Revisiting the Capital Asset Pricing Model: Evidence from CAC 40 Firms

Vincent Desmeulles†

Abstract

Recently from analyses on different databases, the tent-shape of the distribution of firm growth rates has emerged as a robust and universal characteristic of the time evolution of corporations. We add new evidence on this topic and we present a new stochastic model that, under rather general assumptions, provides a robust explanation for the observed regularity.

Keywords: CPAM, Panel Data, CAC 40

†vincent.desmeulles@etu.univ-paris1.fr, Master's 2nd year Applied Economics, Student ID Number: 11906510.

1. Introduction

The *Capital Asset Pricing Model* (CPAM) is one of the most recognized models in financial econometrics. It states the existence of a fundamental relationship between an asset's expected return (mean) and its risk (variance). This relationship is formalized through a β coefficient, which quantifies an asset's sensibility to market movement, which the theory purports to be linear. Since its inception, the model has been used up to this day due to its low information requirement and simplicity. However, the over-simplification of the model bears costs on its reliability, since it relies on restrictive assumptions regarding the behavior of economic agents. The model has been the subject of a large body of theoretical and empirical research, sometimes creating heated debates concerning its practical usefulness. Despite these difficulties, it is still widely used in the field.

The model has been subject to extensive empirical testing. It found early success in the 1970s with *Black et al. (1972)* and *Fama and MacBeth (1973)* finding a linear relationship between β and average returns, which seemed to indicate a validation of the CAPM. This line of success wasn't going to last long, and subsequent studies highlighted the shortcomings of the model. The work of *Banz (1981)* showed that smaller firms tends to have a higher average return than predicted by the model, which challenged the notion that the sensibility to market movement was the only reliable determinant of expected returns, *Jegadeesh and Titman (1993)* showed that stocks past performances explains a large part of the return, independently of β. One of the most famous of these critiques was the several contributions of Fama and French *(1992)*, *(1993)*, *(1996)*, who developed their own eponymous *three-factors* model which incorporated size and valuation as factors into their analysis, with findings suggesting a superior explanatory power of these factors compared to β. Despite these theoretical advances, the CAPM is still used to make empirical analysis of asset return by researchers, highlighting some explanatory power.

This paper attempts to verify whether the CAPM can be considered a valid estimator for the performances of large firms in the French financial market, namely, of CAC 40 firms. To achieve this, a single model is employed, using a panel data of monthly stock returns from 2009 to 2024 and the yield of French 13-week treasury bonds. An estimation of β is provided for each individual firm. Additionally, monthly macroeconomic indicators are incorporated into the analysis to evaluate their potential effects on stock returns.

The next section introduces the dataset and addresses potential econometric issues related to the sample. The third section outlines the two econometric models employed in the analysis. In the fourth section, the results of the study are presented. Finally, the fifth section explores the limitations and practical applications of the CAPM to the French economy.

2. Overview

For the purpose of this article, a dataset was constructed using Yahoo Finance's API to collect monthly stock price of CAC 40 firms. The database follows the trajectory of 40 firms listed on the index as of September 30th 2024, totaling 6817 observations over a time frame ranging from January 2009 up to January 2024. The database was converted to logarithmic returns to account for the large changes in value over the period. This also allowed the observations to follow a stationary trajectory and avoid autoregressive residuals.

For getting the risk-free rate, data on French short-term (13 weeks maturity) treasury bonds[1]

[1] More specifically, the bonds used for getting the risk-free rate were instruments called "Bon du Trésor à taux fixe et à intérêt précompté".

Table 1: Variables of interest in the first model.

Variable	Description
Equity Return	Monthly evolution of the value of an equity plus dividends.
Market Return	Monthly evolution of the value of the CAC 40 index.
Risk-Free Return	Yield of 13 weeks-maturity French treasury bond.

Table 2: Supplementary variables in the second model.

Variable	Description
Sector	Industry Classification Benchmark
Inflation	Monthly evolution of the consumer price index.
IPI	Monthly evolution of the industrial production index.
USD/EUR	Monthly evolution of the exchange rate of the euro against the dollar.

was collected from the Banque de France website. Their rate were also converted to logarithmic return for consistency.

Data on Consumer Price Index (CPI) was collected from the INSEE website, as well as the monthly Industrial Production Index (IPI). The IPI was chosen instead of GDP growth because it is usually reported earlier (which gives less time for market to price-in the information) and because it is reported monthly instead of quarterly.

Finally, the exchange rate from euro to US dollar was taken from the ECB website. Because most of the firms on the CAC 40 are deeply internationalized, the effect of the evolution of exchange rate could have a large effect on the ability of CAC 40 to export their product abroad, which could translate into their returns.

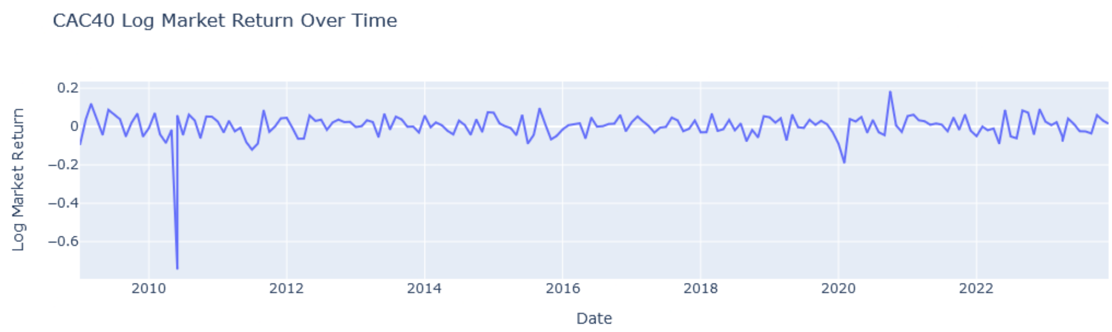

Figure 1: Log market return over time.

Since this analysis focus exclusively on large internationalized firms, it can be assumed that the effects observed for small firms by *Banz (1981)* will be negligible and that omited variable bias will be small. The process of transforming the dataset into log values also made the model robust against autocorrelation.

3. Empirical Model

Since this study purports to analyze the return of firms with heterogenous characteristics endogenous with firm-specific equity return over a long time frame, a panel data model with fixed-effect was chosen. It is of the form:

$$r_{i,t} - rm_t = \alpha_i + \beta_m(rm_t - rf_t) + \beta_2 inf_t + \beta_3 ipi_t + \beta_4 fx_t + \beta_5 s_i \varepsilon_{i,t}$$

With :

- $r_{i,t}$: Log return of asset i at time t.
- rm_t: Log market return at time t.
- fr_t: Log risk-free rate at time t.
- inf_t: Log inflation at time t.
- ipi_t: Log industrial production index growth at time t.
- fx_t: Log EUR/USD exchange rate growth at time t.
- α_i: Intercept term specific to asset i.
- β_m: Sensitivity of asset i to market return.
- β_2: Sensitivity of asset i to inflation.
- β_3: Sensitivity of asset i to industrial production index growth.
- β_4: Sensitivity of asset i to EUR/USD exchange rate growth.
- $\varepsilon_{i,t}$: Idiosyncratic error term of asset i at time t.

Some issues must be addressed regarding the validity of any results obtained with this model. Since this analysis is founded on a panel data ordinary least square model, the following conditions must be respected:

- Zero Conditional Mean Assumption: $\mathrm{E}(\varepsilon_{i,t}|rm_t, rf_t, \ldots, fx_t) = 0$. That is, the conditional distribution of the residuals must be equal to zero for any value of the independent variables. This could be linked to measurement error of the observations, which is unlikely, or to omited variable biais. This last point is more likely, given that log returns seem to have a non-normal distribution (see figure 1). The introduction of variables such as IPI growth could have a decisive effect in neutralizing effects endogenous with the main CAPM equation.
- The sample of CAC 40 firms is identically and independently distributed: It is common knowledge that stock prices are autocorrelated with their previous value, which is why the logarithmic transformation was useful to neutralize this effect. CAC 40 firms are also heterogenous, which is why the study also uses fixed-effect to account for this. Finally, macroeconomic shocks are also accounted for, preventing any biais in the distribution for a given time t.

- Large outliers are unlikely: As seen in figure 1, the distribution of returns, while stationnary, seems to exhibit a non-normal distribution. This paper's use of the OLS model lies on the assumption that the introduction of macroeconomic variable could resoulve this problem.

- No perfect collinearity: This issue is non-existent in this model, in part because no dummy variables are used.

4. Results

The results derived from the model exhibit significant weaknesses, characterized by a general lack of robust correlations among the variables under consideration. Almost none of the predictors demonstrate a meaningful or definitive relationship with the dependent variable, with the notable exception of the logarithm of the US dollar to euro exchange rate. This variable alone emerges with a statistically significant correlation, with a p-value of 0.0004.

Despite this isolated significant finding, the overall explanatory power of the model remains exceptionally low. The coefficient of determination indicate that the model accounts for only 0.19% of the total squared variability in the dependent variable. This figure highlights the model's inability to capture meaningful patterns or relationships within the dataset, suggesting that most variability remains unexplained. While low R squarred are not uncommon in financial models, this is still very low compared to other models.

Table 3: Regression Results

Variable	**OLS Estimate**	**P-value**
Market Excess Return	0.0052	0.7513
Log Inflation	-0.0306	0.8939
Log IPI	-0.0011	0.7423
Log USD/EUR	0.1495	0.0004
Intercept	0.0087	0.0000

Model Summary:
F-Test P-value: 0.0304
R-squared: 0.0019

5. Discussion

The results of this study highlight the inability of the Capital Asset Pricing Model to significantly explain any the returns of the CAC 40 firms. One possible explanation for the weak explanatory power of the CAPM model lies in the presence of non-linear residuals. The model assumes linear relationships between the independent variables and the dependent variable. However, financial markets often exhibit non-linear relations that cannot be fully explained by linear models such as the one employed here.

Another takeaway from this study is that French CAC40 firms may reflect the efficient market hypothesis, which states that markets already contain all the informations in prices. Therefore, models such as CAPM wouldn't be able to consistently predict returns beyond chance. The lack of robust relationship may be a reflection of this principle.

This study seems to reflect a more modern academic consensus that CAPM doesn't hold great explanatory power. And while the results may appear weak, they show critical areas for improvement and the need for more advanced methodologies to better understand financial market dynamics beyond usual linear models.

References

- Banz, R.W. (1981) 'The Relationship Between Return and Market Value of Common Stock', Journal of Financial Economics, 9(1), 3–18.

- Black, F., Jensen, M.C. and Scholes, M. (1972) 'The Capital Asset Pricing Model: Some Empirical Tests', in M.C. Jensen (ed.), Studies in the Theory of Capital Markets, Praeger, New York.

- Fama, E.F. and French, K.R. (1993) 'Common Risk Factors in the Returns on Stocks and Bonds', Journal of Financial Economics, 33, 3–56.

- Fama, E.F. and French, K.R. (1996) 'Multifactor Explanations of Asset Pricing Anomalies', Journal of Finance, 47, 426–465.

- Fama, E.F. and MacBeth, J. (1973) 'Risk, Return and Equilibrium: Empirical Tests', Journal of Political Economy, 81(3), 607–636.

- Jegadeesh, N. and Titman, S. (1993) 'Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency', Journal of Finance, 48(1), 56–91.

Code

```
import yfinance as yf
import pandas as pd
import numpy as np
import time
import plotly.express as px
import xml.etree.ElementTree as ET

def api_request(x, a, b, c):
    x = yf.Ticker(x)
    y = x.history(start=a, end=b, interval=c)
    return y

def get_stocks(x, a, b, c, d):
    l = len(x)
    y = []
    for i, row in x.iterrows():
        n = row["Name"]
        s = row["Sector"]
        j = row["Ticker"]
        print(f"[{i + 1}/{l}] {j}...", end="")
        z = api_request(j, a, b, c)
        z = z.reset_index()
        z.insert(0, "Ticker", j)
        z.insert(1, "Name", n)
        z.insert(2, "Sector", s)
        y.append(z)
        print(" X")
        time.sleep(d)
    y = pd.concat(y, ignore_index=True)
    y["Date"] = pd.to_datetime(y["Date"], utc=True).dt.strftime('%Y-%m')
    return y

tickers = pd.read_csv("data/raw/CAC40.csv", sep=";")
cac = get_stocks(tickers, "2008-12-01", "2024-02-01", "1mo", 1)

cac

bk = api_request("^FCHI", "2008-12-01", "2024-02-01", "1mo")
k = pd.DataFrame()
k["Benchmark"] = bk["Close"]
k = k.reset_index()
k["Date"] = pd.to_datetime(k["Date"], utc=True).dt.strftime('%Y-%m')

cac40 = pd.merge(cac, k, on="Date", how="left")
del bk
cac40
```

```
tmb = pd.read_csv("https://webstat.banque-france.fr/export/csv-columns/fr/selection/5
tmb = tmb[~tmb.isin(["-"]).any(axis=1)]
tmb = tmb[[0, 1]]
tmb.rename(columns={0: "Date", 1: "Bond Yield"}, inplace=True)
tmb["Date"] = pd.to_datetime(tmb["Date"]).dt.strftime('%Y-%m')
tmb["Bond Yield"] = tmb["Bond Yield"].str.replace(',', '.').astype(float)

tmb

data = pd.merge(cac40, tmb, on="Date", how="left")
del tmb

data = data.drop(columns=["Open", "High", "Low", "Volume", "Stock Splits"])
data.rename(columns={"Close": "Price"}, inplace=True)
data

inf = pd.read_csv("data/raw/IPC.csv", skiprows=4, sep=";", header=None)
inf = inf[[0, 1]]
inf.rename(columns={0: "Date", 1: "IPC"}, inplace=True)
inf["Date"] = pd.to_datetime(inf["Date"]).dt.strftime('%Y-%m')
inf = inf.sort_values(by="Date")

inf["Log Inflation"] = np.log(inf["IPC"] / inf["IPC"].shift(1))
inf = inf.drop(columns=["IPC"])
inf

data = pd.merge(data, inf, on="Date", how="left")
del inf

data

ibpi = pd.read_csv("data/raw/IBPI.csv", skiprows=4, sep=";", header=None)
ibpi = ibpi[[0, 1]]
ibpi.rename(columns={0: "Date", 1: "IPI"}, inplace=True)
ibpi["Date"] = pd.to_datetime(ibpi["Date"]).dt.strftime('%Y-%m')
ibpi = ibpi.sort_values(by="Date")

ibpi["IPI"] = np.log(ibpi["IPI"] / ibpi["IPI"].shift(1))
ibpi.rename(columns={"IPI": "Log IPI"}, inplace=True)

ibpi

data = pd.merge(data, ibpi, on="Date", how="left")
del ibpi

data
```

```
usd = ET.parse("data/raw/USD.xml")

# Get the root element
r = usd.getroot()

n = {
    "message": "http://www.SDMX.org/resources/SDMXML/schemas/v2_0/message",
    "exr": "http://www.ecb.europa.eu/vocabulary/stats/exr/1"
}

ob = r.find(".//exr:DataSet", n).find(".//exr:Series", n).findall("exr:Obs", n)

usdeur = []
for i in ob:
    a = i.get("TIME_PERIOD")
    b = i.get("OBS_VALUE")
    usdeur.append({"TIME_PERIOD": a, "OBS_VALUE": float(b)})

usd = pd.DataFrame(usdeur)
del usdeur
del a
del b
del ob
del n
del r

usd.rename(columns={"TIME_PERIOD": "Date", "OBS_VALUE": "USD/EUR"}, inplace=True)

usd

usd["Date"] = pd.to_datetime(usd["Date"])
usd["Date"] = usd["Date"].dt.to_period("M")
usd = usd.groupby("Date", as_index=False)['USD/EUR'].mean()
usd["Date"] = usd["Date"].astype(str)
usd["Date"] = pd.to_datetime(usd["Date"]).dt.strftime('%Y-%m')
usd = usd.sort_values(by="Date")

usd["USD/EUR"] = np.log(usd["USD/EUR"] / usd["USD/EUR"].shift(1))
usd.rename(columns={"USD/EUR": "Log USD/EUR"}, inplace=True)
usd

data = pd.merge(data, usd, on="Date", how="left")
del usd

data

data["Price"] = np.log((data["Price"] + data["Dividends"]) / data["Price"].shift(1))
data = data.drop(columns=["Dividends"])
```

```
data.rename(columns={"Price": "Log Return"}, inplace=True)
data

data["Benchmark"] = np.log(data["Benchmark"] / data["Benchmark"].shift(1))
data.rename(columns={"Benchmark": "Log Market Return"}, inplace=True)
data

data["Bond Yield"] = np.log(1 + data["Bond Yield"] / 100)
data.rename(columns={"Bond Yield": "Log Risk-Free Return"}, inplace=True)
data

low = pd.Timestamp("2009-01")
high = pd.Timestamp("2024-02")

data["Date"] = pd.to_datetime(data["Date"])

data = data[(data["Date"] >= low) & (data["Date"] < high)]
data

data["Date"] = pd.to_datetime(data["Date"]).dt.strftime('%Y-%m')
data = data.set_index(["Ticker", "Date"])
data

data.to_csv("data/data.csv", index=True, sep=";")

###

import pandas as pd
import numpy as np
import statsmodels.api as sm
from linearmodels.panel import PanelOLS

df = pd.read_csv("data/data.csv", sep=";")
df["Excess Return"] = df["Log Return"] - df["Log Market Return"]
df["Market Excess Return"] = df["Log Market Return"] - df["Log Risk-Free Return"]
df["Date"]= pd.to_datetime(df["Date"])
df = df.set_index(["Ticker", "Date"])

df["Intercept"] = 1

dependent = df["Excess Return"]
independent = df[[
    "Market Excess Return",
    "Log Inflation",
    "Log IPI",
    "Log USD/EUR",
    "Intercept"
]]
```

```
model = PanelOLS(dependent, independent, entity_effects=True)
results = model.fit()

print(results.summary)
```