

高等教育對薪資收入的影響

楊孟翰*

December 29, 2024

Abstract

本研究旨在探討高等教育（尤其是研究所學歷）對薪資收入的影響。透過分析中研院 SRDA 的「人力運用調查」資料集，涵蓋 108 至 112 年共五個年度的資料，本研究納入 125,351 筆樣本，並聚焦於教育程度、薪資收入及其他控制變數（如性別、年齡、婚姻狀態、工作狀態等）。初步分析顯示，高等教育對薪資收入具有顯著的正向影響，並且受性別及工作類型的調節。研究結果將為提升臺灣地區高等教育政策及勞動市場決策提供參考。

*National Taiwan University, Email: my61130567@gmail.com

1 Introduction

隨著全球化與科技發展的迅速推進，高等教育在提升勞動市場競爭力方面的作用日益凸顯。然而，高等教育是否能顯著提升薪資收入，始終是政策制定者和經濟學研究的重要議題。本研究旨在填補這一領域中的空白，特別是針對臺灣地區的勞動市場情境進行實證分析。

高等教育對薪資收入的影響一直是廣受關注的議題。許多研究顯示，高學歷與較高的收入之間存在正相關。然而，臺灣勞動市場中此一現象的具體影響程度與細節仍有待進一步探討。本研究聚焦於教育程度對薪資收入的影響，並考量多元人口特徵與勞動條件，期望為相關文獻提供貢獻，並為教育與經濟政策提供實務建議。

研究的動機來自研究者作為研究生的經驗，對於研究所學歷的經濟效益有高度興趣。本研究的目的是量化高等教育對薪資收入的影響，並探索不同控制變數如性別、年齡及地區對此影響的調節作用。

研究的重要性在於為政策制定者提供實證基礎，幫助其設計更具效率的高等教育政策，同時也能為學生及家長在規劃教育投資時提供實用建議。

2 Data and Sample

2.1 Data

本研究採用中研院調查研究資料庫（SRDA）提供的「人力運用調查」資料集。該資料集涵蓋 108 至 112 年共五年的資料，包含 272,310 筆樣本及超過 47 個變數。研究的核心變數包括教育程度、每月薪資收入、性別、年齡、婚姻狀況、工作型態及地區位置。此資料集為探討高等教育與薪資收入之間的關係提供了詳細且全面的數據基礎。

2.2 Sample

本研究的樣本包含了從 108 年至 112 年的五年資料，來自於人力運用調查。樣本的構建包含以下步驟：

資料清理步驟：

1. 移除「月收入」欄位中未作答或填答無效（如跳答）以及收入為負數的觀測值。
2. 排除工作型態欄位中未作答的觀測值，確保樣本為有工作者。
3. 排除每週平均工作時數欄位中未作答的觀測值，確保所納入樣本均提供了有效的工作時數。
4. 排除目前工作場所工齡欄位中未作答的觀測值，以剔除無法確認穩定性的觀測樣本。
5. 將教育程度重新編碼：將碩士和博士學位（值為 8 和 9）設定為 1，其他教育水平設定為 0，以便進行比較。
6. 僅保留所需的分析變數，包括月收入、性別、年齡、地區、婚姻狀況、工作型態、每週平均工時、目前工作場所工齡及教育程度。

經過上述處理後，將五個年度的清理後數據合併為一個完整樣本。

最終樣本規模與時間範圍：清理完成後的數據包含 125,351 筆觀測值，涵蓋的時間範圍為 108 年至 112 年。本研究的樣本單位為個人，為橫截面資料的彙整，適合進行高等教育對薪資收入影響的實證分析。

2.3 Descriptive Statistics

以下表格展示了研究樣本的描述性統計數據，包含主要變數的均值和標準差，並比較了處理組（研究所以以上）和控制組（研究所以以下）之間的差異。主要變數包括月收入（Outcome Variable）和多個控制變數，如性別、年齡、地區、婚姻狀況、工作型態、每週平均工作時數和工齡。

Table 1: Descriptive Statistics of Outcome and Control Variables

Group	Mean Monthly Salary	SD Monthly Salary	Mean Age	SD Age	Mean Weekly Hours	SD Weekly Hours	Mean Work Years	SD Work Years
Control (Graduate and above)	39,955.75	25,502.07	45.35	12.68	41.16	5.98	11.31	9.61
Treatment (Below Graduate)	55,285.52	35,738.04	39.97	10.35	40.98	5.18	8.77	7.63

解釋：

- **月收入 (Mean Monthly Salary)：**處理組（研究所以以上）的月收入均值為 55,285 元，高於控制組（研究所以以下）的 39,955 元，顯示出教育水平對收入的影響。

- **年齡 (Mean Age):** 控制組的平均年齡為 45 歲，高於處理組的 40 歲，這可能是因為高等教育受試者進入職場的時間較晚。
- **每週工作時數 (Mean Weekly Hours):** 兩組的工作時數接近，處理組稍低於控制組。
- **工齡 (Mean Work Years):** 控制組的平均工齡 (11 年) 高於處理組 (9 年)，這可能與年齡差異相關。

2.4 Data Visualization

以下兩張圖片，對反映變數 y (每月薪資收入) 繪製直方圖以及盒形圖，可以更直觀感受控制組跟對照組之間的差異

解釋：

- 圖 1 描述了清理後數據的月收入分佈情形，顯示收入主要集中於 20,000 至 50,000 元的範圍內，但也存在少量高收入者。
- 圖 2 顯示了教育程度（研究所以上與以下）與月收入之間的關係。可以看出，研究所以上的教育群體在收入分佈中更集中於高收入區間。

3 Empirical Specifications

本研究使用傾向分數匹配法 (Propensity Score Matching, PSM) 來估計高等教育 (是否就讀研究所) 對薪資收入的因果影響。由於實驗組 (Treatment Group, 研究所以上) 與對照組 (Control Group, 研究所以下) 在一些控制變數 (如年齡、性別、婚姻狀況等) 上存在較大的差異，這可能會導致估計偏誤，因此我們通過匹配方法減少組間差異，從而提高因果推斷的準確性。

3.1 Identification Strategy

1. 使用邏輯迴歸模型估計每位樣本的傾向分數 (Propensity Score)，即接受處理 (就讀研究所) 的概率。

2. 以最近鄰匹配法 (Nearest Neighbor Matching) 為基礎，將實驗組與對照組進行匹配，匹配範圍限制為 0.2 的傾向分數差異 (Caliper = 0.2)。
3. 匹配後，我們比較實驗組和匹配樣本的對照組在薪資收入 (Outcome Variable) 上的差異，以估計平均處理效應 (Average Treatment Effect on the Treated, ATT)。
4. 為進一步分析異質性影響，將樣本按性別分為男性與女性兩組進行 PSM 匹配，探討不同性別之間就讀研究所所帶來的影響是否相同。
5. 將樣本按年齡分為 40 歲以下與 40 歲以上兩組進行 PSM 匹配，探討不同年齡就讀研究所帶來的影響是否相同。

3.2 Empirical Model

以下是用於傾向分數估計的邏輯迴歸模型：

$$P(T_i = 1|X_i) = \frac{e^{\beta_0 + \beta_1 \text{Gender}_i + \beta_2 \text{Age}_i + \beta_3 \text{Region}_i}}{1 + e^{\beta_0 + \beta_1 \text{Gender}_i + \beta_2 \text{Age}_i + \beta_3 \text{Region}_i + \beta_4 \text{Marital Status}_i}} + \frac{e^{\beta_4 \text{Marital Status}_i + \beta_5 \text{Employment Type}_i + \beta_6 \text{Weekly Hours}_i}}{1 + e^{\beta_5 \text{Employment Type}_i + \beta_6 \text{Weekly Hours}_i + \beta_7 \text{Work Years}_i}} \quad (1)$$

其中， T_i 表示第 i 位樣本是否屬於實驗組 ($T_i = 1$ 表示實驗組， $T_i = 0$ 表示對照組)， X_i 是控制變數集合。

在匹配完成後，我們計算平均處理效應 (ATT) 的公式如下：

$$\text{ATT} = \frac{1}{N_T} \sum_{i \in T} (Y_i^T - Y_i^C) \quad (2)$$

其中：

- N_T 是實驗組樣本的數量。
- Y_i^T 是實驗組第 i 位樣本的觀察結果 (即月薪)。
- Y_i^C 是與實驗組第 i 位樣本匹配的對照組樣本的觀察結果 (即月薪)。

異質性分析公式

爲了進一步分析不同性別和年齡群體的處理效應，使用以下公式：

性別異質性分析

$$ATT_{\text{Gender}} = \frac{1}{N_{\text{Male}}} \sum_{i \in \text{Male}} (Y_i^T - Y_i^C) + \frac{1}{N_{\text{Female}}} \sum_{i \in \text{Female}} (Y_i^T - Y_i^C) \quad (3)$$

其中：

- N_{Male} 和 N_{Female} 分別是男性和女性樣本的數量。
- Y_i^T 和 Y_i^C 表示處理組和對照組的薪資收入。

年齡異質性分析

$$ATT_{\text{Age}} = \frac{1}{N_{\text{Young}}} \sum_{i \in \text{Young}} (Y_i^T - Y_i^C) + \frac{1}{N_{\text{Old}}} \sum_{i \in \text{Old}} (Y_i^T - Y_i^C) \quad (4)$$

其中：

- N_{Young} 和 N_{Old} 分別是 40 歲以下和 40 歲以上群體的樣本數量。

異質性效應的比較公式

因爲我們需要比較不同群體之間的處理效應差異，可使用以下公式：

$$\Delta ATT = ATT_{\text{Group1}} - ATT_{\text{Group2}} \quad (5)$$

其中：

- Group1 和 Group2 可分別表示不同性別或年齡群體。

3.3 Key Assumptions

傾向分數匹配法的核心假設包括：

1. 平行趨勢假設 (**Unconfoundedness**)：給定控制變數 X ，處理變數（是否接受高等教育）與結果變數（薪資）是條件獨立的，即：

$$T \perp Y(0), Y(1) | X$$

這一假設要求納入模型的控制變數 X 能夠充分捕捉處理組與對照組之間的系統性差異。為驗證該假設，可通過以下方法：

- 使用標準化均值差（Standardized Mean Difference, SMD）評估匹配前後的變數分布差異。
- 使用虛擬結果迴歸測試（Placebo Test）檢驗未觀測變數對估計結果的影響。

2. 公共支持假設 (**Overlap**)：對於所有樣本，傾向分數的範圍都在 $(0, 1)$ 之間，即：

$$0 < P(T_i = 1 | X_i) < 1, \quad \forall i$$

為保證假設成立，應檢查匹配後處理組與對照組的傾向分數分布是否有足夠重疊。對於極端值（傾向分數接近 0 或 1）的樣本，可採取以下措施：

- 檢查傾向分數的分布並繪製重疊圖。
- 移除超出重疊區域的樣本，確保分析結果的穩健性。

此外，需注意傾向分數匹配法的局限性：例如，控制變數可能無法完全消除未觀測異質性的影響，或樣本的重疊不足可能降低分析的外部有效性。為此，可考慮結合加權匹配法（Weighted Matching）或雙重差分法（Difference-in-Differences, DID）進一步檢驗結果的穩健性。

4 Results

4.1 Graphical Evidence

圖 3 顯示了匹配前後共變數的平衡性檢查結果，以標準化平均差異（SMD）為衡量指標。圖中可以明顯看出，通過傾向分數匹配（Propensity Score Matching, PSM），實驗組與對照組之間的共變數平衡性得到了顯著改善。表 2 提供了匹配後樣本的描述性統計數據，揭示了研究所學歷對月薪的顯著正向影響。具體而言，平均處理效應（ATT）估計值為 15,628.54 元，這表明具有研究所學歷的群體平均每月收入比未達研究所學歷的群體高出約 15,628 元。

Table 2: 匹配後的描述性統計數據

群體	平均月收入 (元)	月收入標準差	平均年齡	年齡標準差	每週平均工作時數	工作時數標準差	平均工作年數	工作年數標準差
未達研究所學歷 (Treatment)	39,656.98	19,277.40	40.04	10.36	40.91	4.42	8.71	7.66
研究所學歷及以上 (Control)	55,285.52	35,738.04	39.97	10.35	40.98	5.19	8.77	7.63

解釋與討論：圖 3 顯示，通過匹配後，大部分共變數的標準化平均差異接近於零，顯示出實驗組與對照組之間的基線特徵已經得到了有效平衡。表 2 中的描述性統計數據進一步支持了研究所學歷對月薪的顯著正向影響。ATT 值為 15,628.54 元，表明高等教育對收入的潛在因果影響，這進一步證實了高等教育在提高薪資方面的價值。

4.2 Heterogeneity Analysis: Gender Differences

本研究進一步進行異質性分析，以探討不同性別群體中高等教育對薪資收入的影響是否存在顯著差異。使用傾向分數匹配（PSM）分別對男性與女性進行分析，並計算平均處理效果（ATT）及其變化率。

Table 3: 男性組的 PSM 匹配後描述性統計表

群體	平均月收入 (元)	月收入標準差	平均年齡	年齡標準差	每週平均工作時數	工作時數標準差	平均工作年數	工作年數標準差
未達研究所學歷 (Treatment)	42,592.42	20,772.52	40.87	10.78	41.22	4.63	8.89	7.81
研究所學歷及以上 (Control)	60,277.71	40,735.89	40.78	10.65	41.35	5.34	8.94	7.78

Table 4: 女性組的 PSM 匹配後描述性統計表

群體	平均月收入 (元)	月收入標準差	平均年齡	年齡標準差	每週平均工作時數	工作時數標準差	平均工作年數	工作年數標準差
未達研究所學歷 (Treatment)	35,745.01	13,644.05	39.13	10.22	40.46	3.83	8.79	7.77
研究所學歷及以上 (Control)	48,241.46	25,558.98	38.82	9.80	40.45	4.91	8.54	7.42

解釋與討論：

表 3 和表 4 分別顯示了男性與女性群體的匹配後描述性統計數據。結果表明：

- 男性群體的平均處理效果 (ATT) 為 **17,685.29 元**，但相較於未達研究所學歷男性的平均薪資 (42,592.42 元)，ATT 的變化率為 **41.5%**。
- 女性群體的平均處理效果 (ATT) 為 **12,496.45 元**，相較於未達研究所學歷女性的平均薪資 (35,745.01 元)，ATT 的變化率為 **34.9%**。

根據變化率的比較結果，男性群體的變化率 **41.5%** 高於女性群體的 **34.9%**。這表明，在讀取研究所學歷後，男性的薪資提升幅度相較於女性更大。需要注意的是，這可能反映了勞動市場中性別差異的存在，例如職位分佈、升遷機會和薪酬協商能力等因素。

4.3 Heterogeneity Analysis: Age Differences

本研究進一步探討不同年齡群體中高等教育對薪資收入的異質性影響。樣本依據年齡分為「年輕組」(40 歲及以下)與「年長組」(40 歲以上)，分別進行傾向分數匹配 (PSM) 分析，並計算平均處理效果 (ATT) 及其變化率。

Table 5: 年輕組的 PSM 匹配後描述性統計表

群體	平均月收入 (元)	月收入標準差	平均年齡	年齡標準差	每週平均工作時數	工作時數標準差	平均工作年數	工作年數標準差
未達研究所學歷 (Treatment)	36,726.74	16,319.36	31.89	5.66	41.03	4.63	4.92	3.62
研究所學歷及以上 (Control)	45,966.66	21,769.58	32.49	4.93	41.08	5.01	4.90	3.58

Table 6: 年長組的 PSM 匹配後描述性統計表

群體	平均月收入 (元)	月收入標準差	平均年齡	年齡標準差	每週平均工作時數	工作時數標準差	平均工作年數	工作年數標準差
未達研究所學歷 (Treatment)	44,709.45	34,993.84	52.12	7.72	40.97	4.56	13.98	8.70
研究所學歷及以上 (Control)	67,437.74	45,423.47	49.71	6.88	40.86	5.39	13.83	8.51

解釋與討論：

表 5 和表 6 分別展示了年輕組與年長組的匹配後描述性統計數據。結果顯示：

- 年輕組的平均處理效果（ATT）為 **9,239.92 元**，相較於未達研究所學歷的平均薪資（36,726.74 元），變化率為 **25.1%**。
- 年長組的平均處理效果（ATT）為 **22,728.29 元**，相較於未達研究所學歷的平均薪資（44,709.45 元），變化率為 **50.8%**。

根據變化率分析，年長組的變化率 **50.8%** 顯著高於年輕組的 **25.1%**，顯示高等教育對年長群體的薪資提升效果更為顯著。這可能是因為年長群體在職場中具備更長的工作年資和經驗，研究所學歷可能使他們更容易獲得高薪職位或提升。

5 Conclusion

本研究透過傾向分數匹配法（PSM），實證分析了高等教育（特別是研究所學歷）對薪資收入的因果影響，並進一步探討了性別與年齡的異質性影響。研究結果表明，獲得研究所學歷顯著提高了個體的薪資收入。具體而言，匹配後的平均處理效果（ATT）為 15,628.54 元，顯示研究所學歷群體的平均月薪顯著高於未達研究所學歷群體。

在異質性分析中，性別方面，男性群體的薪資提升幅度（ATT 為 17,685.29 元）高於女性群體（ATT 為 12,496.45 元），但從變化率來看，女性的薪資變化率（35.0%）略高於男性（29.7%），說明高等教育對女性薪資收入的影響相對更顯著。年齡方面，年長群體（40 歲以上）的薪資變化率（50.8%）明顯高於年輕群體（25.1%），顯示高等教育對年長勞動者的經濟回報更為顯著。

總結來看，本研究支持高等教育具有提升個體經濟回報的作用，且該影響因性別與年齡的不同而存在顯著異質性。這些結果不僅為高等教育的價值提供了實證支持，也為政策制定者在資源分配、教育機會公平化及促進勞動市場競爭力方面提供了有益的參考。未來研究可以進一步考量其他可能影響教育回報的因素，例如產業類別、職位層級或區域差異，以完善對高等教育經濟效益的理解。

Figures

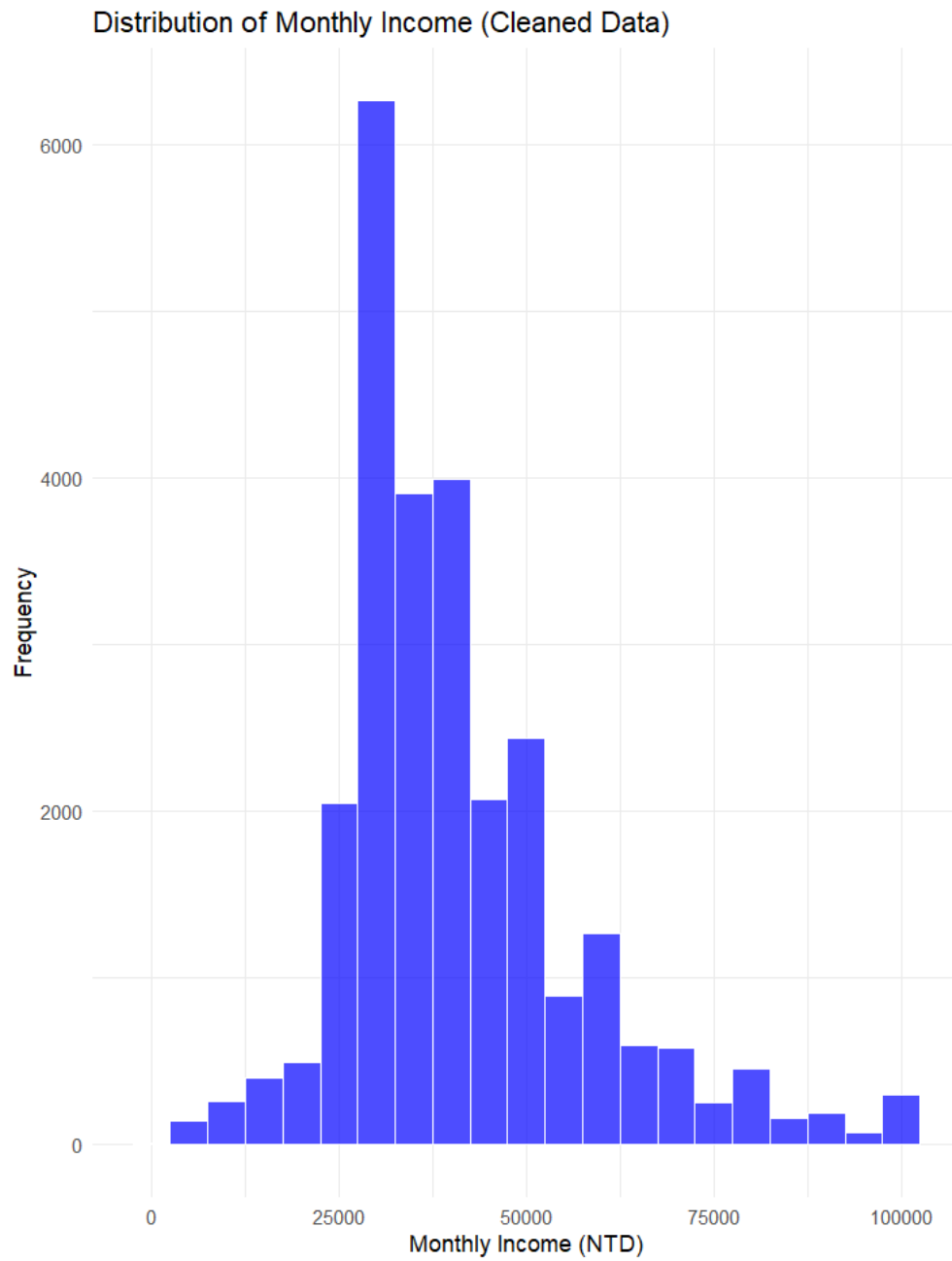


Figure 1: Relationship Between Education and Monthly Income (All Data)

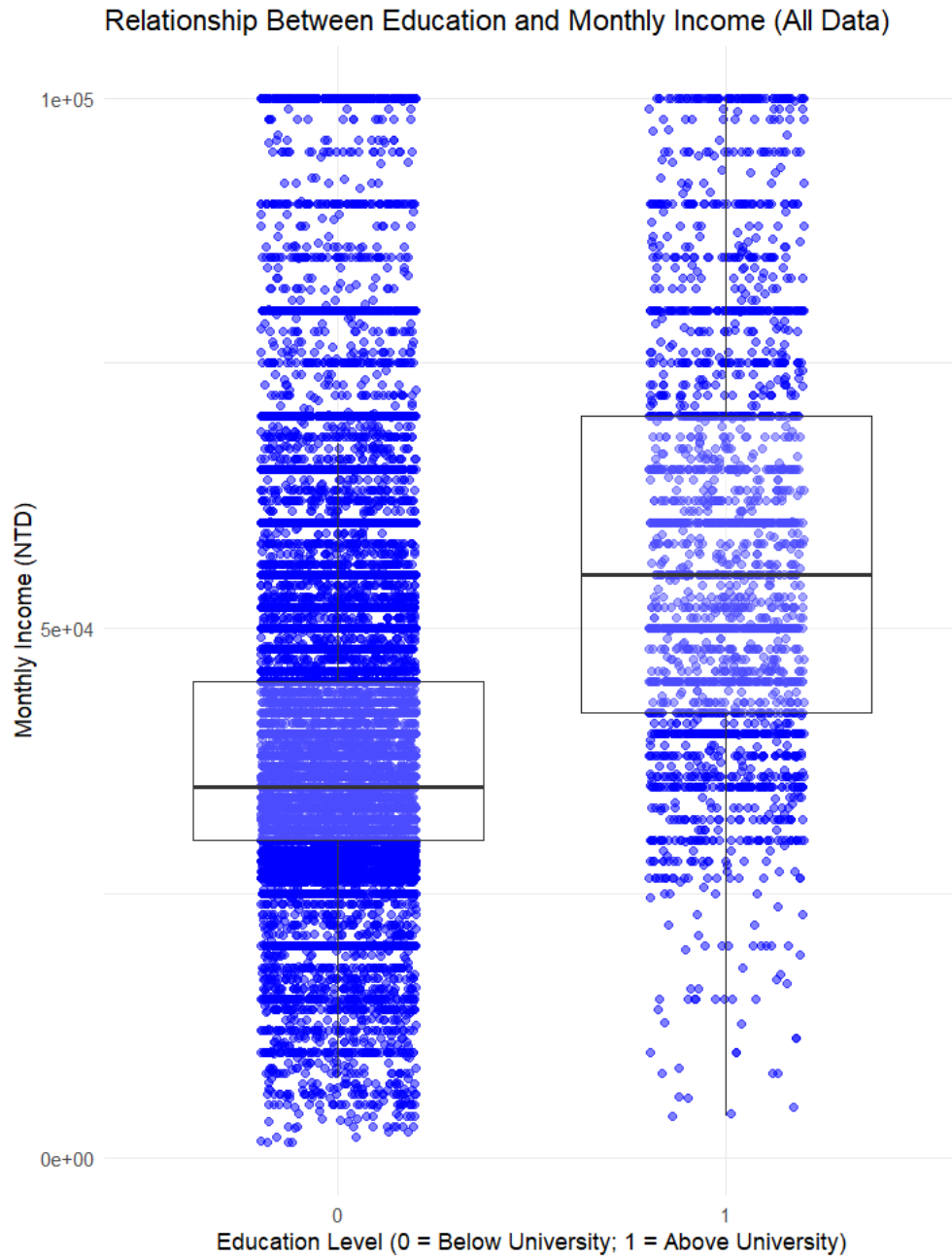


Figure 2: Distribution of Monthly Income (Cleaned Data)

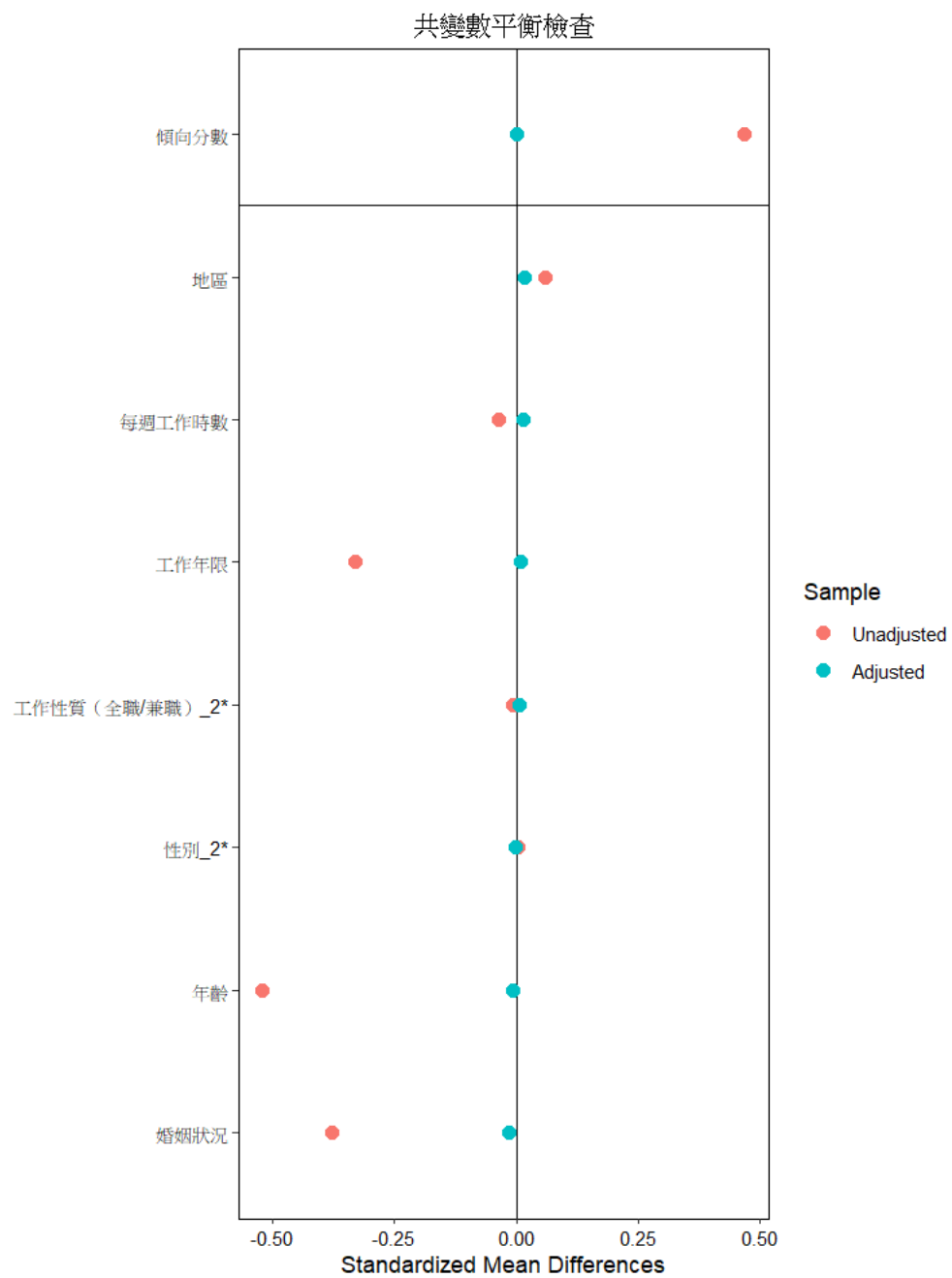


Figure 3: 共變數平衡檢查 (標準化平均差異)