							
S2	0.01	-0.37	1.00						
S4	0.05	-0.31	-0.26	1.00					
AT_F	0.15	-0.02	0.03	0.06	1.00				
AT_R	0.003	-0.14	0.08	-0.01	-0.22	1.00			
AT_V	-0.22	0.21	-0.03	-0.17	-0.30	-0.36	1.00		
Price	0.14	0.12	-0.02	0.01	0.18	-0.14	-0.11	1.00	
E_F	0.04	-0.09	0.03	0.03	-0.02	0.02	-0.03	0.13	1.00

表 6-11 模型二之變數相關係數矩陣(無控制變數)

	R	S1	S2	S4	AT_F	AT_R	AT_V	Price	Dur	E_F	R_P
R	1.00										
S1	0.02	1.00									
S2	0.06	-0.27	1.00								
S4	-0.05	-0.34	-0.38	1.00							
AT_F	0.15	0.08	0.08	-0.11	1.00						
AT_R	0.004	0.03	-0.09	0.07	-0.23	1.00					
AT_V	-0.22	-0.13	-0.05	0.009	-0.30	-0.36	1.00				
Price	0.13	0.009	0.07	-0.09	0.18	-0.14	-0.11	1.00			
Dur	0.09	0.04	0.12	-0.12	0.25	-0.17	-0.27	0.61	1.00		
E_F	0.02	0.007	-0.03	0.05	-0.02	0.01	-0.01	0.10	-0.004	1.00	
R_P	0.04	-0.24	0.57	-0.66	0.09	-0.07	0.03	0.07	0.09	-0.05	1.00

從表 6-10 得知，在全區資料的模型一中，自變數之間的相關係數普遍落在低與極低的相關程度，而表 6-11 得知，在全區資料的模型二中自變數之間的相關係數矩陣，變數間關係大部份處於低或極低的相關程度，而第二季與汛期屬於中相關程度，第四季與汛期屬於高相關程度。

6.3 回歸模型檢定

本節針對兩個模型做 Breusch-Pagan-Godfrey 檢定，檢定模型誤差項是否有異質變異，若有異質變異則需再作 Hetero Robust test，做異質調整。

若檢定出模型之誤差項有異質變異，則模型的估計式就會產生偏

誤。Breusch-Pagan-Godfrey 與 White 檢定的虛無假 H_0 設為同質變異，故如果 p 值小於 0.05，可以代表模型有顯著的異質變異。從表 6-13 可看出模型一及二的全區資料在 1%信心水準下皆拒絕 H_0 同質性的假設，因此本研究認為兩模型皆必須使用異質穩健標準誤 (Heteroskedasticity-consistent standard errors) 的設定重新回歸分析，以 Hetero Robust test 進行之後假說驗證之顯著性分析。

表 6-12 回歸模型之 Breusch-Pagan-Godfrey 與 White 檢定表

BPG Test	F-statistic	Probability
模型一(Bid_Delay)	59.377	0.0000
模型二(Project_Delay)	38.421	0.0000
White Test	F-statistic	Probability
模型一(Bid_Delay)	23.120	0.0000
模型二(Project_Delay)	10.672	0.0000

6.4 回歸分析結果與討論

本研究上述針對了各變數的相關係數矩陣的探討，亦對誤差項進行異質性檢定，確定需以 Hetero-Robust 的設定進行回歸模型分析。本研究最主要的還是針對各個自變數與應變數 “Bid_Delay” 與 “Project_Delay” 做探討，以了解到底什麼因素影響(1)發包遲延及(2)逾期狀況？

6.4.1 模型一全區回歸結果與假設驗證

表 6-13 模型一回歸結果整理

	Model 1 (Bid_Delay)	
	Coefficient (回歸係數)	P-Value
Rural	0.0136	(0.0958)
Season1	-0.1727	(0.0000)***
Season2	-0.0176	(0.0352)
Season4	-0.1721	(0.0000)***
Attr_flood	-0.0616	(0.0000)***
Attr_roadimprove	-2.88×10^{-5}	(0.9977)
Attr_villagerenew	0.1337	(0.0000)***
Price	9.57×10^{-6}	(0.0000)***
BD_flex	-0.1526	(0.0000)***
Year103	0.1024	(0.0000)***
Year104	0.1654	(0.0000)***
Year106	-0.0320	(0.0465)
Year107	0.2841	(0.0000)***
Year108	0.2045	(0.0000)***
Area_north	-0.0923	(0.0000)***
Area_south	-0.0264	(0.0000)***
Area_east	0.0378	(0.0000)***
Adjusted R-squared	0.1439	(0.0000)***

註： *P<0.01；**P<0.005；***P<0.001

表 6-14 統整了模型一全區分析結果，由表 6-14 可得 Model1 針對發包遲延的 Adjusted R-squared 為 14.39%，意指 Model1 的設計能藉由此份分析資料解釋本研究探討的現象 14.39%，此結果對於社會科學的回歸分析而言，解釋力尚可偏低，但在此模型變數中，可以發現一些影響因素。在 Model1 的各變數中，第一季、第四季、流域治理、

農村更新、工程規模之金額及設計預算彈性皆為顯著，而對照於本研究於先前所立定的模型一假說，季節變數於 H2a 得到證實 ($P(\text{Season1}) < 0.001$ 、 $P(\text{Season4}) < 0.001$)、工程屬性變數於 H3a 得到證實 ($P(\text{Attr_flood}) < 0.001$ 、 $P(\text{Attr_villagerenew}) < 0.001$)、工程規模之金額變數為顯著相關 ($P(\text{Price}) < 0.001$)、設計預算彈性變數於 H5a 得到證實 ($P(\text{BD_flex}) < 0.001$)，意指此些變數皆對於發包遲延有顯著相關，其餘變數顯著性低，較無關聯性，整理如表 6-15。

表 6- 14 Model1 假說驗證表

假說	假說內容	模型 假說	實證 假說	驗證 符合
H1a	工程區域 與發包遲延有顯著正相關	正相關	不顯著	X
H2a	季節 與發包遲延有顯著相關	顯著相關	顯著	O
H3a	工程屬性 與發包遲延有顯著相關	顯著相關	顯著	O
H4a	工程規模金額 與發包遲延有顯著負相關	負相關	顯著 正相關	X
H5a	設計預算彈性 與發包遲延有顯著負相關	負相關	顯著 負相關	O

再者針對顯著自變數的回歸係數進行討論。在固定其他變數情況下，對於季節變數而言，第一季的表現最佳、第四季表現次之，發包遲延的機率約小 17.2%，並且從結果指出，第三季表現與第二季沒有太大的差異，表現較差。在固定其他變數情況下，對於工程屬性而言，流域治理順利發包的狀況最佳，發包遲延的機率約小 6%，而農村更新的表現最差，發包遲延的機率約大 13.37%。在固定其他變

數情況下，對於工程規模之金額而言，係數為 9.57×10^{-6} ，若在平均數 290 萬、標準差 270 萬的條件下，每提升 270 萬的工程金額，發包遲延的機率約大 2.6%。在控制變數年度下，可發現各年度相較於比較基礎(105 年度)皆有所差異，而控制變數地區則顯示，北區相較於中區而言，發包遲延的機率最低(-9.2%)，南區次之(-2.6%)，而東區發包遲延的機率最高，約大 3.8%。

6.4.2 模型二全區回歸結果與假設驗證

表 6- 15 模型二回歸結果整理

	Model 2 (Project_Delay)	
	Coefficient (回歸係數)	P-Value
Rural	0.0115	(0.2795)
Season1	-0.0247	(0.1558)
Season2	0.0522	(0.0001)***
Season4	-0.1195	(0.0000)***
Attr_flood	-0.0140	(0.3031)
Attr_roadimprove	0.0632	(0.0000)***
Attr_villagerenew	0.2263	(0.0000)***
Price	1.98×10^{-5}	(0.0000)***
BD_flex	-0.0227	(0.3531)
Rainday_percent	-0.0918	(0.0000)***
Year103	-0.0145	(0.3285)
Year104	0.0068	(0.6378)
Year106	0.0011	(0.9460)
Year107	-0.0137	(0.4168)
Year108	-0.0482	(0.0026)**
Area_north	0.1341	(0.0000)***
Area_south	-0.0382	(0.0000)***

Area_east	-0.0213	(0.1738)
Adjusted R-squared	0.07	(0.0000)***

註： *P<0.01；**P<0.005；***P<0.001

表 6-16 統整了模型二全區之結果，由表 6-16 可得 Model2 針對工程逾期狀況的 Adjusted R-squared 為 7%，意指 Model2 的設計能藉由此份分析資料解釋本研究探討的現象約 7%，此結果的解釋程度稍微偏低，但亦能在模型中發現重要變數的關聯性。在 Model2 的各變數中，第二季、第四季、農路改善、農村更新、工程規模之金額、汛期皆為顯著，而對照於本研究於先前所立定的模型二假說，季節變數於 H2b 得到證實($P(\text{Season2}) < 0.001$ 、 $P(\text{Season4}) < 0.001$)、工程屬性於 H3b 得到證實($P(\text{Attr_villagerenew}) < 0.001$ 、 $P(\text{Attr_roadimprove}) < 0.001$)、工程規模之金額 H4b 得到證實($P(\text{Price}) < 0.001$)、汛期於 H6b 得到證實($P(\text{Rainday_percent}) < 0.001$)，意指此些變數皆對於工程逾期狀況有顯著相關，其餘變數顯著性低，較無關聯性，整理如表 6-17。

表 6- 16 Model2 假說驗證表

假說	假說內容	模型假說	實證假說	驗證符合
H1b	工程區域與逾期狀況有顯著正相關	正相關	不顯著	X
H2b	季節與逾期狀況有顯著相關	顯著相關	顯著	O
H3b	工程屬性與逾期狀況有顯著相關	顯著相關	顯著	O
H4b	工程規模之金額與逾期狀況有顯著正相關	正相關	顯著正相關	O
H5b	設計預算彈性與逾期狀況有顯著負相關	負相關	不顯著	X

H6b	汛期與逾期狀況有顯著正相關	正相關	負顯著	X
-----	---------------	-----	-----	---

再者顯著自變數的回歸係數進行討論。在固定其他變數情況下，對於季節變數而言，第四季的表现最佳，工程逾期的機率約小 12%，第二季的表现較差，工程逾期的機率約大 5.2%，而第一季的表现與第三季沒有太大的差異。在固定其他變數情況下，對於工程屬性而言，農村更新的表現最差，工程逾期的機率約大 22.6%，農路改善的表現次差，工程逾期的期率約大 6.32%，其於工程屬性差異不大。在固定其他變數情況下，對於工程規模之金額而言，其係數為 1.98×10^{-5} ，若在平均數 290 萬、標準差 280 萬的條件下，每提升 280 萬的工程金額，工程逾期的機率約大 5.54%。在固定其他變數情況下，對於汛期而言，若汛期佔施工期間的比例越高(若為 100%)，其工程逾期約下降 9.7%。在控制變數年度下，可發現各年度相較於比較基礎(105 年度)除 108 年外皆無差異，亦指工程逾期在各年度為普遍發生的現象，而控制變數地區則顯示，北區相較於中區而言，工程逾期的機率約大 13.41%，而南區相較於中區而言，工程逾期的機率約小 3.82%，表現最佳。

6.4.3 模型一各區回歸結果與假設驗證

表 6-17 模型一各區回歸結果整理(北區)

	Model 1 (Bid_Delay)	
	Coefficient (回歸係數)	P-Value
Rural	0.0283	(0.3155)
Season1	-0.1847	(0.0000)***
Season2	-0.0807	(0.0234)**
Season4	-0.3070	(0.0000)***
Attr_flood	-0.2038	(0.0000)***
Attr_roadimprove	0.0112	(0.7437)
Attr_villagerenew	0.2738	(0.0000)***
Price	2.65×10^{-6}	(0.4234)
BD_flex	-0.0101	(0.9193)
Year103	0.0795	(0.0497)**
Year104	0.0401	(0.3432)
Year106	-0.0563	(0.2421)
Year107	0.3934	(0.0000)***
Year108	0.3242	(0.0000)***
Adjusted R-squared	0.2452	(0.0000)***

註： *P<0.1；**P<0.05；***P<0.01

表 6-18 統整了模型一北區分析結果，由表 6-18 可得 Model1 針對流標情況的 Adjusted R-squared 為 24.52%，意指 Model1 的設計能藉由此份分析資料解釋本研究探討的現象 24.52%，此結果對於社會科學的回歸分析而言，解釋力尚可。在 Model1 的各變數中，第一季、第二季、第四季、流域治理、農村更新皆為顯著，其自變數顯著性結果與模型一全區雷同。而在固定其他變數的條件下，在季節變數中，第四季表現最佳，發包遲延的機率約小 30.7%，第一季表現次之，發

包遲延的機率約小 18.47%，第二季尚可，發包遲延的機率約小 8.07%，而第三季的表現最差。在工程屬性變數，流域治理的表現最佳，發包遲延的機率約小 20.38%，農村更新的表現最差，發包遲延的機率約大 27.38%。

表 6-18 模型一各區回歸結果整理(中區)

	Model 1 (Bid_Delay)	
	Coefficient (回歸係數)	P-Value
Rural	0.0292	(0.0127)
Season1	-0.1853	(0.0000)***
Season2	-0.0057	(0.6241)
Season4	-0.1806	(0.0000)***
Attr_flood	-0.0521	(0.0005)***
Attr_roadimprove	-0.0356	(0.0106)
Attr_villagerenew	0.1509	(0.0000)***
Price	1.08×10^{-5}	(0.0000)***
BD_flex	-0.1809	(0.0000)***
Year103	0.1720	(0.0000)***
Year104	0.2170	(0.0000)***
Year106	0.0029	(0.9022)
Year107	0.3590	(0.0000)***
Year108	0.2882	(0.0000)***
Adjusted R-squared	0.1862	(0.0000)***

註： *P<0.01；**P<0.005；***P<0.001

表 6-19 統整了模型一中區分析結果，由表 6-19 可得 Model1 針對發包遲延的 Adjusted R-squared 為 18.62%，意指 Model1 的設計能藉由此份分析資料解釋本研究探討的現象 18.62%，此結果對於社會科學的回歸分析而言，解釋力尚可。在 Model1 的各變數中，第一季、

第四季、流域治理、農村更新、工程規模之金額、設計預算彈性皆為顯著，其自變數顯著性結果與模型一全區雷同。而在固定其他變數的條件下，在季節變數中，第一季表現最佳，發包遲延的機率約小 18.53%，第四季表現次之，發包遲延的機率約小 18.06%，第二季與第三季的差異不大，表現較差。在工程屬性變數，流域治理的表現最佳，發包遲延的機率約可小 5.21%，農村更新的表現最差，發包遲延的機率約大 15.09%。對於工程規模之金額而言，其係數為 1.08×10^{-5} ，若的平均數 272 萬、標準差 240 萬的條件下，每提升 240 萬的工程金額，發包遲延的機率約大 2.6%。對於設計預算彈性而言，其彈性越大（若比值為 1），則發包遲延的機率約小 18.09%。

表 6-19 模型一各區回歸結果整理(南區)

	Model 1 (Bid_Delay)	
	Coefficient (回歸係數)	P-Value
Rural	-0.0030	(0.8411)
Season1	-0.1758	(0.0000)***
Season2	-0.0167	(0.2578)
Season4	-0.2011	(0.0000)***
Attr_flood	-0.0188	(0.4218)
Attr_roadimprove	0.0178	(0.3747)
Attr_villagerenew	0.0617	(0.0012)**
Price	1.99×10^{-5}	(0.0000)***
BD_flex	-0.1938	(0.0002)***
Year103	0.0264	(0.1981)
Year104	0.1080	(0.0000)***
Year106	-0.1012	(0.0006)***

Year107	0.1849	(0.0000)***
Year108	0.0020	(0.9401)
Adjusted R-squared	0.1017	(0.0000)***

註： *P<0.01；**P<0.005；***P<0.001

表 6-20 統整了模型一南區分析結果，由表 6-20 可得 Model1 針對流標情況的 Adjusted R-squared 為 10.17%，意指 Model1 的設計能藉由此份分析資料解釋本研究探討的現象 10.17%，此結果對於社會科學的回歸分析而言，解釋力偏低。在 Model1 的各變數中，第一季、第四季、農村更新、工程規模之金額、設計預算彈性皆為顯著，其自變數顯著性結果與模型一全區雷同。而在固定其他變數的條件下，在季節變數中，第四季表現最佳，發包遲延的機率約小 20.11%，第一季表現次之，發包遲延的機率約小 17.58%，第二季與第三季的差異不大，表現較差。在工程屬性變數，農村更新的表現最差，發包遲延的機率約大 6.17%。對於工程規模之金額而言，其係數為 1.99×10^{-5} ，若的平均數 257 萬、標準差 244 萬的條件下，每提升 244 萬的工程金額，發包遲延的機率約大 4.86%。對於設計預算執行彈性而言，其彈性越大(若比值為 1)，則發包遲延的機率約小 19.38%。

表 6-20 模型一各區回歸結果整理(東區)

	Model 1 (Bid_Delay)	
	Coefficient (回歸係數)	P-Value
Rural	0.0058	(0.7604)
Season1	-0.3013	(0.0000)***
Season2	-0.0997	(0.0000)***

Season4	-0.1112	(0.0001)***
Attr_flood	-0.0964	(0.0005)***
Attr_roadimprove	0.0542	(0.0262)**
Attr_villagerenew	0.2429	(0.0000)***
Price	9.76×10^{-6}	(0.0061)***
BD_flex	-0.1330	(0.1768)
Year103	0.1058	(0.0030)***
Year104	0.2947	(0.0000)***
Year106	0.0460	(0.2716)
Year107	0.2262	(0.0000)***
Year108	0.2620	(0.0000)***
Adjusted R-squared	0.2292	(0.0000)***

註：*P<0.1；**P<0.05；***P<0.01

表 6-21 統整了模型一東區分析結果，由表 6-21 可得 Model1 針對流標情況的 Adjusted R-squared 為 22.92%，意指 Model1 的設計能藉由此份分析資料解釋本研究探討的現象 22.92%，此結果對於社會科學的回歸分析而言，解釋力尚可。在 Model1 的各變數中，第一季、第二季、第四季、流域治理、農路改善、農村更新、工程規模之金額皆為顯著，其自變數顯著性結果與模型一全區雷同。而在固定其他變數的條件下，在季節變數中，第一季表現最佳，發包遲延的機率約小 30.13%，第四季表現次之，發包遲延的機率約小 11.12%，第二季尚可，發包遲延的機率約小 9.97%，而第三季的表現最差。在工程屬性變數，流域治理的表現最佳，發包遲延的機率約小 9.64%，農村更新的表現最差，發包遲延的機率約大 24.29%，農路改善的表現次差，發

包遲延的機率約大 5.42%。對於工程規模之金額而言，其係數為 9.76×10^{-6} ，若在平均數 311 萬、標準差 287 萬的條件下，每提升 287 萬的工程金額，發包遲延的機率約大 2.8%。

前者分析完模型一各區變數之顯著關係，後進行模型二各區變數之顯著關係分析。

6.4.4 模型二各區回歸結果與假設驗證

表 6-21 模型二各區回歸結果整理(北區)

	Model 2 (Project_Delay)	
	Coefficient (回歸係數)	P-Value
Rural	0.0077	(0.8241)
Season1	0.0625	(0.2179)
Season2	0.0654	(0.1248)
Season4	0.0087	(0.8772)
Attr_flood	0.0566	(0.2012)
Attr_roadimprove	0.0435	(0.3163)
Attr_villagerenew	0.1107	(0.0096)***
Price	9.97×10^{-6}	(0.0142)**
BD_flex	-0.0528	(0.7034)
Rainday_percent	-0.1045	(0.1367)
Year103	-0.1601	(0.0016)***
Year104	0.1292	(0.0072)***
Year106	0.0615	(0.2359)
Year107	0.0117	(0.8334)
Year108	0.0024	(0.9651)
Adjusted R-squared	0.0552	(0.0000)***

註： *P<0.1；**P<0.05；***P<0.01

表 6-22 統整了模型二北區分析結果，由表 6-22 可得 Model2 針

對工程逾期狀況的 Adjusted R-squared 為 5.5%，意指 Model2 的設計能藉由此份分析資料解釋本研究探討的現象 5.5%，此結果對於社會科學的回歸分析而言，解釋力稍弱。在 Model2 的各變數中，農村更新、工程規模之金額為顯著，其於變數皆不顯著，自變數顯著性結果與模型二全區部份雷同。在固定其他變數情況下，對於季節變數而言，四季的表現沒有太大的差異，與工程逾期的關聯性不顯著。在工程屬性變數，農村更新的表現較差，工程逾期的機率約大 11%，其餘屬性差異不大。在固定其他變數情況下，工程規模之金額變數其係數為 9.97×10^{-6} ，若在平均數 482 萬、標準差 383 萬的條件下，每提升 383 萬的工程金額，工程逾期的機率約大 3.81%。

表 6-22 模型二各區回歸結果整理(中區)

	Model 2 (Project_Delay)	
	Coefficient (回歸係數)	P-Value
Rural	0.0389	(0.0345)
Season1	-0.0172	(0.5548)
Season2	0.0456	(0.0350)
Season4	-0.1515	(0.0000)***
Attr_flood	-0.0258	(0.2315)
Attr_roadimprove	0.0669	(0.0012)**
Attr_villagerenew	0.3028	(0.0000)***
Price	2.1×10^{-5}	(0.0000)***
BD_flex	0.0572	(0.3674)
Rainday_percent	-0.1745	(0.0000)***
Year103	-0.0070	(0.7666)
Year104	0.0268	(0.2344)

Year106	0.0375	(0.1686)
Year107	0.0236	(0.3965)
Year108	-0.0256	(0.3066)
Adjusted R-squared	0.0995	(0.0000)***

註： *P<0.01；**P<0.005；***P<0.001

表 6-23 統整了模型二中區分析結果，由表 6-23 可得 Model2 針對工程逾期狀況的 Adjusted R-squared 為 9.95%，意指 Model2 的設計能藉由此份分析資料解釋本研究探討的現象 9.95%，此結果對於社會科學的回歸分析而言，解釋力尚可。在 Model2 的各變數中，第四季、農路改善、農村更新、工程規模之金額、汛期皆為顯著，其自變數顯著性結果與模型二全區部份雷同。在固定其他變數情況下，對於季節變數而言，第四季的表現最佳，工程逾期的機率約小 15.15%，其餘三季無顯著差異。在工程屬性變數，農村更新的表現較差，工程逾期的機率約大 30.28%，其次為農路改善，工程逾期的機率約大 6.69%。在固定其他變數情況下，工程規模之金額變數其係數為 2.1×10^{-5} ，若在平均數 268 萬、標準差 262 萬的條件下，每提升 262 萬的工程金額，工程逾期的可能性約上升 5.5%。對於汛期而言，若汛期佔施工期間的比例越高(若為 100%)，其工程逾期約下降 17.45%。

表 6-23 模型二各區回歸結果整理(南區)

	Model 2 (Project_Delay)	
	Coefficient (回歸係數)	P-Value
Rural	-0.0239	(0.1418)
Season1	-0.1249	(0.0000)***
Season2	0.0699	(0.0015)**
Season4	-0.1623	(0.0000)***
Attr_flood	-0.0128	(0.5698)
Attr_roadimprove	0.0864	(0.0000)***
Attr_villagerenew	0.2029	(0.0000)***
Price	2.62×10^{-5}	(0.0000)***
BD_flex	-0.0537	(0.0334)
Rainday_percent	-0.1308	(0.0000)***
Year103	0.0065	(0.7748)
Year104	-0.0322	(0.1533)
Year106	-0.0583	(0.0287)
Year107	-0.0244	(0.3477)
Year108	-0.0868	(0.0006)***
Adjusted R-squared	0.0662	(0.0000)***

註： *P<0.01；**P<0.005；***P<0.001

表 6-24 統整了模型二南區分析結果，由表 6-24 可得 Model2 針對工程逾期狀況的 Adjusted R-squared 為 6.62%，意指 Model2 的設計能藉由此份分析資料解釋本研究探討的現象 6.62%，此結果對於社會科學的回歸分析而言，解釋力稍差。在 Model2 的各變數中，第一季、第二季、第四季、農路改善、農村更新、工程規模之金額、汛期皆為顯著，其自變數顯著性結果與模型二全區雷同。在固定其他變數情況下，對於季節變數而言，第四季的表現最佳，工程逾期的機率約小

16.23%，第一季表現次佳，工程逾期的機率約小 12.49%，第二季的表現較差，工程逾期的機率約大 6.99%。在工程屬性變數，農村更新的表現較差，工程逾期的機率約大 20.29%，農路改善的表現次差，工程逾期的機率約大 8.64%，其餘無顯著差異。在固定其他變數情況下，工程規模之金額變數其係數為 2.62×10^{-5} ，若在平均數 266 萬、標準差 260 萬的條件下，每提升 260 萬的工程金額，工程逾期的機率約大 6.81%。對於汛期而言，若汛期佔施工期間的比例越高(若為 100%)，其工程逾期的機率約小 13.08%。

表 6-24 模型二各區回歸結果整理(東區)

	Model 2 (Project_Delay)	
	Coefficient (回歸係數)	P-Value
Rural	0.0151	(0.5968)
Season1	-0.1218	(0.0210)**
Season2	-0.0078	(0.8321)
Season4	-0.1616	(0.0005)***
Attr_flood	-0.0016	(0.9655)
Attr_roadimprove	0.0332	(0.3497)
Attr_villagerenew	0.1948	(0.0000)***
Price	9.73×10^{-6}	(0.0526)*
BD_flex	0.0772	(0.5914)
Rainday_percent	-0.0742	(0.1394)
Year103	-0.0531	(0.2332)
Year104	-0.0394	(0.3718)
Year106	-0.0446	(0.3520)
Year107	-0.0824	(0.0694)*
Year108	-0.0282	(0.5507)
Adjusted R-squared	0.0386	(0.0000)***

註： * $P < 0.1$ ；** $P < 0.05$ ；*** $P < 0.01$

表 6-25 統整了模型二東區分析結果，由表 6-25 可得 Model2 針對工程逾期狀況的 Adjusted R-squared 為 3.86%，意指 Model2 的設計能藉由此份分析資料解釋本研究探討的現象 3.86%，此結果對於社會科學的回歸分析而言，解釋力稍低，但依舊可探討出重要變數間之顯著性。在 Model2 的各變數中，第一季、第四季、農村更新、工程規模之金額為顯著，其自變數顯著性結果與模型二全區部份雷同。在固定其他變數情況下，對於季節變數而言，第四季的表現最佳，工程逾期的機率約小 16.16%，第一季表現次佳，工程逾期的機率約小 12.18%。在工程屬性變數，農村更新的表現較差，工程逾期的機率約大 19.48%，其餘無顯著差異。在固定其他變數情況下，工程規模之金額變數其係數為 9.73×10^{-6} ，若在平均數 311 萬、標準差 290 萬的條件下，每提升 290 萬的工程金額，工程逾期的機率約大 2.82%。

6.4.5 模型一綜合討論

針對模型一之分析結果，首先討論較為特別的幾個變數。

模型一之季節變數以專案核定時間作為判定的依據，即以此為出發點進行討論，在固定其餘所有的變數後，全區的回歸結果中呈現第一季的核定季節為工程順利發包的最佳時間，其次是第四季。第一季為整年度的開頭，對於廠商而言，是衡量整年度公司營運是否順利，

以及規劃案件執行期程的重要時間，廠商會為了公司的業務盤算，以確保一整年不會沒有案件執行，因此盡可能的投標，而間接影響發包的成功率，另外，在第一季案件投標的排擠壓力較低，也會間接影響到發包成功率；而某些機關常在前一年度的第四季即開始安排隔年度的工程及計畫，因此對於廠商而言，第四季亦是一個重要的期間，可提前安排公司隔年的業務。相較於第二季、第三季，又以第三季表現最差，可能因為年度業務規劃或案量已達負荷，減少廠商投標意願，進而降低機關發包成功率。

模型一之工程屬性變數以流域治理為發包最為順利之工程屬性，而農村再生為發包最不順利之工程屬性。由本研究整理之四大類工程屬性定義，流域治理應屬較為複雜且風險較高之專案，然而對於分析結果而言卻截然相反，本研究認為承攬水保局之廠商應已累積一定的流域治理工程經驗，執行機關該考慮的不確定性及風險也都有事先考慮完善，因此對於此類的專案駕輕就熟；反倒是農村再生的屬性，因農村再生類似執行社區營造，不僅景觀類工程，老舊村落社區的營造與新式發展也居多，對於執行工程的廠商而言並非強項，部份農村再生的設計及規格十分特殊，亦不是所有廠商皆有能力執行，有能力、有意願投標的廠商極大的可能性為特定廠商，且特定廠商屬少數，因此對於投標的制度而言十分不利，普遍首次開標需要三家廠商，若特

定廠商不足，首次開標即決標的可能性低，甚至有高機率流標，本研究認為此為水保局案件當中工程屬性所呈現出的潛在現況，並且與機關經驗有所出入。而農路因其工程內容多為防災工程及路面工程，材料以鋼筋、混凝土、模板、瀝青混凝土為主，項目單純且入門門檻低，多數廠商均有能力投標。

對於工程規模之金額而言則與原先設定之假說有所出入，原先就機關的經驗和認定來說，案件越大的越容易發包，然而就分析結果而言，案件越大，其工程遲延的機率大 2.6%，雖然數值不大，但以案件的平均金額而言，此結果可建議主管機關在未來核定案件費用時參考，以制定一個成功率較高的工程金額。而設計預算彈性即十分直觀，設計工程款與工程預算的比例越高，表示後續發包在一定的折扣下，廠商能運用的工程款費用越多，執行上亦較為彈性，風險隨之降低，與本研究所設立之假說方向一致。然而如同第五章節所述，發包遲延的計算標準礙於管考系統欄位的限制，並非全部代表發包的行為，其中包含發包前的執行機關準備期以及行政程序，工程金額越大的準備期間需較多時間，因此本研究認為，後續於發包遲延的應變數計算上需配合執行機關提供更細節資訊，稍作修正，才可更準確的分析出其潛在的根本影響因子。

在控制變數所呈現的結果來說，不同年度間、地區間的顯著性十

分明顯，亦代表著不同的時空背景無法清楚解釋的差異，以及地區間無法清楚表示的文化習慣，是需要被控制，以減少其他自變數的偏差以及落入偽因果的錯誤判斷。

針對工程區域變數得知，偏遠與發包遲延沒有顯著相關性，這與實證前的思緒有所出入，以往認為較為偏遠的地區施作廠商的意願可能比較低，經本研究討論過後，在所有條件均一致的狀況下，可能有某部份的廠商對於專案已累積足夠的經驗施作，只要給定充足的專案條件，因此不管偏遠與否，皆有廠商來進行承接，因此工程區域變數的效果不顯著。

6.4.6 模型二全區綜合討論

針對模型二之分析結果，亦先討論較為特別的幾個變數。

模型二之季節變數以開工時間作為判定的依據，即以此為出發點進行討論，在固定其餘所有的變數後，全區的回歸中呈現第四季的開工季節為工程執行最佳時間，其次第二季為最差的開工季節，對於第四季的结果而言，本研究認為第四季對於主辦機關及廠商皆是非常重要的時間點，主辦機關被要求達一定的執行率，因此第四季發包之工程為有利廠商提早進場施工，除法定等標公告期限以外，執行機關可針對決標後至訂約期間以及訂約後至開工期間進行行政程序及期程加速，以利提升執行效率，而在其他季因時間充足，而較不會加速期

程及提升效率；另對於廠商而言，年前執行工程數量的多寡，會直接影響其申請工程款項的數目，因此在執行機關要求之下，必然會影響到廠商的對於工程的積極度，以利年度請款有益於資金調度。而第二季執行狀況較差，在於第二季廠商累積的案子越來越多，案子執行的優先順序將會影響到執行成效。

模型二之工程屬性變數以農村再生為工程執行狀況最不佳之屬性，其次為農路改善。有鑑於模型一所提到農村再生同時為容易發包延遲的原因，本研究認為基於其介面設計的複雜度、工程材料種類繁多且部份特殊性高、施工方式難度較高，時常需要特殊工法亦或是材料需要廠商工班支援，在機具、工班、材料的調度上容易影響施工進度，因此農村的工程逾期可能性最高，因此可使執行機關重新評估，實際上是否有農村再生案件給予之執行條件不足的現象存在，可於未來參考並改善。相較於農村再生，農路改善雖工程項目單純且入門門檻低，工程較容易控制，但亦有逾期的風險存在，但可能性不高。

模型二對於工程規模之金額之結果十分直觀，在其餘變數固定的條件下，工程金額越大的逾期機率越高。

模型二對於汛期而言，其結果與本研究設定之假說有所出入，就常理來說，施工期間包含汛期的時間越長，逾期的可能性應該越大，因為梅雨、颱風將會影響工程進行，並且若非一定程度的嚴重，給予

展延的可能性較低，然而就分析結果表示，汛期佔施工期間越長的，工程逾期的可能性越低，本研究認為，應是在規劃案件時已知會遭遇汛期，因此給予的執行條件及需承擔的風險皆已納入考量，甚至給的條件十分充足，因此才會有結果的方向與假說相反的現象產生。同第五章節所述，在應變數工程逾期的計算上礙於管考系統裡的資料欄位有限，本研究僅採用開工時間、核定工作日與實際完工日期作比較，計畫執行途中合理的工期展延並非記錄進去，並且缺乏工程變更的次數與資料，亦會影響到逾期效果的分析，因此本研究建議在工程逾期的辨定上需由機關協助提供工程專案更細部執行記錄，使對於工程逾期的分析結果更為精準。

針對模型二之分析結果，在控制變數上依舊有關聯性，因此需要被控制以確定自變數不會被其他因素所影響。另外，工程區域的偏遠對於工程執行狀況之效果不顯著，就實務上而言，會執行較為偏遠地區的廠商對於施作環境應累積充足經驗，廠商有足夠的經費、足夠的工期即可作，不因工程的偏僻程度而受影響，以致於並非影響執行狀況重要的因素。

然而模型二全區的解釋力雖僅達 7%，意指就工程管考系統內的資料欄位以及執行機關人員及長官所關心之變數，並不是主要影響工程逾期狀況的原因，但亦有討論出一些特殊的因素。而本研究驗證

Model2 假說之影響，大部份並非真正影響工程執行之原因，與執行機關之想法有所衝突，扣除掉工程施工區域、季節、工程屬性、工程規模之金額、汛期等自變數，主要影響的變數皆在殘差項內，本研究合理猜測主要影響工程逾期狀況的因素並非為工程管考系統內記錄的系統化資料，而是其他未控制的因素，如廠商品質、專案管理能力、案子之特殊性等人為或不易量化之原因，換句話說，未解釋之原因為管理面(管理能力、管理策略)及廠商選擇之問題，亦可延伸至專案管理的重要性，包括事前的規畫、選商、至包商管理皆是影響工程執行控管上的主要工作，工程規模越大、複雜性越高，與管理面之能力關聯性亦非常大。再者，水保局內的制度，包含工程費用的發放、管理人員的執行能力(包括管理單位、人員、人力、負責的案量大小)，亦有很大的關聯性，也是本研究這次模型無法納入的變數。而案件的特殊狀況，如地質情況、環境因素、水文水理的影響及天災極端情況的發生，亦是影響工程執行逾期情況的重要關鍵，此些資料欄位也不易從工程管考系統中取得，因此本研究建議於後續研究上，可減少分析樣本數，增加分析變數，以案例分析討論的方法深入討論本研究提出未納入模型之變數是否為影響的重要因子，並且與執行機關合作，調閱案子的細部資料，以供後續管考系統欄位新增參考。

本研究參考水土保持工程管考大數據資料集分析暨智慧應用先

期計畫產出，其工程預警系統之 S-curve 四個變數之實際累積資料為資料分析基礎，並持續修正模型以完善其準確率，此與本研究之主旨不同，本研究透過結果分析期望貢獻真正影響發包遲延以及工程逾期狀況的實際問題，並深入討論是否有其他變數是會影響但並未納入，以驗證執行機關之疑慮之真實性。

6.5 策略擬定

基於 6.4 所得之分析與綜合討論之結果，本研究針對發包遲延及工程逾期擬定初步改善與策略方針：

針對發包效能：

- (1)由模型一之結果得知，無論在專案的核定季節或專案的開工季節皆有顯著的關聯性，本研究認為針對此變數，可依照案件重要性、規模等，有策略規畫核定時間於不同季節，增加發包率。
- (2)由模型一之結果得知，雖然本次採用之工程屬性為大範圍之計畫，但可明顯得知在農村再生的相關專案中，其發包與執行的狀況皆不甚理想，針對農村再生的工程特性與執行內容來說，景觀等特殊工法可透過工法交流研討會議，並且提升廠商能力避免因僅少數廠商具備施工能力影響發包效能。
- (3)綜合模型一之結果，本研究若要減少發包之遲延時間，可搭配有效的發包策略與方法，如提高評分及格制發包比例(減少單純價格標)，

以鼓勵並保障能力好的廠商。

針對工程進度：

- (1)由模型二之結果得知，第四季之執行狀況最佳，執行機行可以第四季為範例，優化其他季執行模式，並以估驗機制或獎勵刺激廠商效率。
- (2)由模型二之結果得知，農村再生的工程逾期狀況最為嚴重，建議重新評估農村再生計畫類別給予的金額、工期、設計是否過於嚴苛，並調整核定內容。
- (3)綜合模型二之結果，本研究認為若要改善工程逾期之狀況，可針對進度超前訂定獎勵制度如趕工獎金或優先承攬（配合評分及格最低標）。

整體而言，機關可於計畫核定時，把時間花費、工程複雜度、氣候影響之潛在風險納入，並配合廠商生態於不同時間點投入不同資源，並以資格保護發包、獎勵刺激執行，以此修改管理面的方針。

本研究對於管考系統可新增之欄位及潛在影響發包遲延及工程逾期之變數進行整理，從發包、成本、進度的角度下進行討論：

針對發包遲延的欄位資料，包含工程分類、流標次數記錄、開標會議日期、底價的更動記錄、案件特殊性記錄等。

針對工程逾期的欄位資料，包含工程分類、申請展延許可之情況、契約變更(工期、金額)之記錄、工程遭遇之爭議、工程項目調解事項、

合解金額、訴訟金額、案件重大行為等。

針對工程成本的欄位資料，可增加每個階段里程碑的預估成本、實際花費金額、結案的情形等。

針對廠商品質的資料，可建置承接過執行機關之廠商資料庫，記錄廠商規模，並納入工程案件檢討報告，針對工程執行完成後之品質進行評等，作為後續投標之參考依據，亦可與營建署、工程會已建立之資料庫進行連結，分享廠商能力及品質之記錄，如表 6-25 所示。

表 6-25 新增資訊欄位建議表

議題	新增欄位資訊
發包遲延	精準工程分類、流標次數記錄、招標會議記錄、底價的更動記錄、案件特殊性記錄等
工程逾期	精準工程分類、申請展延許可之情況、契約變更(工期、金額)之記錄、工程遭遇之爭議、案件重大行為、工程項目調解事項、合解金額、訴訟金額等
成本控管	各階段里程碑的預估成本、實際花費工程款、結案情形等
廠商品質	建置廠商資料庫、附帶廠商工程案件檢討報告、完工品質評等、與他部會分享資訊等

第七章 結論與建議

7.1 結論

本研究採用大數據計量分析，以大數據統計為基礎，大數據計量回歸為方法，分析過程中由專家先協助判斷重要欄位，以設立假說、選擇應變數、自變數及控制變數，系統欄位並非全部使用至變數，而沒有意義或直觀判斷沒有影響的欄位即排除之外。

本研究於完成模型一與模型二之實證分析，透過參考文獻及水土保持局過往委辦計畫作為發想之依據，探索水土保持局工程管考系統資料庫，當中包含北、中、南、東各區以及全區之大數據計量分析，以此期望獲得有別於數值上之潛在關聯性。

模型一陳述在應變數流標情況與自變數之間的關係，發現顯著的變數包含：季節(第一季、第四季)，又以第一季的表現情況最好，其主要原因為廠商承接每年度案件的時程規劃；工程屬性(流域治理、農村再生)，又以流域治理的表現情況最佳，其主要原因為廠商充份的經驗加上執行單位將應考量的因素妥善納入，而較不容易流標，然而農村再生的表現情況最差，其主要原因在於農村再生的設計過於特定，在施作的廠商中要願意且有能力的並非多數，因而造成發包狀況不佳；工程規模之金額呈現越大的案件越容易流標；執行彈性越大的案件，則越不容易流標，此結果可供執行機關作為未來發包策略之擬定。

模型二陳述在應變數工程逾期狀況自變數之間的關係，發現顯著的變數包含：季節(第二季、第四季)，又第四季的表現最佳，主要原因本研究認為第四季發包工程為有利廠商提早進場施工，除法定等標公告期限以外，執行機關可針對決標後至訂約期間以及訂約後至開工期間進行行政程序及期程加速，以利提升執行效率，增加執行率，進而促使廠商積極施作，亦有利於廠商年底請款進行資金調度，而第二季因為屬年中，執行機關與施作廠商較容易不積極執行相關業務或進行工項施作；工程屬性(農路改善、農村再生)又以農村再生的執行狀況最差，主要原因是基於農村再生的特殊性，在其介面設計的複雜度、工程材料種類繁多且部份特殊性高、施工方式難度較高，時常需要特殊工法亦或是材料需要廠商工班支援，在機具、工班、材料的調度上容易影響施工進度，亦需由執行機關於未來評估是否給予農村再生的執行條件不足，而農路改善雖施作內容較為單純，但亦有逾期之風險存在；工程規模之金額與工期，此自變數為逾期的重要影響因子，而在各別固定及其餘變數固定的情況下，金額越高逾期風險越高，而工期越多逾期風險越低，與實際上狀況與經驗吻合；汛期的變數與原先設定之假設方向相反，本研究認為執行單位在規劃案件時已知遭遇汛期，並將其風險妥善納入考量並且給予足夠之條件，以致於汛期佔施工期限百分比越大的案

件，其逾期的風險越低。

而在年度以及地區的變數設計下，年度間難以解釋的差異，以及各區間的差異即被納入考量，也增加了回歸模型解釋的精確度。

模型二因解釋程度僅 7%，雖本研究於實證假說之探討因素下現幾項關鍵因子，其餘大部份並非影響工程執行逾期狀況之主要原因，變數間並沒有顯著的關聯性，進而探討廠商素質、專案管理能力、案件的特殊狀況或許才是影響工程逾期的可能原因。

7.2 建議

- 1.本研究於研究過程中發現，部份資料在某些重要欄位常缺少資料，應完整每件案件的資訊，以增進未來進行分析之資料完整性。
- 2.本研究於工程管考系統中所採用的欄位已接近全部，未來可新增更多的資料欄位或與外界開放的資料進行連結，包含廠商的規模、廠商的品質、更細部的工程屬性、更精準的流標次數、更完整的工程變更資料、每一次的招標記錄等等，以提供更多潛在量化的因素進行計量的分析。
- 3.因本研究較屬於以大數據進行因素探索，將比較有可能影響逾期、流標的變數納入，並且比較不相關、不應該放、沒有影響的變數從模型中拿掉，日後研究，可針對本研究找出來的變數，再以個案的資料(更細部的資料，為紙本有但數據庫沒有的)進行分析，透

過減少分析的資料數量，來增加精準度，從探索式研究進入驗證性研究，以驗證一萬多筆資料的結果是否正確。

4.本研究認為，基於本次研究結果，後續可作的研究方向可用個案研究方法，針對表現最好的及表現不好的個案去深入了解，探討與本次結果是否相呼應，亦可從各案中得到不同的關聯性，並增加為新的假說進行驗證。因本次模型二解釋力大概 7%，有不少的重要的因素在資料庫中無法發掘，亦可以透過個案研究跟計量分析去發現。

5 基於本研究所提出之自變數的影響，可供執行機關於未來制定工程發包策略，與工程執行狀況之管理方針。

參考文獻

1. 計量經濟學(Introductory Econometrics A Modern Approach), Jeffrey M. Wooldridge (胥愛琦 譯), Cengage, six edition, 2017
2. 鐘俊凱, “工程逾期造成工程爭議問題之研究”, 國立高雄第一科技大學, 碩士論文, 2014。
3. 王華弘、謝彥安, “道路鋪面工程中展延工期問題之研究”, 鋪面工程, 第 14 卷, 第 1 期, 15-22 頁, 2016。
4. 戴貞德、林信宗, “政府採購雙贏決策之研究-以台電高雄區營業處為例”, 商業現代化學刊, 第 5 卷, 第 1 期, 73-84 頁, 2009。
5. 黃正翰, “以 BIM 為基礎建構水保工程應用 4D 與 5D 模型於施工階段之研究”, 農委會水土保持局 107 年計畫, 期中報告書
6. 陳美心、周天穎、林峰正、林伯勳, “水土保持工程管考大數據資料集分析暨智慧應用先期計畫”, 農委會水土保持局 104 年計畫
7. 國立臺灣大學委辦, “集水區保育治理巨量資料分析先期示範計畫”, 農委會水土保持局 104 年計畫
8. Anuradha J. A brief introduction on Big Data 5Vs characteristics and Hadoop technology. Procedia Computer Science. 2015; 48:319-24.
9. A. Oguntimilehin and O. Ademola, ‘A Review of Big Data Management , Benefits and Challenges A Review of Big Data Management , Benefits and Challenges’ , J. Emerg. Trends Comput. Inf. Sci., vol. 5, no. July 2014, pp. 433 – 438, 2015.
10. Bottles, K., Begoli, E. & Worley, B., 2014, ‘Understanding the

pros and cons of big data analytics’ , Physician Executive 40(4),
6 – 12.

11. White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *econometrica*, 48(4), pp. 817-838.
12. 內政部，“山地、平地原住民及離島等偏遠地區一覽表(附件二十七)”，
https://www.cdc.gov.tw/File/Get/oqXE2BJJpJ4L_MmBKWtRw
13. 李玫郁，“計量大數據分析、人工智慧基礎概念”，2018 月 4 月 21, ,
<http://econcloud.blogspot.com/2018/04/blog-post.html>

附錄

附錄一、期初審查會議紀錄暨回覆辦理情形

項次	審查意見	回覆辦理情形
報告內容審查意見：		
一	P. 1-1 參考基準(如…)等，在既有管考系統資料中，如何取得該項資料。	於工作執行計畫書中先提出一個方向，後續將透過申請工程管考系統權限，並針對系統內有的欄位進行討論。
二	P. 1-2 研究範圍以臺灣北部地區，請列出行政區域名稱。	北部地區主要著重於山地原住民鄉，包含烏來、復興、五峰、尖石、大同、南澳，其餘依照欄位登記之所有台北分局執行專案。
三	P. 2-6 第四節方法概述與第三節方法概述有何關聯性之步驟，其中P. 2-7與P. 2-2第節資料類型相似。	已更正第四節為研究流程
四	P. 1-2 本計畫擬將研究範圍設定於台灣北部地區，收集歷年包含易淹水地區水患治理計畫、流域綜合治理計畫、整體性治山防災計畫、重劃區外緊急農路設施改善計畫等各項不同類型之水土保持工程管考資料進行分析；其中，北部地區之工程數量佔全體約3萬5000件之比例為何，是否能代表全體特性。	本研究首要針對北區之資料建立模型與變數，後續將納入全區資料，探討代表全體的變數關係與特性。

項次	審查意見	回覆辦理情形
	P.1-2 第四節目前計畫作業流程建議可與第三章工作進度合併，並詳細列出圖1-1計畫整體之作業流程。	將於期中報告書中進行修正
	P.2-2 關聯分析中，進行工程進度與其他可能影響因子之關聯關係分析，是否亦會探討可能影響因子等自變數間之關聯性。	於分析結果的章節會呈現自變數之間的相關矩陣
	P.2-2 本計畫擬利用過去特定區間歷史資料來訓練出預估工期品質模型，再以目前正在執行的工程案件做為新資料輸入，用以預測預估工期品質。建議增加訓練出模型之準確度評估，以瞭解模型之適用性。	將納入後續研究過程參考
	P.2-7 檢視資料特性與水土保持議題之相關性與完整性中，說明數據來源自管考系統資料庫尋找，包括統計年鑑、相關數據之論文作者。其中統計年鑑、相關數據之論文作者所指為何？這些資料是否能涵蓋全部工程案件？	此為一個研究開頭之方向，包含各年度統計資料、並蒐集是否有相關文獻之研究概念與本計畫相似。若納入全區的資料僅能說涵蓋工程管考系統中擁有的工程案件。
	P.2-4工程影像及日誌類資料於工程管考系統內較不齊全，此部分若分析成果精度或較難做後續採用，可於本研究內針對分析成果做後續管考系統收納資料及分析方向建議。	將納入研究過程參考

項次	審查意見	回覆辦理情形
	P. 2-7預期分析主題規劃需要分析資料此研究暫列24項，建議其中可納入「開工日期」，另「變更次數」於108年以前，管考系統上無確切紀錄登載，該資料可否取得請再評估。	將納入後續研究過程參考並修正
	本局104年度辦理「水土保持工程管考大數據資料集分析暨智慧應用先期計畫」有本研究類似研究方法與成果，請本研究參考並於內容闡明與該計畫已有成果不同之處，或為針對該計畫成果不足及可延伸議題之後續研究分析方向。	將納入後續研究過程參考，並於期中報告書進一步說明

附錄二、期中審查會議紀錄暨回覆辦理情形

項次	審查意見	回覆辦理情形
報告內容審查意見：		
一	P. 1-2應說明大尺度全台灣崩塌規模及中尺度流域土砂量分析與原先計畫目標之關聯。	本研究參考大尺度全台灣崩塌規模及中尺度流域土砂量分析著重在其已完成收集巨量資料，符合原先本研究目標“巨量資料探勘之本意。
二	P. 1-4大尺度全台灣崩塌規模及中尺度流域土砂量分析中，應說明分析目的為何？應變數、自變數、控制變數為何？	本研究在探討完“大尺度全台灣崩塌規模及中尺度流域土砂量分析”之資料集後，經過討論認為其資料類型與本研究所採用之類型不同，亦與本研究想探討的方向有所出入，因此將此部份概念從本計畫中移除，並於期末報告中進行說明。
三	P. 5-2第三行錯誤！找不到參照來源等文字遺漏處，請修正。	已修正
	P. 5-3 Hetero Robust test 中，判斷參數和消費者對於智慧建築的購買意願之間的相關性，與本研究之關聯？	已修正錯誤敘述部份

項次	審查意見	回覆辦理情形
	P.5-4 發包順利與否以天數小於90天為0，天數大於90天為1，請說明採用90天之依據為何？若將發包順利與否改為發包天數，亦即將應變數已連續變數進行分析，是否更能探討自變數對發包之影響？	90天之依據來自與水土保持局承辦人員討論，亦有討論採用60天，而期中設計以90天為先，並於結果分析後討論是否修正判別依據。若改為發包天數，與本研究設定之模型概念有所出入，呈現的係數與解釋的方向會有所不同，而將會納入後續修正參考
	P.5-5 工程屬性以流域綜合治理計畫、重劃區外緊急農路設施改善計畫等四個類別進行分類，是否能真正反應水土保持工程的屬性？	本研究因採用工程管考系統中之資料欄位，於期中採用大方向分類之類別，後續可討論是否採用其中更細的計畫類別
	P.5-7 模型一包含一個控制變數，但後文又包含發包彈性與年度，請確認。	已修正
	P.5-8 模型二包含工程變更，應說明此自變數如何衡量，又年度於模型二屬於自變數？	已修正並加強敘述。本研究於討論後移除“工程變更”變數。年度於模型二屬控制變數

項次	審查意見	回覆辦理情形
	P. 6-8 模型一中以105年度做為基礎，分析結果不同年間有所差異，其可能潛在因素為何？	於期中報告書中 5.3.3 控制變數之選定 P5-6 年度項目進行說明，控制變數本研究不多作敘述，確切原因無法明確從模型中得知，但若不納入此控制變數將會使其他欲了解的自變數偏差
	期初審查會亦請補充回覆辦理情形。	已補上
	PI摘要本計畫並未做水保局相關單位專訪，而是前人研究成果，應予釐清；另最後第2行本期初…以後均移除，英文部分亦同。	已修正
	研究方法中哪個才是本創研計畫研究與其他委辦計畫不同的內容。	研究方法中，僅採用委辦計畫之數據及專家訪談結果進行方向參考，其餘方法皆與委辦計畫不同
	採104年委辦計劃研究所得內容與5年後工程管考之內容，差異頗大，如單價、無法發包原因、採開口約辦理工程等等，如此如何給水保局有幫忙，應有說明。	將納入討論，於期末報告進行說明

項次	審查意見	回覆辦理情形
	P3-6小節部分，受訪者為誰，是本次計畫辦理之受訪者嗎？	此結果為文獻回顧之參考
	最近幾年發包中有許多案件採評分及格最低標，其發包期程變長，對影響順利發包之假說變化很大，是否影響實證模型。	因工程管考系統中無法蒐集部份資料，因此發包方式並未納入本次研究，將於後續建議進行說明
	P5-3應變數之選定中本局發包天數降為60天為基礎。	將納入後續修正
	P5-9資料蒐集年度區間請說明。	採用 103-108 年度之資料

附錄三、期末審查會議紀錄暨回覆辦理情形

項次	審查意見	回覆辦理情形
報告內容審查意見：		
一	建議有案例分析深入探討以提高可讀性	本研究屬大數據探索性實證分析，如 7.2 建議第 3 點所述，後續研究可採用較少樣本數，增加分析變數，以深入探討方式提升研究議題的精準程度。
二	對於研究計畫成果之應用建議加強說明	針對本研究之分析結果，提出現階段策略改善或建議方針，計畫之成果供執行機關參考，並可由執行機關討論是否有管理上能調整的地方，後續若執行機關欲深入探討亦可協助辦理。
三		
四		