

Cash Estimation Modeling

Introduction:

Understanding the pivotal role of cash holdings in shaping a firm's future performance and market dynamics is crucial in modern financial analysis. Our analysis aims to estimate optimal cash levels for individual firms. Through rigorous statistical modeling, we seek to ascertain whether companies hold surplus or inadequate cash compared to their estimated targets.

Our methodology involves meticulous data preprocessing, cleaning, and merging to ensure the integrity and reliability of our analysis. By scrutinizing key variables and their correlations, we have crafted predictive models using SAS, shedding light on the significance of various factors influencing cash reserves within firms.

The outcomes of our analysis offer actionable insights for firms, guiding strategies to optimize cash flow, refine investment decisions, and capitalize on managerial expertise and sales growth for sustained financial health. However, these recommendations should be approached with a nuanced understanding of industry-specific contexts.

This collaborative endeavor serves as a guiding reference for firms seeking to fortify their financial foundations and navigate the complex role between cash management and long-term performance in today's dynamic business landscape.

Dataset Description:

Dataset:

The data used for our analysis is sourced from the WRDS portal, specifically the Compustat - Fundamentals Annual section for the North America region, covering the period from 1960 to 2022.

Source:

<https://wrds-www.wharton.upenn.edu/pages/get-data/compustat-capital-iq-standard-poors/compustat/north-america-daily/fundamentals-annual/>

Data Preprocessing:

Filter criteria:

- Time period is filtered from 1975 to 2022.
- Firm-years have non-missing data for cash and cash equivalents
- Positive total assets
- Positive net sales
- Non-missing common shareholders' equity
- Excluded firms with a year-end stock price of less than \$2.
- Excluded the utility firms (SIC codes 40-49) and financial firms (SIC codes 60-67) due to their unique additional regulatory constraints on the cash and investments.

Cleaning:

To ensure data integrity, we executed systematic cleaning and refining procedures for each variable. This included sorting the data by 'gvkey' and 'nyear', removing duplicates, and handling missing values.

Data Transformation:

To enhance data processing in our analysis, we initiate a transformation on the 'gvkey' variable, originally coded as a character. The conversion involves multiplying 'gvkey' by 1, resulting in the creation of a new numeric variable labeled as 'gvkey2.' Additionally, we created a new variable 'nyear' derived from the 'datadate' variable to represent the year corresponding to each 'datadate' entry. Another refinement involved the creation of 'sic2' by dividing the existing 'sic' variable by 100 and extracting the integer part of the result. The outcome is a 'sic2' variable capturing the first two digit Standard Industrial Classification (SIC) Codes. This transformation process enhances the granularity of our dataset, facilitating more insightful and streamlined analyses.

Merging:

The data merging process in this study was a critical step in integrating multiple datasets to create a comprehensive analytical dataset for our analysis. The merging was coordinated based on unique identifiers, 'gvkey' and 'nyear', ensuring a cohesive combination of relevant information across different years and firms. The merged dataset serves as the foundation for subsequent analyses, providing a holistic view of firm characteristics over time and enabling robust empirical investigations into the factors influencing financial performance and corporate decisions.

Winsorization:

To ensure data reliability and consistency, we addressed outliers from the merged dataset by calculating percentile breakpoints to manage extreme values and winsorization was applied to cash and all continuous variables (except for dummies) at the 1st and 99th percentiles. This approach mitigates the impact of outliers, contributing to the robustness of the dataset for subsequent analyses.

Descriptive statistics:

Table 1:
Descriptive statistics for the cash estimation model variables:

Variable	N	Mean	Standard Deviation	25th Percentile	Median	75th Percentile
Cash	93,289	0.352	1.280	0.023	0.080	0.241
Firm Size	93,289	6.054	2.155	4.483	5.967	7.533
Net Working Capital	93,289	0.067	0.498	-0.031	0.083	0.226
Capital Expenditures	93,289	0.077	0.086	0.026	0.050	0.093
Industry Sigma	93,289	20.488	42.486	1.677	5.367	16.580
Research & Development	93,289	0.039	0.090	0.000	0.000	0.038
Dividend Dummy	93,289	0.393	0.393	0.000	0.000	1.000
Sales Growth	93,289	0.257	0.852	-0.008	0.092	0.256
Cash from Operations	93,289	0.023	0.552	0.026	0.087	0.147
Firm Age	93,289	2.558	0.902	1.946	2.639	3.296
Tax Burden on Foreign Income	93,289	0.169	0.550	0.000	0.000	0.000
Leverage	93,289	0.320	0.432	0.128	0.260	0.408
Managerial Ability	93,289	0.537	0.286	0.300	0.500	0.800
Interest Rates	93,289	1.026	17.045	0.003	0.020	0.126
Market To Book	93,289	1.547	3.634	0.532	0.940	1.691

The descriptive statistics reveal insightful patterns in the dataset used for our cash estimation model. The average cash level across firms is 0.352, with a notable standard deviation of 1.280, indicating considerable variation. Firms, on average, have a size of 6.054, while net working capital shows a mean of 0.067, with some companies experiencing negative values. Capital expenditures exhibit a mean of 0.077, displaying moderate investment activities. In terms of R&D expenditure, firms allocate a relatively small portion of their resources to R&D activities with a mean of 0.039. Notably, a significant portion of firms (25th percentile) report zero R&D expenditures, suggesting a diverse landscape in R&D Investment, with some firms either not engaging in such activities or investing minimally. Industry sigma, exhibits substantial diversity with a mean of 20.488. Dividend payments are prevalent, as indicated by the mean value of 0.393. Sales growth varies significantly (mean = 0.257, std dev = 0.852), and firms generate positive cash from operations with a mean of 0.023. The dataset comprises relatively young firms, with a mean age of 2.558. The managerial ability has a mean of 0.537, suggesting a moderate level of perceived competence. The dataset spans diverse interest rate environments (mean = 1.026, std dev = 17.045), and the market-to-book ratio averages 1.547, indicating a generally positive relationship between market and book values.

Correlation Matrix:

Correlation specifically measures the strength and direction of a linear relationship between two variables. In the context of descriptive statistics, correlation provides information about the degree to which two variables vary together. It is often expressed as a correlation coefficient, for example, the Pearson correlation coefficient, which quantifies the strength and direction of a linear relationship between two continuous variables.

Table 2:

Variable	Cash	Net Working Capital	Capital Expenditures	Size	Sales Growth	Cash from Operations	Research & Development	Tax Burden on Foreign Income	Firm Age	Industry Sigma	Dividend Dummy	Managerial Ability	Interest Rates	Leverage	Market to Book
Cash	1.000	-0.450 <.0001	0.126 <.0001	-0.127 <.0001	0.158 <.0001	-0.620 <.0001	0.359 <.0001	0.008 0.013	-0.154 <.0001	0.080 <.0001	-0.140 <.0001	-0.011 0.001	0.125 <.0001	0.349 <.0001	0.133 <.0001
Net Working Capital	-0.450 <.0001	1.000	-0.118 <.0001	-0.018 <.0001	-0.084 <.0001	0.418 <.0001	-0.256 <.0001	-0.002 0.524	0.065 <.0001	-0.100 <.0001	0.049 <.0001	0.036 <.0001	-0.115 <.0001	-0.508 <.0001	-0.230 <.0001
Capital Expenditures	0.126 <.0001	-0.118 <.0001	1.000	-0.108 <.0001	0.168 <.0001	-0.062 <.0001	0.051 <.0001	-0.049 <.0001	-0.216 <.0001	-0.033 <.0001	-0.095 <.0001	0.048 <.0001	0.012 0.000	0.005 0.167	0.063 <.0001
Size	-0.127 <.0001	-0.018 <.0001	-0.108 <.0001	1.000	-0.143 <.0001	0.188 <.0001	-0.214 <.0001	0.186 <.0001	0.391 <.0001	0.165 <.0001	0.434 <.0001	0.036 <.0001	-0.097 <.0001	0.020 <.0001	-0.157 <.0001
Sales Growth	0.158 <.0001	-0.084 <.0001	0.168 <.0001	-0.143 <.0001	1.000	-0.138 <.0001	0.116 <.0001	-0.030 <.0001	-0.239 <.0001	0.021 <.0001	-0.148 <.0001	0.043 <.0001	0.034 <.0001	0.052 <.0001	0.149 <.0001
Cash from Operations	-0.620 <.0001	0.418 <.0001	-0.062 <.0001	0.188 <.0001	-0.138 <.0001	1.000	-0.480 <.0001	0.068 <.0001	0.140 <.0001	-0.053 <.0001	0.133 <.0001	0.108 <.0001	-0.144 <.0001	-0.392 <.0001	-0.180 <.0001
Research & Development	0.359 <.0001	-0.256 <.0001	0.051 <.0001	-0.214 <.0001	0.116 <.0001	-0.480 <.0001	1.000	0.040 <.0001	-0.136 <.0001	0.049 <.0001	-0.205 <.0001	0.071 <.0001	0.070 <.0001	0.184 <.0001	0.229 <.0001
Tax Burden on Foreign Income	0.008 0.013	-0.002 0.524	-0.049 <.0001	0.186 <.0001	-0.030 <.0001	0.068 <.0001	0.040 <.0001	1.000	0.115 <.0001	0.064 <.0001	0.041 <.0001	0.081 <.0001	-0.013 0.000	-0.006 0.086	0.030 <.0001
Firm Age	-0.154 <.0001	0.065 <.0001	-0.216 <.0001	0.391 <.0001	-0.239 <.0001	0.140 <.0001	-0.136 <.0001	0.115 <.0001	1.000	0.045 <.0001	0.365 <.0001	0.046 <.0001	-0.051 <.0001	-0.027 <.0001	-0.121 <.0001
Industry Sigma	0.080 <.0001	-0.100 <.0001	-0.033 <.0001	0.165 <.0001	0.021 <.0001	-0.053 <.0001	0.049 <.0001	0.064 <.0001	0.045 <.0001	1.000	0.015 <.0001	-0.001 0.785	0.009 0.009	0.071 <.0001	0.053 <.0001
Dividend Dummy	-0.140 <.0001	0.049 <.0001	-0.095 <.0001	0.434 <.0001	-0.148 <.0001	0.133 <.0001	-0.205 <.0001	0.041 <.0001	0.365 <.0001	0.015 <.0001	1.000	0.055 <.0001	-0.043 <.0001	-0.059 <.0001	-0.081 <.0001
Managerial Ability	-0.011 0.001	0.036 <.0001	0.048 <.0001	0.036 <.0001	0.043 <.0001	0.108 <.0001	0.071 <.0001	0.081 <.0001	0.046 <.0001	-0.001 0.785	0.055 <.0001	1.000	0.005 0.094	-0.067 <.0001	0.076 <.0001
Interest Rates	0.125 <.0001	-0.115 <.0001	0.012 0.000	-0.097 <.0001	0.034 <.0001	-0.144 <.0001	0.070 <.0001	-0.013 0.000	-0.051 <.0001	0.009 0.009	-0.043 <.0001	0.005 0.094	1.000	0.022 <.0001	0.200 <.0001
Leverage	0.349 <.0001	-0.508 <.0001	0.005 0.167	0.020 <.0001	0.052 <.0001	-0.392 <.0001	0.184 <.0001	-0.006 0.086	-0.027 <.0001	0.071 <.0001	-0.059 <.0001	-0.067 <.0001	0.022 <.0001	1.000	0.115 <.0001
Market to Book	0.133 <.0001	-0.230 <.0001	0.063 <.0001	-0.157 <.0001	0.149 <.0001	-0.180 <.0001	0.229 <.0001	0.030 <.0001	-0.121 <.0001	0.053 <.0001	-0.081 <.0001	0.076 <.0001	0.200 <.0001	0.115 <.0001	1.000

From the above table results, we know the strength and magnitude of the relation between Cash and other variables. Key relationships are as below:

- Cash from Operations: A strong negative correlation (-0.62) suggests a robust negative linear relationship, indicating that as cash from operations increases, cash estimates tend to decrease.
- Research & Development: A moderate positive correlation (0.359) implies that there is a moderate positive linear relationship between research and development and cash estimates.
- Leverage: A moderate positive correlation (0.349) suggests a moderate positive linear relationship between leverage and cash estimates.
- Net Working Capital: A negative correlation (-0.45) indicates a moderate negative linear relationship, suggesting that as net working capital increases, cash estimates tend to decrease.

Results:

Cash Estimation Model 1:

Cash = Intercept + Firm size + NetWorkingCapital + Capital Expenditures + IndustrySigma + R&D +Dividend Dummy + SalesGrowth + CashfromOperations + FirmAge + TaxBurdenonForeignIncome+ YearDummies +IndustryDummies + Managerial Ability + Error

Cash Estimation Model 2:

Cash = Intercept + Firm size + NetWorkingCapital + Capital Expenditures + IndustrySigma + R&D +Dividend Dummy + SalesGrowth + CashfromOperations + FirmAge + TaxBurdenonForeignIncome+ YearDummies +IndustryDummies + Managerial Ability + Leverage + MarketToBook + InterestRates + Error

Table 3: Regression of Cash level on various firm characteristics

Dependent Variable: Cash and cash equivalents

Variable	Model 1		Model 2	
	Estimate	Pr > t	Estimate	Pr > t
<i>Intercept</i>	0.498	<.0001	0.270	0.0002
<i>Firm size</i>	-0.039	<.0001	-0.009	<.0001
<i>Net Working Capital</i>	-0.826	<.0001	-0.544	<.0001
<i>Capital Expenditures</i>	1.806	<.0001	1.546	<.0001
<i>Industry Sigma</i>	0.001	<.0001	0.001	<.0001
<i>Research & Development</i>	-1.194	<.0001	-0.293	<.0001
<i>Dividend Dummy</i>	-0.070	<.0001	-0.057	<.0001
<i>Sales Growth</i>	0.064	<.0001	0.064	<.0001
<i>Cash from Operations</i>	-1.210	<.0001	-1.110	<.0001
<i>Managerial Ability</i>	0.408	<.0001	0.225	<.0001
<i>Firm Age</i>	-0.101	<.0001	-0.067	<.0001
<i>Tax Burden on Foreign Income</i>	0.067	<.0001	0.068	<.0001
<i>Leverage</i>			0.144	<.0001
<i>Market To Book</i>			-0.022	<.0001
<i>Interest Rates</i>			0.002	<.0001
Observations	115,664		93,289	
F-Score	1017.30	<.0001	825.32	<.0001
R-square	46.82%		47.46%	
Adj R-square	46.77%		47.40%	

Notes: We include year and industry dummies (based on 2-digit SIC Code) that are not shown in the above table.

Interpretation of Results:

Model 1:

Model Significance:

The overall model fit is assessed using the R-squared value, which is 0.4682. This indicates that approximately 46.82% of the variance in cash estimation can be explained by the variables included in the model.

The F-statistic for the model is 1017.30 with a p-value $< .0001$, signifying that the model is statistically significant in explaining the variance in cash estimation. This indicates that the joint effect of all the predictors in the model is statistically significant in predicting cash level.

Predictors Interpretation:

Positive predictors such as Capital Expenditures, Industry Sigma, Sales Growth, Managerial Ability, and Tax Burden on Foreign Income contribute significantly to increased cash flow by their respective coefficient values in model 1. Higher capital expenditures positively impact cash flow, reflecting strategic investments in long-term assets and business expansion. Sales growth is associated with enhanced cash flow, highlighting the importance of expanding customer bases or increased demand. A positive coefficient for Industry Sigma indicates that being part of a strong and growing industry enhances cash generation. Additionally, positive coefficients for managerial ability and tax burden on foreign income suggest that effective management practices and international business operations positively influence cash generation.

Conversely, negative predictors such as Firm Size, Net Working Capital, Research and development, Dividend Dummy, Firm Age, and Cash from Operations are associated with decreased cash flow by their respective coefficient values in model 1. Larger firm size may pose challenges in managing cash efficiently, and negative net working capital indicates potential issues in working capital management. A negative coefficient for firm age suggests that older firms may face challenges maintaining optimal cash flow. Reduced cash flow linked to research and development expenses suggests the substantial costs of innovation. A negative coefficient indicates that dividend-paying firms tend to hold less cash. Lastly, a decrease in cash from operations implies potential difficulties in generating cash through regular business activities.

Model 2:

We expanded our analysis by incorporating additional factors including Leverage, Market to book, and Interest rates to consider the factors that were previously overlooked. The inclusion of these variables potentially enhances the explanatory power of the model, offering a more realistic representation of the complexities involved in cash estimation. We also see consistent results in terms of positive and negative signs for all the variables in both models.

Model Significance:

The overall model fit is assessed using the R-squared value indicating approximately 47.82% of the variance in cash estimation can be explained by the variables included in the model.

The F-statistic for the model is 825.32 and $Pr > F$ is $<.0001$, signifying that the model is statistically significant in explaining the variance in cash estimation. This indicates that the joint effect of all the predictors in the model is statistically significant in predicting cash.

Predictors Interpretation:

Positive predictors such as Capital Expenditures, Industry Sigma, Sales Growth, Managerial Ability, Tax Burden on Foreign Income, Interest rates, and Leverage contribute significantly to increased cash flow by their respective coefficient values in model 2. Higher interest rates are associated with increased cash flow, signaling increased financial costs impacting overall cash availability.

Conversely, negative predictors such as Firm Size, Net Working Capital, Research & Development, Dividend Dummy, Firm age and Market To Book are associated with decreased cash flow by their respective coefficient values in model 2. Dividend payments, financial leverage, and market-to-book ratios negatively impact cash flow, revealing the trade-offs between shareholder distributions, financial structure, and market perceptions. A positive coefficient for leverage may indicate that as a company takes on more debt, its cash levels increase, attributed to the immediate infusion of cash into the business facilitated by loans. Lastly, a negative coefficient for market-to-book suggests that if the market values a company more than its book value, it might negatively impact cash flow.

All the above insights provide valuable guidance for businesses aiming to optimize their cash flow dynamics.

Recommendations:

To bolster cash reserves, companies should implement effective strategies for managing working capital by implementing strategies to improve accounts receivable collection will ensure timely cash inflows. On the flip side, strategically managing accounts payable by negotiating favorable terms will optimize cash utilization. A company's working capital is a core part of funding its daily operations; however, it's important to analyze both the working capital and the cash flow of a company to determine whether the financial activity is a short-term or long-term event.

Emphasizing the effectiveness of management practices is crucial, as firms with stronger managerial abilities tend to experience higher cash flows.

To drive sustainable revenue growth, we recommend introducing strategies that diversify revenue streams. This includes expanding market share, introducing new products, and enhancing sales efforts. By embracing these strategies, firms can position themselves for long-term success.

Making strategic capital investments can positively impact cash positions, allowing for increased flexibility and room for growth. Firms must prioritize projects aligned with long-term goals, optimize financing, and carefully monitor returns to ensure effective management of capital expenditures which might help to acquire, upgrade, and maintain physical assets such as property, plants, buildings, technology, or equipment.

Efforts to optimize operational efficiency are vital for generating cash from operations and providing liquidity for day-to-day business activities. Focusing on efficient working capital management, timely collections, and cost-effective operations contributes to the enhancement of cash from operations and overall financial stability.

Monitoring changes in interest rates is essential, given their potential impact on financing costs. Firms should remain vigilant to market conditions and incorporate interest rate dynamics into financial planning and risk management.

While leveraging can contribute to cash, prudent management of debt levels is essential. Striking a balance between leveraging for growth and maintaining financial stability requires careful consideration of potential financial risks associated with increased leverage. Managing debt levels carefully is essential for balancing growth and financial stability. Regularly reviewing and adjusting our leverage strategies will help firms adapt to changing market conditions.

Limitations:

Our analysis faced constraints arising from varied data availability across timeframes. Notably, 'Cash from operations' data was unavailable before 1987, limiting insights into historical periods. Furthermore, numerous missing values in existing variables necessitated replacement with zeros, potentially introducing bias into our findings.

Additionally, our findings and recommendations do not universally apply across all industries or diverse business environments. Certain sectors or unique company characteristics may exhibit distinct cash management patterns not entirely captured by our models.

Our model may not fully capture the influence of external macroeconomic factors or unforeseen events on cash reserves. Economic fluctuations, regulatory changes, or crises could impact cash levels in ways beyond our modeled variables. This highlights the need for a comprehensive understanding of external influences.

These limitations underscore the need for caution in generalizing findings and suggest a need for broader data coverage and consideration of external factors in future analyses. The recommendations and insights we present do not universally apply to all industries or companies. Each organization is unique, and our model's sensitivity to context emphasizes the need for careful consideration of specific circumstances i.e. adjustments should be made based on the unique characteristics of each business.

By addressing these limitations transparently, we ensure a nuanced and informed interpretation of our findings, fostering more robust decision-making.

Conclusion:

In conclusion, our analysis offers a robust framework for estimating optimal cash levels in firms, highlighting key predictors influencing cash reserves, and providing actionable insights to firms to optimize cash flow, refine investments, and enhance financial health. While acknowledging limitations, the study serves as a valuable guide for navigating the intricate dynamics of cash management and long-term performance in today's business landscape.

Contributors:

Barkha Sharma
Sri Lakshmi Mallipudi
Amrapali Samanta
Rabiya Fatima

Literature Citation:

DEREK K. OLER, MARC P. PICCONI, Implications of Insufficient and Excess Cash for Future Performance

Appendix:

Variable description and calculation of variables	
Firm Size	Natural logarithm of total assets, adjusted using the Consumer Price Index.
Net Working Capital	Net working capital excluding cash, scaled by total assets less cash.
Capital Expenditures	Capital expenditures scaled by total assets less cash.
Industry Sigma	Average standard deviation of "cash flow" within 2-digit SIC code industries over the prior 20 years.
Research and Development	Research and development expenses scaled by total assets.
Dividend Dummy	Binary variable indicating whether the firm paid dividends to common shareholders in the prior year.
Sales Growth	Growth rate of net sales, calculated as the difference between current and prior year net sales, scaled by prior year net sales.
Cash from Operations	Cash flow from operations scaled by total assets less cash.
Firm Age	Natural logarithm of the number of years the firm has been publicly traded.
Tax Burden on Foreign Income	Ratio of (Foreign pretax income - U.S. statutory tax rate) to foreign income taxes paid, scaled by total assets less cash.
Interest Rates	Interest rate is calculated as interest expense divided by long-term debt plus current debt scaled by total assets less cash.
Leverage	Total debt scaled by total assets less cash.
Market To Book	Market value of equity scaled by total assets.
Gvkey	Global Company Key (gvkey) is a unique number assigned to each company in the Compustat-Capital IQ database
Nyear	The year extracted from 'Datadate' variable.

Year dummies	Dummy variables representing different years or periods.
Industry dummies	Dummy variables representing various industry categories or sectors.