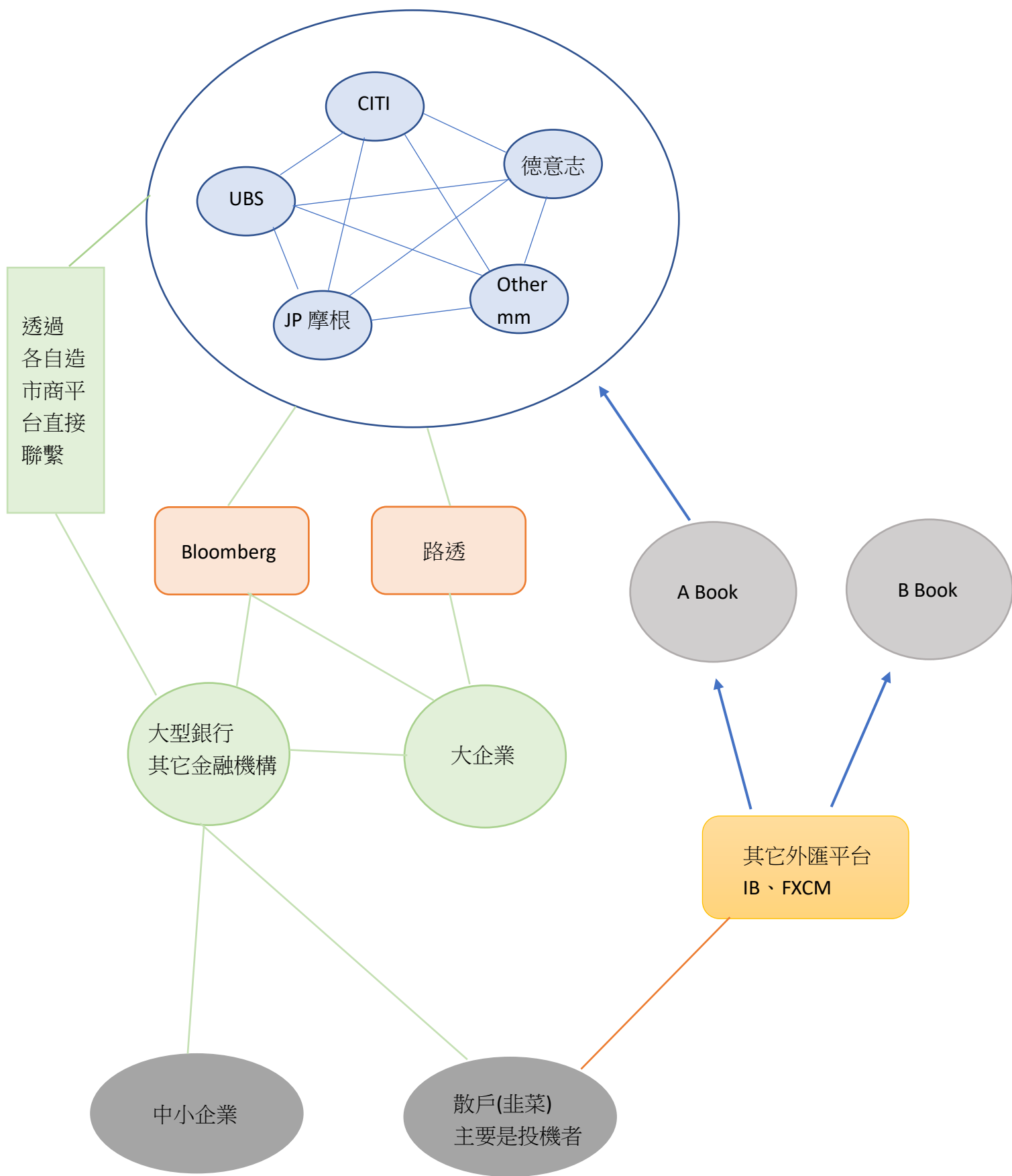


外匯市場架構



關於研究方向

1. 機器學習

目前主要學術論文都是透過歷史價格資料訓練出未來走勢¹。

主觀想法：影響匯率的因素很多，只透過歷史價格資料訓練出來的模型往往在 out-sample test 表現不好，部分模型之所以有效可能是針對部分貨幣對找到動量趨勢的 pattern，意味著使用的演算法可能只是跟著動量走，並不是本身演算法很有效。類似經驗是我和同學討論計量分析專題發現 piecewise linear regression 表現比任何機器學習模型要好，也有實證研究發現價格資料所隱含的資訊只符合 AR(1)。另外，匯率市場雖然長期而言技術分析有效²，但也可能只是所有交易員都在使用所以變成有效，一些指標在股票市場比較明顯³。

研究匯率走勢還有個困難點就是匯率分離迷思，就是透過總體數據很難有效預測短期走勢。

The exchange rate disconnect puzzle

- Studies finds that short-run exchange rates do not consistent with the IRP.
- Underlying fundamentals such as interest rates, inflation rates, and output don' t explain the short-term volatility in exchange rates.

2. 波動度研究

以下分兩部份討論

A. 波動度預測

透過機器學習

主觀想法：許多研究是透過市場資料做模型訓練，或是建模相關性預測，不同幣別的波動性可以體現在實際市場資料。

透過 market sentiment

主觀想法：透過主流外匯論壇如 Forex Factory、DailyFX、MT5 Forum，裡面許多交易員會主觀表達自己的看法，透過爬蟲擷取資料建立並建立模型。

類似研究有 [News implied volatility and disaster concern](#)。他透過雜誌標題並透過 SVM 建立 NVIX(based on news 的波動度指數)，並發現在預測 VIX 統計上顯著。

我和同學之前有參加南山研究計畫(實務課)，透過爬取 FOMC 文章，建立利率預測。主觀認為即使建立了有預測能力的模型可能也無法實際交易。

Implied volatility and Realized volatility

一般而言 Implied volatility 具有 forward-looking 的特性。而 Realized volatility 是 backward-looking。然而，我看許多相關研究並沒有結論確定是哪種波動度有比較好的預測力。這部份研究主要圍繞在時間序列分析。

¹ [A Review on Recent Advancements in FOREX Currency Prediction](#) 整理了所有目前關於外匯預測的研究方向。

² 與台新和一銀的外匯交易員討論過，他們主要也是透過技術指標去判斷。

³ [The 52-week high and momentum investing](#) The Journal of Finance, 2004

B. 波動度交易策略

均值迴歸

研究不同幣別的均值迴歸特性。可以衍生出如 Strangle、Straddle、calendar spread 策略。

Implied volatility and Realized volatility

不管是波動率的方向性策略還是套利性策略，一定程度上都是帶有方向性地賭未來歷史或隱含波動率。那麼最簡單的自然有以下三種：

1、Long/Short Gamma (Bet 未來的實際波動率)

2、Long/Short Vega (Bet 未來的隱含波動率)

3、Long Gamma Short Vega / Short Gamma Long Vega (日曆價差 Calendar Spread 策略：同時 Bet 實際波動率和隱含波動率) 該策略可以用不同月份的 ATM straddle (或 single leg) 合約來構建。

3. 市場微結構(Microstructure of Foreign Exchange)

參考資料：[FX Trading and the Exchange Rate Disconnect Puzzle Martin D. D. Evans\(2018\)](#)

他透過研究認為匯率分離迷思不存在反而「quite well “connected” to interest differentials.」。

因為 inner tier⁴的 trading flow，主要是 sophisticated trading entities。某種程度上他們的交易行為會隱含總體經濟資訊。日內匯差主要體現在兩國間的短率差異和流動性風險。

重要研究結果：

1. interest differentials appear to have no forecasting power.

2. liquidity risk 有預測能力。

透過 30 秒的高頻資料(limit order book and price dynamics)，構建 VAR 模型來預測匯率。

⁴ The FX dealers working at major banks trade with each other in the inner tier, while trades between dealers and end-users take place in the outer tier. And sophisticated trading entities have information advantages.

使用文字探勘針對波動度預測

參考論文為：News Implied Volatility and Disaster Concerns

本篇作者想要在歷史的宏觀角度上研究投資人對未來的不確定性是一個什麼樣的關係，並想辦法去測出來。方法是對 VIX 和 Wall Street Journal 封面上的字出現的頻率去做迴歸。這樣模型做出來的 volatility 我們稱為 News-implied volatility index (NVIX)。那至於為何作者是選擇 vix 是因為 merton 1973 說 options implied volatility 可以做為股價報酬的參考指標。報章雜誌的用字遣詞可以準確反映出普通投資人的擔憂，理由是名聲好的新聞媒體，在用字遣詞上能夠比較好的去反應真實事件。

數據處理部分可以參考下圖。下面就是波動度迴歸式， \mathbf{x} 是一個向量。以 financial 為例， \mathbf{x} 的第一個 entry 的分子是 Financial 在這個月出現幾次，然後分母就是這個月出現了多少字。

很顯然，如果我們直接對這迴歸式做 OLS 會有 overfitting 的問題出現。因為我們的觀察值很少。但是維度會很大(就是 financial 相對於母體出現的次數差太多)。

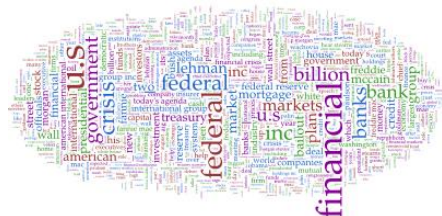
front-page titles and abstracts of the Wall Street Journal, 1890-2009

Date	Title	Abstract
2008-09-16	AIG Faces Cash Crisis As Stock Dives 61%	American International Group Inc. was facing a severe cash ...
2008-09-16	AIG, Lehman Shock Hits World Markets ...	The convulsions in the U.S. financial system sent markets ...
2008-09-16	Business and Finance	Central banks around the world pumped cash into money ...
2008-09-16	Keeping Their Powder Dry: Draft Boards ...	The Selective Service System has the awkward task of ...
2008-09-16	Old-School Banks Emerge Atop New ...	Banks are heading "back to basics - to, if you like, the core ...

Tue 30 Sep 2008

$$VIX_t - \overline{VIX} = w_0 + \mathbf{w} \cdot \mathbf{x}_t + v_t$$

42.75



Raw word frequencies



Weighted word frequencies

處理 High dimensionality 的方法：

1. 帶入懲罰項的迴歸：透過在模型中帶入懲罰項(Penalty Term)，減少高維度帶來的誤差。如 Ridge Regression、Lasso Regression
2. 建立在機器學習上的迴歸模型：如 Decision Tree Regression 或是 Support Vector Regression，在很多時候這些模型的表現都遠勝過前兩種迴歸模型

隱含波動與資訊

根據定義，任何隨機變數 X_t 可以被分解為兩個部分

- Expected value conditional on an information set Φ_{t-m} available m periods earlier
- Zero mean forecast error ϵ_t

$$X_t = E[X_t | \Phi_{t-m}] + \epsilon_t \text{ Where } E[\epsilon_t | \Phi_{t-m}] = 0$$

有了以上結果，我們可以有迴歸關係式如下

$$X_t = \alpha + \beta F(\Phi_{t-m}) + \epsilon_t$$

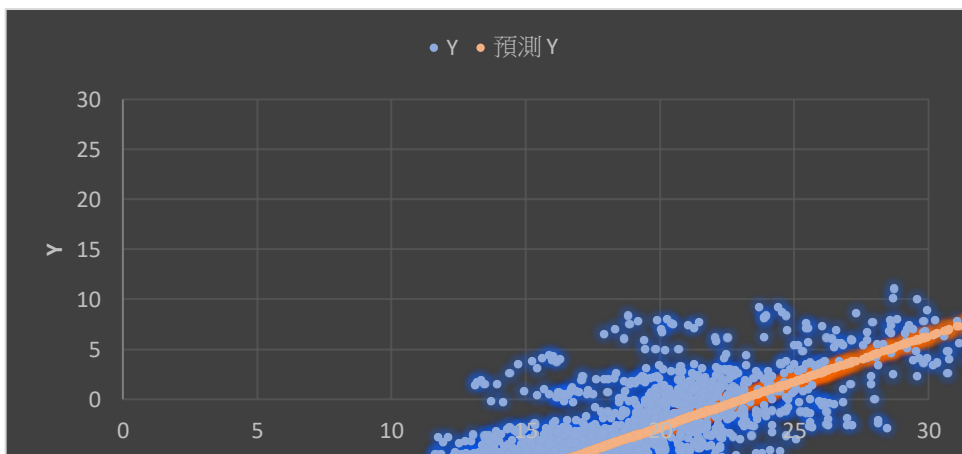
對應到我們波動度分析我們可以有以下結果

$$Realized_t = \alpha + \beta Implied_t + \epsilon_t$$

各幣別回測結果如下

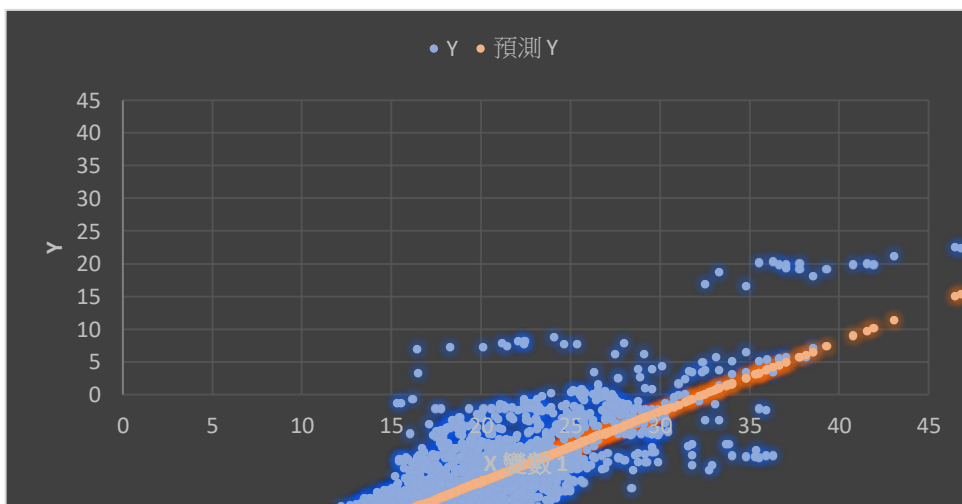
EUR/USD（高度相關）

迴歸統計	
R 的倍數	0.883063
R 平方	0.7798
調整的 R 平方	0.779742
標準誤	1.651304
觀察值個數	3794



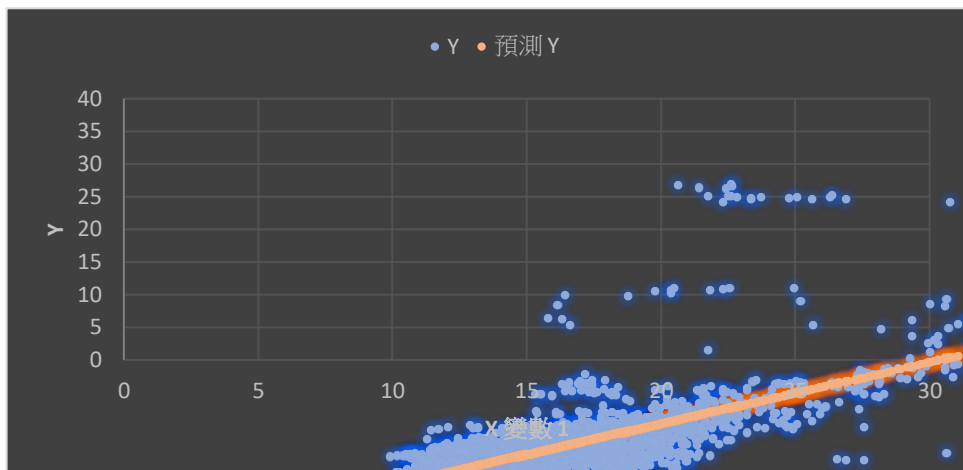
USD/JPY（中度相關）

迴歸統計	
R 的倍數	0.835544
R 平方	0.698133
調整的 R 平方	0.698053
標準誤	2.353937
觀察值個數	3794



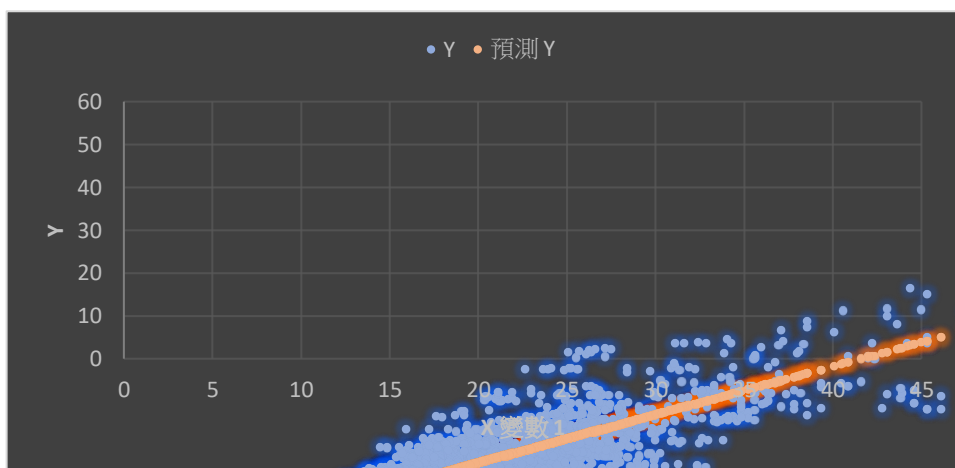
GBP/USD（中度相關）

迴歸統計	
R 的倍數	0.777714
R 平方	0.604839
調整的 R 平方	0.604735
標準誤	2.51503
觀察值個數	3793



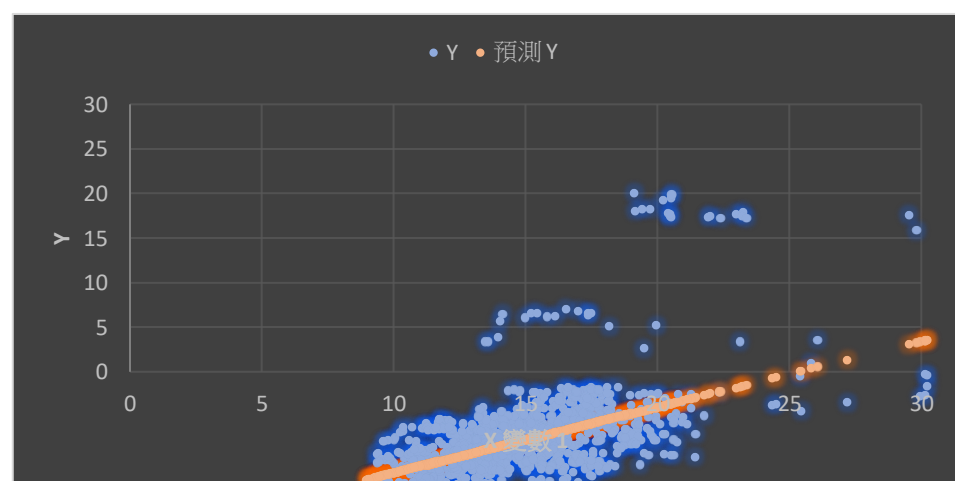
EUR/JPY（高度相關）

迴歸統計	
R 的倍數	0.869902
R 平方	0.75673
調整的 R 平方	0.756666
標準誤	2.850349
觀察值個數	3794



EUR/GBP（低度相關）

迴歸統計	
R 的倍數	0.562603
R 平方	0.316522
調整的 R 平方	0.315988
標準誤	2.722503
觀察值個數	1284



可以觀察出長期來看，隱含波動度和歷史波動有相關性。先進經濟體和交易量大的幣別相關性會更高。代表隱含波動度確實隱含了對市場預期的資訊從而可以衍生出相關交易策略。