

Stockholm School of Economics

Department of Accounting

Master's thesis in accounting and financial management

December, 2012

Forecasting cash flow in capital budgeting by integrating a tailored Huff model with the net present value technique – a case study in the Swedish retail industry^π

Oscar Schagerström^θ

Joachim Wingårdh^ω

ABSTRACT

We build a capital budgeting model which is able to compute the future cash flows of a proposed new store in the Swedish retail industry – rather than relying on case-by-case, experience-based cash flow forecasts by management. The coordinates of a proposed new store are inputted into the model whereby future cash flows and net present value (NPV) are calculated by the cash flow forecasting capital budgeting model (CFF-CB Model). The CFF-CB Model is also able to calculate “comprehensive NPV” for a proposed new store, which is the NPV of the proposed store less any lost NPV of already existing stores due to cannibalization effects.

The CFF-CB Model consists of a tailored Huff (1962) model, used for volume forecasting, and the NPV capital budgeting appraisal technique to convert the Huff model's volume forecasts into forecasts of cash flows and, subsequently, NPV and comprehensive NPV for a proposed new store.

The cash flow forecasts of the CFF-CB Model are back-tested against actual cash flows and the case company's *ex ante* forecasts. Average absolute forecast error (deviation from actual) is 20 percent for the CFF-CB Model, compared to 78 percent for the case company, which indicates reasonable accuracy in the model's forecasts. Average (non-absolute) forecast error is –6 percent for the CFF-CB Model, compared to 78 percent for the case company, demonstrating that forecasting cash flow with a model can mitigate the empirically documented problem of optimistically biased cash flow forecasts in capital budgeting. Finally, the features of comprehensive NPV are explored.

Keywords: capital budgeting, cash flow forecasting model, comprehensive NPV, Huff model

Tutor: Assistant Professor Katerina Hellström

^π In this thesis edition certain sensitive case-company data have been concealed, in accordance with our non-disclosure agreement with the case company and our agreement with the Department of Accounting at Stockholm School of Economics. The concealment does not substantially diminish any important aspect of our thesis. We may be able to grant access to the unconcealed version to researchers interested in obtaining it. Please send such a query to any of us per e-mail.

^θ 21545@student.hhs.se

^ω 21547@student.hhs.se

Acknowledgements

We would like to express gratitude to our tutor, Assistant Professor Katerina Hellström, for offering us guidance and inspiration throughout the process of writing this thesis. Her suggestions improved, above all, the structure and clarity of our thesis. In addition, we would like to thank the case company for providing us with important accounting data, necessary for conducting our research. We are particularly grateful to the CFO of the case company for his generous contribution of time, patience and helpful advice.

Contents

Definitions and abbreviations	iii
1. Introduction	1
2. Previous research	4
2.1. Volume forecasting in the retail sector.....	4
2.2. Cash flow forecasting in capital budgeting.....	6
2.3. Capital budgeting appraisal techniques	7
2.4. Research contribution	7
3. The case company and its industry	8
3.1. The case industry	8
3.2. The Company.....	8
4. Data and methodology	10
4.1. Building the Tailored Huff Model for estimating store volume	10
4.2. Building the cash flow forecasting capital budgeting model (the CFF-CB Model)	19
5. Results from estimating the parameters of the Tailored Huff Model.....	26
5.1. The first, broad, iteration round.....	26
5.2. The second, narrow, iteration round.....	27
5.3. Robustness discussion.....	28
6. Results from applying the CFF-CB Model.....	29
6.1. Back-test 1: Evaluating the cash flow forecasts at a point in time – accuracy and bias of the forecasts	29
6.2. Back-test 2: Evaluating the forecasting of growth in cash flow – development of forecasts over time in the changing competitive landscape.....	32
6.3. Hypothetical new store – exploring the comprehensive NPV concept	33
6.4. Robustness discussion.....	34
7. Conclusions	35
Appendix A. Supportive data and calculations.....	36
Bibliography	40
Additional data sources	42

Definitions and abbreviations

basic NPV	NPV of a proposed new store before cannibalization effects are taken into account. That is, NPV as it is commonly thought of. The term “basic” is used to distinguish it from comprehensive NPV (which takes cannibalization effects into account).
cash flow	Cash flow to equity as used in the NPV capital budgeting appraisal technique.
the CFF-CB Model	The cash flow forecasting capital budgeting model. A model that integrates the Tailored Huff Model with the NPV capital budgeting appraisal technique to forecast cash flow, NPV and comprehensive NPV for a proposed new store.
the Company	The case company that we write this thesis in collaboration with.
comprehensive NPV	Basic NPV of a proposed new store less NPV cannibalization on existing stores.
the Market Leader	A company in the case industry with a very dominant market position.
NPV cannibalization on existing stores	Lost NPV of one’s own existing stores due to the establishment of a new store. Calculated by the CFF-CB Model as specified in section 4.2.
product A	The most common product type in the industry. Accounts for 73 percent of the Company’s total volume and for 72 percent of the Market Leaders total volume.
the Tailored Huff Model	A Huff model tailored to the specifics of the case industry. The Tailored Huff Model is used to forecast volume for stores.
TAVD	Total absolute volume difference. This metric is minimized in section 5 in order to estimate the parameters of the Tailored Huff Model.
TTM	Trailing twelve months.
volume	Product-A volume unless otherwise stated.

1. Introduction

This thesis is written in collaboration with a company (the “Company”) active in the Swedish retail industry. We try to improve – with a comprehensive model – the Company’s “cash flow”¹ forecasting in its capital budgeting for new stores. The Company plans to open 10–50 stores per annum over the coming years and capital budgeting is thus a critical management activity.

Forecasting cash flow in capital budgeting is notoriously famous for being unreliable and optimistically biased (see e.g. Pruitt and Gitman, 1987; Statman and Tyebjee, 1985; Flyvbjerg *et al.* 2007). These negative features are also pervasive in the capital budgeting of the Company. Its *ex ante* cash flow forecasts for new stores have on average deviated 78 percent from *ex post* actual cash flows, and all these deviations have been higher than actual. That is, forecasts have been optimistically biased. Moreover, the Company projects its cash flow forecasts on a case-by-case basis, based on a combination of unstructured data and management experience. The Company’s CFO explains the capital budgeting for new stores:²

Well, we have a lot of useful data readily available in this industry, but no structured way of condensing it all into measures we can use when forecasting cash flow and payback time for potential new stores. We look at the array of data we have for an investment case and make a rough estimate for future cash flows. [...] Our estimates become better with experience, but are still not very accurate. [...] Several new stores open up every month in this industry and this can be difficult to keep track of.

In this case study, we build a capital budgeting model that itself *computes* cash flow and net present value (“NPV”) forecasts for a proposed new store if the store’s coordinates are inputted into the model. In other words, the model does not rely on case-by-case, experience-based cash flow forecasts being inputted by management. We call this model the cash flow forecasting capital budgeting model (the “CFF-CB Model”).

In addition, the CFF-CB Model is able to compute “comprehensive NPV” for a proposed new store. With comprehensive NPV we mean basic³ NPV of a proposed new store less any lost NPV of already existing stores due to cannibalization effects of introducing the proposed new store. This feature is made possible as the CFF-CB Model takes the entire competitive landscape into account in its capital budgeting.

The *case* purpose is to investigate whether it is possible to improve the Company’s, case-by-case, experience-based cash flow forecasts in its capital budgeting – with a comprehensive model. We hope to improve the cash flow forecasts in three ways: i) improve the accuracy of the forecasts, ii) mitigate the optimistic bias of the forecasts, and iii) induce a more holistic approach to capital budgeting by taking cannibalization effects into account by computing comprehensive NPV, rather than merely “basic NPV”.

The *general* purpose is to shed some light on the under-researched area within financial management of using models for forecasting cash flow in capital budgeting. Our hope is that some of our case-specific findings can be extrapolated to the general case.

¹ Throughout this thesis, for brevity, “cash flow” refers to “cash flow to equity in capital budgeting” unless otherwise required by the context.

² Nine interviews were conducted with the CFO between May 9, 2012 and November 6, 2012 as specified in section “Additional data sources”.

³ With basic NPV we mean NPV that does *not* take cannibalization effects into account. The term “basic” is used to distinguish it from comprehensive NPV (defined as NPV that *does* take cannibalization effects into account).

The CFF-CB Model consists of two parts: a Huff model for “volume”⁴ forecasting, and the NPV capital budgeting appraisal technique to transform these volume forecasts into forecasts of cash flow, basic NPV and comprehensive NPV.

The volume-estimating Huff model builds on the original Huff model. The original Huff model forecasts customer geographic behavior and can be used for forecasting volume of existing and potential retail stores (Huff, 1962; Huff, 2008). However, our Huff model is tailored to the specifics of the case industry and we henceforth refer to it as the “Tailored Huff Model”. When estimating volume for a store, the Tailored Huff Model takes into account i) the driving times to the store from surrounding locations, ii) the ramp-up level of the store, and iii) the impact of the stickiness of the old industry structure (previously a market leader (the “Market Leader”) enjoyed a very dominant market position) in the industry. In addition, these factors are accounted for in relation to the entire competitive landscape of the case industry. For example, the volume a store obtains from a particular location does not solely depend on the driving time from the location, but also on the driving time from the location to competing stores.

The Tailored Huff Model is suitable for the case industry for two main reasons. First, it estimates volume which is the key driver of cash flow and value. Second, it takes into account the entire competitive landscape, which currently is rapidly changing.

The NPV appraisal technique is just the standard procedure of calculating cash flows and then to discount these cash flows to get to basic NPV. We make no attempt to improve the technicalities of this, but the technique is integrated with the Tailored Huff Model.

What is unique to our study is the *integration* of the NPV technique with the volume-forecasting Tailored Huff Model, into the CFF-CB Model. This enables us to use the coordinates of any proposed new store to *model* its cash flow forecasts, rather than rely on them being *inputted* by management on a case-by-case basis.

Our research should be of practical interest for businesses as capital budgeting for new investments is a crucial management activity, strongly linked to shareholder value (Baker and English, 2011). In addition, forecasting cash flows is the most challenging part of capital budgeting – and, as noted, cash flow forecasts are generally unreliable and optimistically biased (Pruitt and Gitman, 1987; Statman and Tyebjee, 1985; Flyvbjerg *et al.*, 2007).

Moreover, our study is academically relevant because it covers a neglected field of research within financial management. Although previous research on capital budgeting appraisal techniques – NPV, IRR, payback time etcetera – is extensive, little attention has been directed towards constructing models for forecasting the cash flows that go in to these appraisal techniques (Turner and Guilding, 2012; Baker and English 2011).

On the other hand, in the marketing literature, store location is considered to be one of the most important management activities and numerous models for site selection have been developed (Craig *et al.*, 1984; Koontz, 2000). Among these, the Huff (1962) model is frequently used in practice and recognized in academia (Okabe and Okunuki, 2001; Joseph and Kuby 2011). However, despite the apparent link of using the volume estimates of the Huff model in capital budgeting, no papers have made the connection explicit by integrating the Huff model with a capital budgeting appraisal technique, as we do in this study.

The accuracy and bias of the CFF-CB Model’s cash flow forecasts are evaluated in two back-tests.

⁴ Volume refers to volume of product A unless otherwise stated. Product A is the most common product type and accounts for 73 percent of the Company’s total volume.

In the first back-test, the CFF-CB Model computes cash flow forecast for 23 of the Company's stores for the trailing twelve months ("TTM") ending on October 31, 2012. These forecasts are then compared to i) actual cash flows and ii) the Company's *ex ante* forecasts. The average absolute forecast error (an estimate's absolute percentage deviation against actual) for the CFF-CB Model is 20 percent, compared to 78 percent for the Company – for stores opened more than twelve months. This result indicates a reasonable accuracy in the CFF-CB Model's forecasts.

The average non-absolute forecast error for the CFF-CB Model is –6 percent, compared to 78 percent for the Company. This shows that modeling cash flow forecasts may overcome the pervasive problem of optimistically biased cash flow forecasts in capital budgeting.

In the second back-test, the CFF-CB Model computes cash flow forecasts for six of the Company's stores for the TTM ending on 30 April, and for the TTM ending on 31 October 2012. Thereby the forecasted growth in cash flow for these six stores over the six-month review period is obtained. The cash flow growth forecasted by the CFF-CB Model is compared to actual cash flow growth. Average absolute deviation is rather large at 2.7 percent but correlation is 0.78, indicating a relationship between growth in forecasted cash flow and growth in actual cash flow (excluding one outlier).

Finally the concept comprehensive NPV is explored. The CFF-CB Model computes capital budgeting for a hypothetical new Company store within the actual competitive landscape of a mid-sized Swedish city. We show that even though the hypothetical new store's basic NPV is positive, its comprehensive NPV is negative as it is placed too close to one of the Company's existing stores.

The results are within a reasonable degree of accuracy both in terms of point forecasts and the growth of the point forecasts over time. Moreover, the comprehensive NPV concept is theoretically appealing with practical value implications. Thus, our CFF-CB Model demonstrates considerable potential of this neglected academic field, within financial management, of modeled cash flow forecasting in capital budgeting.

The remainder of this study is structured as follows. In section 2 we review previous literature, in particular as it relates to the original Huff model, and to prior research on the accuracy and bias of cash flow forecasting in capital budgeting. Section 3 is an introduction to the case company and its industry. In section 4.1 the methodologies for building the Tailored Huff Model and for estimating its parameters are accounted for. In section 4.2 the methodology for integrating the Tailored Huff Model with the NPV technique, to form the CFF-CB Model is accounted for, as well as the concept comprehensive NPV. In section 5 we present and discuss the results from estimating the parameters of the Tailored Huff Model. In section 6.1 we present and discuss the results from back-testing the CFF-CB Model. In section 6.2 we explore the comprehensive NPV concept by computing capital budgeting for a hypothetical new store. Section 7 concludes our study.

2. Previous research

This section is structured around our research endeavor of integrating a volume-estimating Huff model with the NPV capital budgeting appraisal technique into a model for forecasting cash flow in capital budgeting. First, we present marketing literature on volume-forecasting methods and models in the retail sector. Second, we present literature on cash flow forecasting in capital budgeting. Third, we present previous research on capital budgeting appraisal techniques. Finally, we comment on our research contribution.

2.1. Volume forecasting in the retail sector

2.1.1. Overview of volume-forecasting methods in the retail sector

In the retail marketing literature, store location is considered to be one of the most important management activities and numerous methods for site selection have been developed (Craig *et al.*, 1984; Koontz, 2000). These methods can be used for estimating volume or sales of a proposed new store. According to Koontz (2000), methods for retail-site selection vary from intuition and subjective estimates to quantitative models based on theory. She identifies five key methods for site selection: the checklist method, analog approach, regression models, analog location allocation models, and retail gravity models. A brief summary of Koontz' (2000) description of these methods follows:

The checklist method is based on the decision maker's intuitive ability to systematically evaluate the relative value of a location by comparing it to other potential locations. The *analog approach* also depends on the decision maker's intuitive judgment. The analog approach is based on the assumption that a proposed store's attractiveness is analogous to similar stores operated by the company making the assessment. *Regression models* relate store sales as the dependent variable to a number of independent variables such as population size and traffic patterns. A general critique to the aforementioned three methods is that they focus on site-specific characteristics and fail to account for important competitive characteristics.

Analog location allocation models shift the focus from site-specific characteristics to evaluating the entire trading or market area by analyzing a proposed retail store in terms of market or profit share. *Retail gravity models* are based the assumption that customers select retail stores in proportion to the attractiveness of the stores and in inverse proportion to the stores' distance from the customers. Attractiveness can be measured by for example square-meter selling area and distance is often measured by driving time, driving distance or Euclidean distance. A key feature of retail gravity models is that they can estimate sales or volume for all retail stores in the entire competitive landscape.

Among the above methods, gravity models are popular both academically and practically (Joseph and Kuby, 2011). William J. Reilly and David L. Huff pioneered retail trade area analysis through the formulation of the two most popular retail gravity models (Joseph and Kuby, 2011). Reilly introduced the "Law of Retail Gravitation" in 1931 and Huff followed with the alternative Huff model in 1962 (Huff, 1962). In particular, the Huff model has become a landmark retail-site-selection model (Joseph and Kuby, 2011; Koontz, 2000).

As we use a tailored application of the Huff model for volume estimation in the CFF-CB Model, we continue this section with a brief description of Reilly's model but put most emphasis on the Huff model.

2.1.2. The Law of Retail Gravitation

Reilly hypothesizes that retail business gravitates from smaller cities and towns to larger cities, and calls this the "Law of Retail Gravitation" (Huff, 1962; Joseph and Kuby, 2011). Reilly argues that two key location factors, city population and distance can be used to identify the breaking points of retail influence

of two competing cities on an intermediate town. More specifically, Reilly reasons that two cities attract retail trade from the intermediate town in the area corresponding to the breaking point approximately in:

- i) direct proportion (one-to-one) to the two cities' populations; and
- ii) inverse proportion to the square of the distances from the two cities to the in-between, smaller, intermediate town.

This implies that if an intermediate town is located exactly between two equally large (population wise) cities, city A and city B, each city attracts 50 percent of the intermediate town's retail business. All else equal, if the population in city A doubles, city A's share of the retail business increases to 66.7 percent. If, all else equal, the distance from city A to the intermediate town doubles, retail trade would decrease to 25 percent for city A (Huff, 1962; Joseph and Kuby, 2011).

2.1.3. The Huff model

The Huff model builds on Reilly's Law of Retail Gravitation and is used to forecast customer geographic behavior and can be used for forecasting volume of existing and potential retail stores, as well as defining and analyzing retail store trade areas (Huff, 1962; Huff, 2008).

Huff (1962) acknowledges that Reilly's model introduces an interesting conceptualization of retail trade gravitation, but nonetheless points out two main limitations of the model. i) It only accounts for inter-urban retail trade gravitation for two cities, whereas intra-urban gravitation within a multiple-store competitive landscape often is more relevant to consider for a company. ii) Reilly's model is *deterministic* – that is, customers are assumed to always go to the same store.

Huff (1962, 2008) addresses the above limitations by presenting an alternative gravitation model. The Huff model recognizes that customer geographic behavior is probabilistic rather than deterministic. That is, when a customer is confronted with a set of alternative retail stores, the probability that a certain store is selected is proportional to the perceived utility of that store. In addition, the Huff model is specified in such a way as to take into account all the stores in the relevant competitive landscape and can thus be used for predicting consumer geographic behavior in an intra-urban setting with many competing stores.

The Huff model is based on empirical evidence suggesting that two variables exert such an influence on the utility, of a retail store to a customer, that they are sufficient to predict consumer geographic behavior. These two variables are:

- i) store attractiveness, represented by the number of items carried by the retail store that are demanded by customers; and
- ii) the distance or travel time from a customer to alternative retail stores.

The first variable, store attractiveness, often estimated as square-meter selling area, exerts a proportionally positive effect on store utility to a customer. The second variable, distance or travel time, exerts an inversely proportional effect on store utility to a customer. (Huff, 1962)

A basic formulation of the Huff model is as follows (Huff, 2008; Huff, 1962):

$$(\text{Probability})_{i,j} = \frac{(\text{Utility})_{i,j}}{\sum (\text{Utility})_{i,j}} = \frac{(\text{Store size})_j^\alpha \times (\text{Travel time})_{i,j}^{-\beta}}{\sum [(\text{Store size})_j^\alpha \times (\text{Travel time})_{i,j}^{-\beta}]}$$

Equation 2.1

where:

j	Retail store j
i	Customer location i
$(\text{Probability})_{i,j}$	The probability that a customer located at i will patronize store j
$(\text{Utility})_{i,j}$	The perceived utility of store j to customer at location i
$(\text{Store size})^{\alpha}_{i,j}$	The size of store j , serving as a proxy for store attractiveness. Raised to the power of the store-size parameter α
$(\text{Travel time})^{-\beta}_{i,j}$	The travel time between customer location i and store j . Raised to the negative power of the distance-decay parameter β to capture the assumed inverse relationship between travel time and store utility

Huff argues that the model's store-size parameter α and distance-decay parameter β should be estimated statistically based on survey data from actual customer preferences in for example census blocks (Huff, 2008).

Empirical studies on gravity models

In empirical studies, gravity models – primarily variants of the Huff model – have been applied in settings ranging from strict consumer goods to service sectors such as universities and hospitals (Gautschi, 1981; Drezner and Drezner, 2002; Bruno and Improta, 2006; McLafferty, 1988). Most of the empirical studies support the usefulness of the Huff model, and similar gravity models, for describing consumer geographic behavior (Craig *et al.*, 1984).

The distance-decay parameter (β in equation 2.1) typically varies by type of retail sector and methodology (Joseph and Kuby, 2011). Huff (1962) estimated the distance-decay parameter in the range of 2.1 to 3.8 for retail and furniture stores in different neighborhoods. Fotheringham (2011) notes that ten different empirical papers estimated the distance-decay parameter in the range of negative 0.5 to positive 5.2. If the distance-decay parameter is set without estimation, it is commonly assumed to be 2 (see e.g. Xu and Liu, 2004; Huff 1962).

2.2. Cash flow forecasting in capital budgeting

2.2.1. Models for cash flow forecasting in capital budgeting

According to Baker and English (2011) projecting cash flows is the most challenging phase of the capital budgeting process. Despite its importance, previous research on models for cash flow forecasting in capital budgeting is scant (Biondi and Marzo, 2011; Turner and Guilding, 2012).

In operational budgeting, research on sales forecasting is extensive including methods such as a jury of executive opinion, moving average, straight-line projection and trend-line analysis (see Mccharty *et al.*, 2006 for a review). However, these methods are usually not translatable to a capital budgeting context due to the lack of historical data when investing in new projects.

The only paper covering models for cash flow forecasting in capital budgeting, discovered by us, is a survey on practices in 232 U.S. firms by Pohlman *et al.* (1988). They find that 90.5 percent of the respondents used management's subjective estimates for cash flow forecasting. 48.3 percent of respondents used sophisticated mathematical models. The authors conclude that firms who combine judgmental and quantitative methods greatly improve their cash flow forecasts.

2.2.2. Accuracy and bias of cash flow forecasts

Several studies have been made on the accuracy and bias of cash flow forecasts. In empirical surveys, Pruitt and Gitman (1987) as well as Statman and Tyebjee (1985), find that capital budgeting cash flow forecasts are optimistically biased in a majority of surveyed firms. Forecasted sales are higher than actual sales, and forecasted costs are lower than actual costs. These findings are also consistent with more recent research by Flyvbjerg *et al.* (2007) who hypothesize that "outright lying" of forecasters rather than poor forecasting techniques is a plausible explanation to the forecast errors. Turner and Guilding (2012), however, emphasize that empirical examination of the *sources* for the poor accuracy in, and the optimistic bias of, cash flow forecasts in capital budgeting has been minimal.

2.3. Capital budgeting appraisal techniques

The most common capital budgeting appraisal techniques are payback time, NPV, and internal rate of return. Among these, NPV is generally considered theoretically superior (Brunzell *et al.*, 2011; Ryan, 2002). In practice, Graham and Harvey (2001) find that among 392 CFOs in U.S. firms, internal rate of return followed by NPV and payback time are most popular. In a similar survey on 313 European CFOs' practices, Brounen *et al.* (2004) find that payback time followed by NPV and internal rate of return are most popular.

2.4. Research contribution

Our research review points to the importance of, and the difficulty in, forecasting cash flows in capital budgeting. In addition, cash flow forecasts are unreliable and prone to be optimistically biased. It is therefore surprising that we did not discover any previous research on models for cash flow forecasting in capital budgeting. By developing and evaluating our CFF-CB Model we hope to shed some light into this identified research gap within financial management.

3. The case company and its industry

In this chapter the Company and the industry it operates in are introduced, for the purpose of conveying two key issues. First, the market size in terms of volume is very stable and predictable, which is a suitable condition for the Tailored Huff Model. Second, the Company's cash flow forecasting in its capital budgeting is currently conducted on a case-by-case basis primarily reliant on management's experience – rather than a model.

Due to our non-disclosure agreement with the Company parts of this section have been concealed.

The sources of information in this chapter are, unless otherwise stated, based on interviews with the Company's CFO as well as Company data.

3.1. The case industry

The case industry is part of the Swedish retail industry and was recently reregulated. Prior to the reregulation the Market Leader had certain regulatory privileges which induced it to a very strong market position. The reregulation favored competition and since it took affect several new companies have entered the market, such as the Company.

The following case-industry characteristics make it particularly suitable for our research:

- The total market volume is very stable
- Very reliable, detailed volume data for the aggregated market is publicly available
- Location is the key volume driver in the industry

3.1.1. Product A volume and total volume

Product A is the most common product and comprises 72 percent of the total volume in the industry. In the Tailored Huff Model we only forecast product-A volume. This is done for two reasons. First, the Company primarily offers product A. Second, product A's share of total volume has been stable over time, and is expected to continue to be so. "Volume" in this paper refers to product A volume unless otherwise stated.

3.1.2. Key value drivers – location is paramount

Location is the key volume and value driver in the industry. The most important factors determining location attractiveness are proximity to customers' home and work, and remoteness to competitors. Other value drivers related to location include accessibility and visibility of a store, as well as proximity to shopping centers and other car clusters.

Additional value drivers, apart from location, include differentiation based on opening hours, pricing, brand, and customer service as well as marketing efforts. However, these factors are of minor importance relative to location.

3.2. The Company

3.2.1. Introduction to the Company

Company history

The Company is private equity owned and operates 20–60 stores and has around 50–500 employees. The Company's objective for the coming two to three years is to continue its rapid expansion by opening approximately 10–50 stores per annum. Thus, capital budgeting and cash flow forecasting for new stores are very important for the Company.

3.2.2. The Company's capital budgeting in brief

The Company's capital budgeting involves the board, senior management (CEO and CFO) and regional managers. Regional managers are responsible for scouting of new stores. For each investment proposal they produce an investment memorandum in close contact with the senior management. The investment memorandum is then presented to the board for review and approval.

The investment memorandum is the cornerstone of the Company's capital budgeting. The most important part in the investment memorandum is a payback-time calculation based on annual cash flow forecasts.

The main problem with the Company's capital budgeting is that cash flow forecasts are unreliable and persistently too optimistic. *Ex ante* cash flow forecasts for new stores have so far on average deviated 78 percent from the *ex post* actual figures, and all these deviations have been higher than actual outcome.

Forecasts on price, expenses, working capital and capital expenditure are not motivated in detail and are relatively close to actual. Volume forecasts are more difficult to produce and deviates the most from actual. The investment memoranda therefore include qualitative sections with some quantitative data supporting the volume forecasts. However, the Company does not have a model for volume or cash flow forecasting, and the Company's CFO describes the forecasts as case-by-case based, primarily reliant on management and store scouts' experience.

4. Data and methodology

This section accounts for the data sources and methodologies used for building the CFF-CB Model. First, this is done for the Tailored Huff Model (section 4.1), which is used by the CFF-CB Model for forecasting volume. Thereafter we account for the data and methodologies used for integrating the Tailored Huff Model with the NPV capital budgeting appraisal technique, to form our integrated CFF-CB Model (section 4.2). The integration is performed in order to transform the volume estimates from the Tailored Huff Model into forecasts of cash flow, basic NPV and comprehensive NPV for the proposed new store which is capital budgeted for. Thus, in section 4.2 we also account for how the CFF-CB Model computes basic NPV and comprehensive NPV.

4.1. Building the Tailored Huff Model for estimating store volume

According to Huff (1962), the Huff model should be tailored to each specific industry and could include other variables than his original travel-time and store-attractiveness variables. This is rather intuitive as one, for example, would expect customer geographic behavior in the fast-food industry to differ from that of the fashion clothing industry.

In this section we construct a Huff model tailored to the specifics of the case industry – the Tailored Huff Model. First we specify the model equation and lay out its intuition. Then we describe the rational, methodologies and data sources for each variable separately. Thereafter we discuss some excluded variables. Finally we describe the methodology we use for estimating the model parameters.

4.1.1. A Huff model tailored to the case industry

Our Tailored Huff Model postulates that store j 's percentage of the total volume emanating from postal code i , during the TTM, is determined by the equation:

$$\left(\text{Store } j\text{'s percentage of postal code } i\text{'s volume in TTM} \right)_{i,j} = \frac{\left(\text{Old-structure stickiness} \right)_j^s \times \left(\text{Driving time} \right)_{i,j}^{-d} \times \left(\text{Ramp up TTM} \right)_j \times \left(\text{Incomplete TTM} \right)_j}{\sum \left[\left(\text{Old-structure stickiness} \right)_j^s \times \left(\text{Driving time} \right)_{i,j}^{-d} \times \left(\text{Ramp up TTM} \right)_j \times \left(\text{Incomplete TTM} \right)_j \right]}$$

Equation 4.1

where:

$\left(\text{Old-structure stickiness} \right)_j^s$	Old-structure stickiness of store j . Assumes that old Market Leader stores attract additional volume as they are well-known. The number of x-attributes is the proxy for old-structure stickiness of old Market Leader stores. Raised to the power of parameter s which is estimated empirically to determine the magnitude of the old-structure-stickiness effect.
$\left(\text{Driving time} \right)_{i,j}^{-d}$	The driving time in minutes from postal code i to store j . Raised to the negative power of parameter d which is estimated empirically to determine the importance of a customer's driving time to stores.
$\left(\text{Ramp up TTM} \right)_j$	The average ramp-up level of store j during the TTM. The number of months opened determine the ramp up level. Includes a parameter r – as explained in section 4.1.4 – which is estimated empirically to determine the ramp-up pattern of newly opened stores.
$\left(\text{Incomplete TTM} \right)_j$	A control variable, adjusting the volume estimates for stores that have only been opened during part of the TTM. The distinction and interac-

tion between the incomplete-TTM variable and the ramp-up variable is explained in section 4.1.5.

Equation 4.1 implies that the Tailored Huff Model distributes the volume emanating from a particular postal code onto its surrounding stores depending on the relative driving times to the stores, the relative ramp-up levels of the stores, and whether any of the stores asserts old-structure stickiness onto the postal code. Thus, the Tailored Huff Model takes the entire competitive landscape into account when computing its volume estimates. Next, we illustrate this with an intuitive example.

Intuitive explanation of the Tailored Huff Model

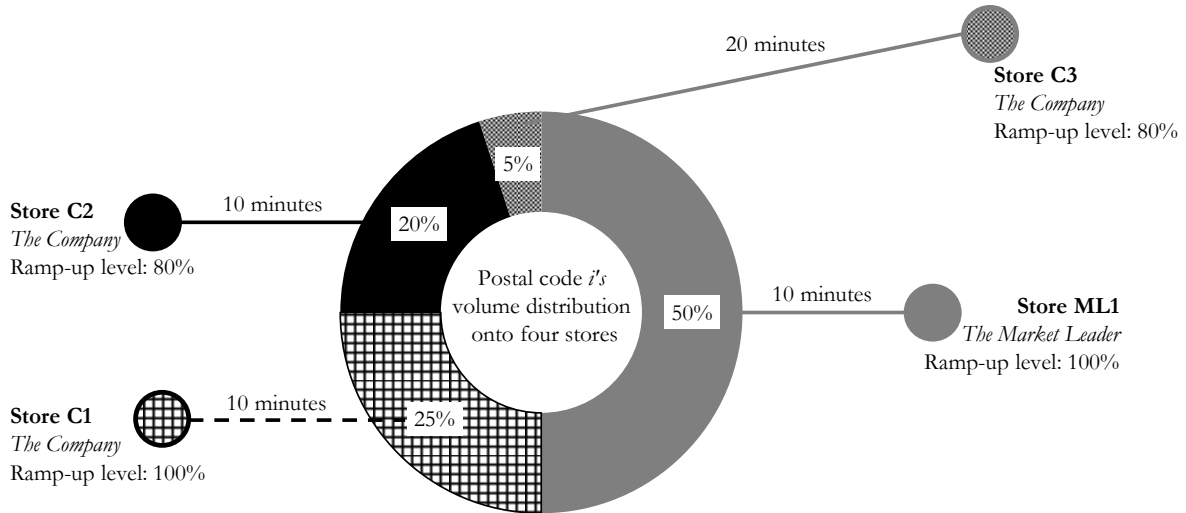
Figure 4.1 below illustrates intuitively how the volume emanating from a fictional postal code is distributed onto its four surrounding fictional stores according to the Tailored Huff Model.

Store ML1 and C1 have the same ramp-up level (100 percent) and are both a ten-minute drive away from the postal code. But ML1 is an old Market Leader store (whereas C1 is a fairly new store of the Company) and asserts old-structure influence on the postal code, because the residents have been patronizing the old Market Leader store for decades and are generally satisfied with its service. Therefore store ML1 receives 50 percent of the postal code's total volume, compared to 25 percent for store C1.

Store C1 and C2 are both owned by the Company and are both a driving time of ten minutes away from the postal code. But store C2 is younger than store C1 and its ramp-up level is only 80 percent. Therefore store C2 receives a volume share of 20 percent, 5 percentage points below that of store C1.

Stores C3 and C2 are both owned by the Company and both have a ramp-up level of 80 percent, but store C3 is located 20 minutes away from the postal code – twice the driving time of store C2. Thus, store C3 only receives 5 percent of the postal code's volume.

Figure 4.1. Intuitive illustration of how the Tailored Huff Model forecasts the distribution of a postal code's volume onto its surrounding stores based on driving times, old-structure stickiness and ramp-up level for the stores



In the example above, the parameters s , d , and r have arbitrarily been set to 1.0, 2.0 and 0.5, respectively, and these values determine the magnitude of the effect of the variables. If, for example, we were to set parameter d to 0.0, this would imply no driving time decay effect at all (which is rather unlikely) and stores C2 and C3 would receive the same volume share.

A key idea with the Tailored Huff Model – as with the original Huff Model – is that the variables' parameters should be estimated empirically. That is, the effects of driving time, old-structure stickiness and ramp

up should be estimated empirically, rather than set based on assumptions. Our estimation technique is explained last in this section and the results are presented in section 5.

Converting percentage of volume to volume

Equation 4.1 expressed store j 's percentage of postal code i 's volume in the TTM. To convert this *percentage of volume* into *volume*, the Tailored Huff Model multiplies it with the total volume emanating from postal code i in the TTM:

$$\left(\begin{array}{c} \text{Store } j\text{'s volume} \\ \text{from postal code } i \\ \text{in TTM} \end{array} \right)_{i,j} = \left(\begin{array}{c} \text{Store } j\text{'s percentage} \\ \text{of postal code } i\text{'s} \\ \text{volume in TTM} \end{array} \right)_{i,j} \times \left(\begin{array}{c} \text{Total volume} \\ \text{from postal code } i \\ \text{in TTM} \end{array} \right)_{i,j}$$

Equation 4.2

To estimate store j 's total volume in the TTM, the Tailored Huff Model summarizes the volume it receives from all postal codes:

$$\left(\begin{array}{c} \text{Store } j\text{'s total} \\ \text{volume} \\ \text{in TTM} \end{array} \right)_j = \sum_i \left(\begin{array}{c} \text{Store } j\text{'s volume} \\ \text{from postal code } i \\ \text{in TTM} \end{array} \right)_{i,j}$$

Equation 4.3

Next we will lay out the rationale behind the variables in the Tailored Huff Model in more detail, explain the technicalities of them, and account for the data sources for them.

4.1.2. The old-structure-stickiness variable

Rationale for including the variable in the Tailored Huff Model

We include the old-structure-stickiness variable in the model because we assume that the old industry structure – in which the Market Leader had a very dominant market position – wields a certain degree of stickiness. All else equal, a customer should be more likely to travel to an “old Market Leader store” than to a newly opened store, for example because of familiarity with that store and habitual behavior. The Market Leader also enjoys high customer satisfaction (the Market Leader’s annual report 2011).

Methodology and technicalities of the variable

We define “old Market Leader stores” as stores that opened before January 1, 2010. Stores that opened after this date we define as “new stores” or “newly opened stores”. These newly opened stores include both new Market Leader stores and new stores opened by new entrants, such as the Company. The motivation for this cut-off date is that it was only in 2010 that the Market Leader started to open many new *small* stores each year, designed for the new industry structure (the Market Leader’s annual reports 2009, 2010 and 2011).

As a proxy for stickiness of an old Market Leader store we use its number of x-attributes. An x-attribute could be seen as a measure of the capacity of a store to service its customers. When the Market Leader enjoyed a very dominant market position it could control and calculate demand to a great extent. For example, chose whether to place one store with ten x-attributes or two stores with five x-attributes each in a particular area, without needing to consider competitors’ actions very much. The more x-attributes an old Market Leader store has the more customers it presumably had in the past. Hence, stickiness of an old Market Leader store should increase with the number of x-attributes it has.

A newly opened store can, by definition, not exert old-structure stickiness. Therefore, all newly opened stores are assigned the value 1.0 for old-structure stickiness (regardless of its number of x-attributes).

To summarize: An old Market Leader store receives a value for the old-structure-stickiness variable equal to its number of x-attributes. New stores all receive the value 1.0 for the old-structure-stickiness variable.

We have data that lends support to our reasoning in the previous paragraphs. Figure 4.2 depicts a positive relationship between the number of x-attributes and the estimated (from the Company's investment memoranda) volume for 21 old Market Leader stores. Conversely, figure 4.3 indicates a non-existing relationship between the number of x-attributes and volume for eleven of the Company's stores, which are all newly opened.

Figure 4.2. Volume as a function of number of x-attributes for old Market Leader stores indicates that old-structure-stickiness exists in the industry

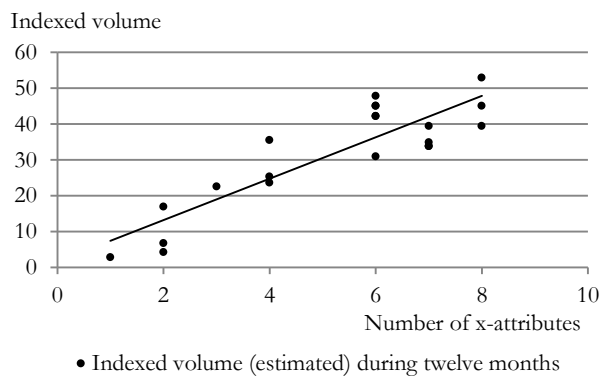
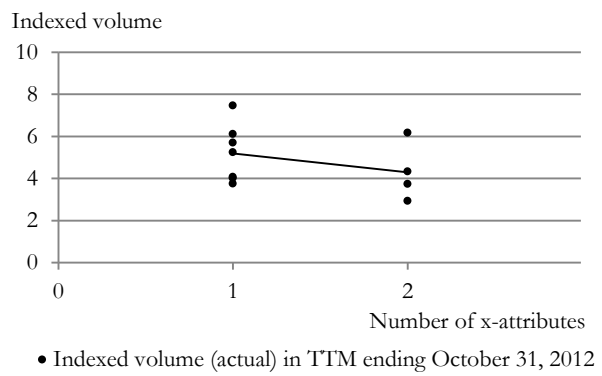


Figure 4.3. Volume as a function of number of x-attributes for new Company stores indicates no clear volume–x-attributes relationship for new stores

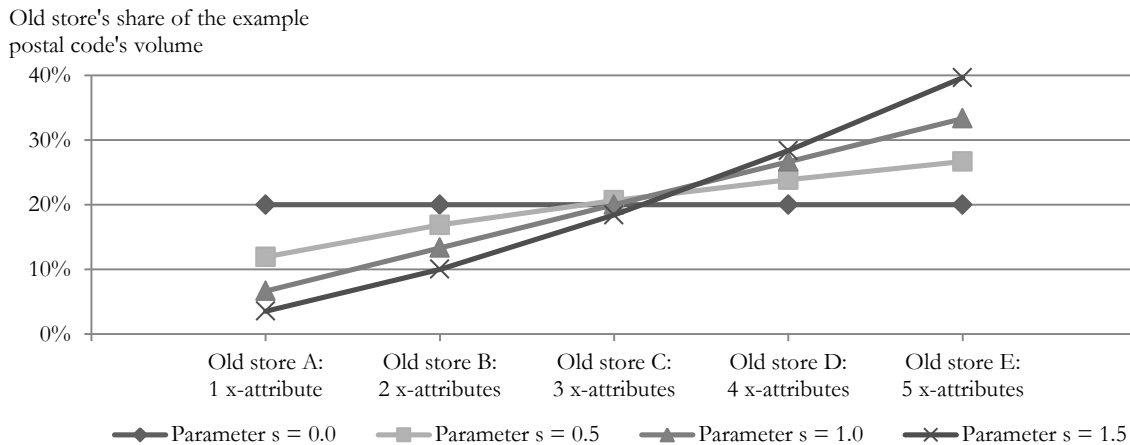


Parameter s and variable properties

The old-structure-stickiness variable is raised to the power of parameter s , for us to be able to estimate the magnitude of the potential old-structure-stickiness effect.

The properties of the old-structure-stickiness variable for different values of the s parameter are illustrated below. The five example stores have the same driving distance from an example postal code, and the same ramp-up values (100 percent). The only thing that differs is the number of x-attributes.

Figure 4.4 Illustration of how different values for the s parameter change the magnitude of the impact of stores' old-structure stickiness on the distribution of a postal code's volume onto its surrounding stores



If the stickiness effect in reality is weak, the s parameter should be estimated at close to zero. For this case ($s = 0.0$), in the example chart, all five stores get the same share, 20 percent each, of the postal code's total volume.

If the stickiness effect in reality is strong and close to proportional to the number of x-attributes of old stores, the s parameter should be estimated at close to one. For this case ($s = 1.0$), in the example chart, the five stores get a volume share proportional to its number of x-attributes.

We estimate the s parameter empirically in section 5.

Data

Number of x-attributes for old Market Leader stores

We need the number of x-attributes for all old Market Leader stores, but not for the newly opened stores as these all receive the value 1.0, regardless of their number of x-attributes.

To collect the number of x-attributes for old Market Leader stores we used Google Street View, the Company's investment memoranda and set some values arbitrarily. Google Street view identified the number of x-attributes for most stores. The Company's investment memoranda identified the number of x-attributes for an additional number of old Market Leader stores. The number of x-attributes for the remaining old Market Leader stores were set arbitrarily, mostly for stores in the north of Sweden which do not impact the volume estimates for the Company's stores (and thus neither impacts the iteration of the parameters of the Tailored Huff Model).

4.1.3. The driving-time variable

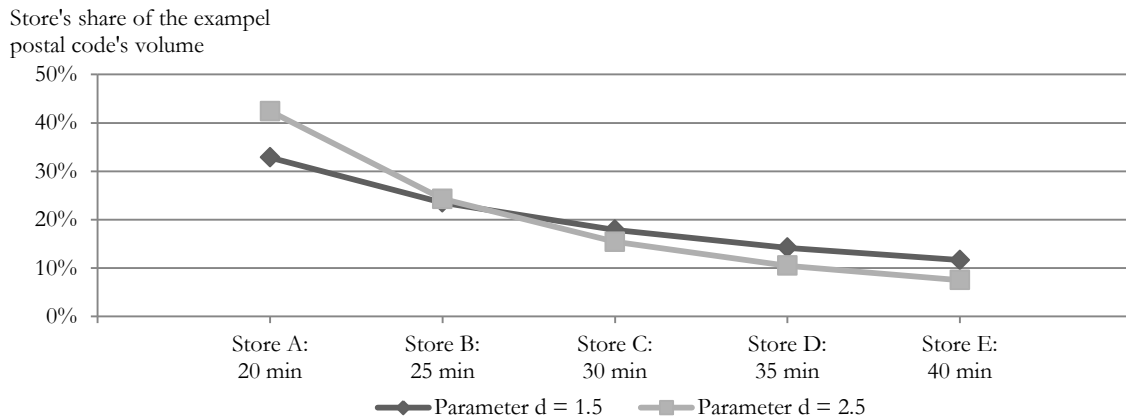
Rationale for including the driving-time variable in the Tailored Huff Model

The driving-time variable is included in the model, raised to the negative power of parameter d . This implies that we assume that driving-time is inversely related to the probability of a customer patronizing a store. The rationale is straightforward. As travelling by car involves effort, time and mileage cost, the average customer should prefer a shorter trip to a longer one, all else equal. The original Huff model included a distance variable (sibling of driving time) with similar motivation as we use (Huff, 1962). The CFO of the Company also believes that proximity to customers in terms of driving time is the most important volume driver in the industry.

Parameter d and variable properties

The stronger the inverse effect of driving time is on customers' store selection, the larger the parameter d should be estimated at (parameter d is positive because of the minus sign preceding it). This property is illustrated in the chart below. The five example stores differ, with regard to the example postal code, only for driving time. All else is equal, for example old-structure stickiness and ramp-up value. When parameter d is stronger, at 2.5 rather than 1.5, driving time is a more important factor and the two closest stores receive a larger share of the volume, at the expense of the three stores that are farther away.

Figure 4.5. Illustration of how different values for the d parameter changes the magnitude of the impact of driving time on the distribution of a postal code's volume onto its surrounding stores



We estimate the d parameter empirically in section 5.

Data and methodology

First we collected coordinates for i) all stores, and ii) all relevant postal codes in Sweden. Thereafter we retrieved the relevant driving times between these postal codes and stores.

We collected coordinates for all stores that were operational in Sweden as of October 31, 2012. First we collected their addresses. Then we ran the “geocode” command in Stata to transform these addresses into coordinates, by following the procedure laid out by Ozimek and Miles (2011).

Then we collected coordinates for 9,675 postal codes in Sweden. Swedish postal codes are provided free of charge by Statistiska centralbyrån (Statistics Sweden). We dropped 52 postal codes for various reasons, for example postal codes for islands from where no meaningful driving times could be retrieved. For the remaining 9,623 postal codes we retrieved coordinates by, once again, running the Stata command “geocode”.

The total store–postal-code pairs of 2–6 million were reduced to 224,555 by applying a cut-off rule based on crow-flight distances. This was done because it is irrelevant to have the driving time between a postal code in Kiruna and a store in Stockholm (950 kilometers apart) – and *calculating* a crow-flight distance is about a million times faster than *retrieving* a driving time from Google Maps.

Driving times for the remaining, relevant, 224,555 store–postal-code pairs were retrieved from Google Maps through the Stata “traveltime” command (Ozimek and Miles, 2011), which took about ten hours.

4.1.4. The ramp-up variable

Rationale for including the ramp-up variable in the Tailored Huff Model

The rationale behind the ramp-up variable is that it takes time for a new store to raise awareness of its existence among potential customers. This implies that, all else equal, the store’s volume will increase gradually during some ramp-up period until full ramp-up level is reached.

Presumably, stores can affect the development of their own ramp-up levels, by marketing efforts for example. We are, however, estimating if there is an average ramp-up level to be detected in the industry. If there is, it is important for us to take this into account in the model as many new stores are being established continually.

Technicalities

The ramp-up variable is constructed as:

$$\left(\text{Ramp up} \right)_{\text{TTM}}^j = r + \frac{(100\% - r)}{\text{Ramp-up period} (= 30 \text{ months})} \times (\text{No of months opened})$$

Equation 4.4

where:

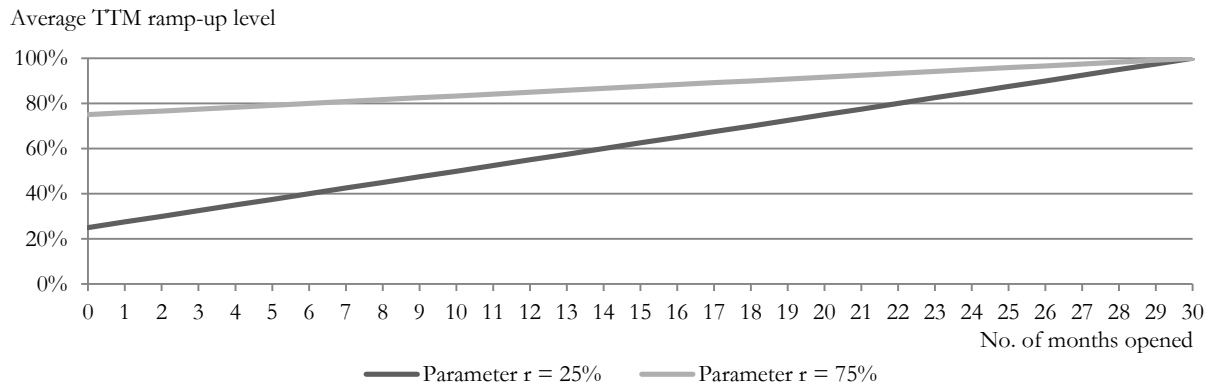
$(\text{Ramp up TTM})_j$	The average ramp-up level of store j during the TTM ⁵
r	Ramp up level at the opening date of a store.
100%	Refers to the fact that a store has a ramp-up level of 100 percent when it is fully ramped up.
Ramp-up period	Refers to the number of months it takes for a store's average TTM ⁵ ramp-up level to reach 100 percent – assumed to be 30 months.

Hence, a store starts at ramp-up level r and then has a linear development of its average TTM ramp-up level during the ramp-up period which we have set to 30 months based on the data analysis in figure 4.7.

Parameter r and variable properties

The parameter r is, then, the ramp-up level of a store at its opening day. If stores are fully ramped up at opening, then r should be estimated at 100 percent. Below is illustrated the average TTM⁵ ramp-up pattern for a store, for parameter r values of 25 percent and 75 percent.

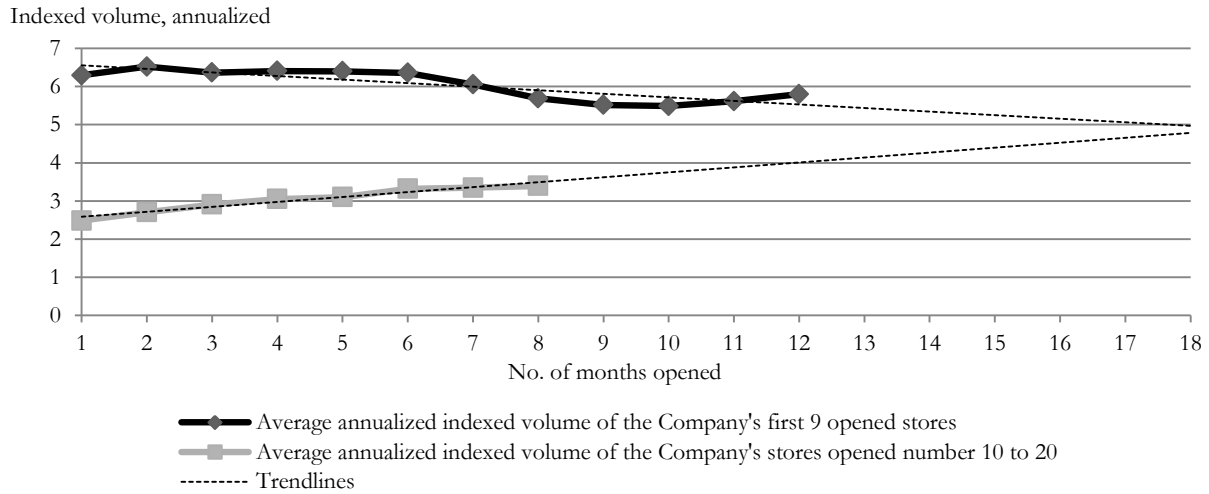
Figure 4.6. Illustration of how different values for the r parameter change the ramp-up pattern of a newly opened store – and thus the estimated volume it receives from the Tailored Huff Model



In analyzing the store P&Ls we noticed that the ramp-up pattern for the Company's stores seem to have changed quite considerably recently, as illustrated in figure 4.7 below. The first nine stores on average start out at a high level, around 6 in indexed volume, and stay around that level. The stores opened number 10 to 20, on average, start out at a 60 percent lower level but are increasing fairly linearly, on average by 5 percent per month. The Company's CFO believes that the reason for this pattern is that the pace of opening up new stores has almost doubled since the first nine stores were opened and there is now less effort spent on advertising for each store before it opens.

⁵ TTM if the store has been opened for twelve months or more, otherwise during the months the store has been opened.

Figure 4.7. The shift in ramp-up pattern of the first 9 opened Company stores and the Company stores opened number 10 to 20



The data in the figure above suggest that there has been a shift in the ramp-up pattern of stores since the establishment pace increased. Therefore it would be improper, to force equation 4.4, onto the Company's first 9 opened stores that were virtually fully ramped up at their opening date. Thus, we set the Company's first nine stores as fully ramped up. Subsequently opened stores receive a ramp-up level according to equation 4.4, which implies that their average TTM ramp-up level reach 100 percent after 30 months.

Because of the data provided in figure 4.7 we expect that when we estimate the parameters of the Tailored Huff Model empirically in section 5 we will detect a considerable ramp-up effect, with a starting ramp-up level r at somewhere around 50 percent. However, we cannot merely use the data in figure 4.7 to set the r parameter because this data does not take the effects of old-structure stickiness and driving-time into account, which the Tailored Huff Model does. In other word, the data in figure 4.7 *indicate* a ramp-up effect and prompt us to include the ramp-up variable, but the ramp-up effect is better estimated while taking other effects on store volume than just the ramp-up level into account.

Data

The number of months opened, for stores opened less than 30 months, is calculated using their opening dates. Opening dates were collected from annual reports, the Company's P&Ls and by Google searches.

4.1.5. The incomplete-TTM variable

Rationale for including the variable in the Tailored Huff Model

The incomplete-TTM variable is a control variable, adjusting the volume estimates for stores that has only been opened part of the TTM. If a store has only been opened for nine months of the TTM, it should receive about nine twelfth of the volume estimate it had received had it been opened during the entire TTM.

The distinction and interaction between the ramp-up variable and the incomplete-TTM variable

The ramp-up variable captures the lower volume that a newly opened store may receive from low awareness among customers. The incomplete-TTM variable captures that a store, which was only opened for, again, nine twelfth of the TTM, should have its volume reduced by about three twelfth over and above any reduction due to a lower ramp-up level.

Data and methodology

While adjusting for if a store has been opened for less than twelve months of the TTM, the incomplete-TTM variable also takes the industry's seasonality effect into account. This implies that the incomplete-

TTM variable will cut the estimate by more if the months in which a particular store was not opened were peak-season months than if they were non-peak-season months. The seasonality effect was calculated based on the P&Ls of four of the Company's fully ramped up stores. The technicalities of this, including the seasonality regression are presented in Appendix A1.

4.1.6. Excluded variables

Several variables that probably have an impact on store volume have been excluded because of resource and time constraints, and lack of data availability. Excluded variables include price, visibility and accessibility of a store. Price should have a negative relation to volume, but as current price variation in the industry is limited we deem the impact of the exclusion of a price variable of minor importance at this point.

4.1.7. Data and methodology for estimating volume by postal code

In the preceding sections we explained the variables that, by postulation, determine a store j 's percentage of the total volume emanating from postal code i , during the TTM, according to equation 4.1. To convert this percentage of volume into volume, the Tailored Huff Model multiplies the percentage of volume figures with the total volume emanating from the postal code in the TTM. To restate equation 4.2:

$$\left(\begin{array}{c} \text{Store } j\text{'s volume} \\ \text{from postal code } i \\ \text{in TTM} \end{array} \right)_{i,j} = \left(\begin{array}{c} \text{Store } j\text{'s percentage} \\ \text{of postal code } i\text{'s} \\ \text{volume in TTM} \end{array} \right)_{i,j} \times \left(\begin{array}{c} \text{Total volume} \\ \text{from postal code } i \\ \text{in TTM} \end{array} \right)_{i,j}$$

Thus, to apply equation 4.2 we need to estimate the total volume emanating from all 9,623 Swedish postal codes in our dataset.

To do this, we first calculated the total volumes for Swedish counties (Swe.: Län) 2011. Then we distributed a county's aggregate volume to its municipalities (Swe.: Kommun) in accordance with the percentage of a particular feature in the county registered in the respective municipality.^{6,7} Finally we distributed a municipality's volume to its postal codes in accordance with the percentage of the municipality's inhabitants residing in the respective postal code.⁸

We calculated the total volume of a county by dividing the Market Leader and the Company's joint volume in that county with their joint Swedish market share in 2011. The Market Leader's volume by county in 2011 is publicly available at their website. We could calculate the Company's volume by county in 2011 as we had access to all their store's monthly P&Ls. Market share data at the country level is publicly available at a government agency's website.

When we forecast future volume (to get forecasts of cash flow and NPV in the CFF-CB Model) we apply a per-annum market volume growth of 2 percent.

4.1.8. Methodology for estimating the parameters

In his original paper from 1962, Huff used an iteration technique to estimate the parameters of his model (see equation 2.1), based on survey data. We employ another type of iteration technique, suitable to the data available to us. Surveying a sufficient number of prospective customers would be prohibitive and most certainly not as reliable as using the actual data we have been granted access to by the Company. Instead of customer surveys, our actual data are the monthly P&Ls – including volumes – for 23 of the Company's stores.

⁶ This "particular feature" is concealed but this data is very reliable and publicly available at a government agency's website.

⁷ Postal code-to-municipality links were retrieved from Postnummerservice's website via Excel macro.

⁸ The number of inhabitants per postal code is publicly provided online by Statistiska centralbyrån.

We let the Tailored Huff Model compute estimates for all stores in Sweden for a particular set of parameters, for the TTM ending on October 31, 2012 (i.e., November 1, 2011–October 31, 2012). Then we calculate the absolute difference between estimated volume and actual volume for each of the 23 Company stores. Finally the absolute differences for these 23 Company stores are summarized into total absolute volume difference (“TAVD”). This procedure is iterated for different sets of values for the parameters, and the set of parameters that produces the minimum TAVD is selected. This process is illustrated below.

Table 4.1. Illustration of our iteration technique for estimating the parameters of the Tailored Huff Model

		ITERATION 1			ITERATION 2		
		parameter s	0.0		parameter s	1.0	
		parameter d	1.0		parameter d	2.0	
		parameter r	1.0		parameter r	0.5	
Store	Actual TTM volume (store P&Ls)	Estimated TTM volume (Tailored Huff Model)	Difference	Absolute difference	Estimated TTM volume (Tailored Huff Model)	Difference	Absolute difference
Store A	13	8	-5	5	14	1	1
Store B	10	15	5	5	9	-1	1
...	...	...	...	...	...	...	...
Store Z	15	10	-5	5	17	2	2
		TAVD*	15*		TAVD*	4*	

* Total absolute volume difference is what we minimize in order to estimate the parameters of the Tailored Huff Model

We perform two rounds of iterations. First a round with broad ranges for the parameter values to roughly ring fence the minimum. Then, a second round with narrow ranges for the parameter values to confine the minimum further. The ranges for the parameter values in the second round are set around those parameter values which produced the minimum TAVD in the first round.

In the first iteration round parameter s goes from 0.0 to 2.4 in steps of 0.3, parameter d goes from 1.0 to 5.0 in steps of 0.5, and parameter r goes from 0.00 to 1.00 in steps of 0.25 (405 iterations). In the second iteration round parameter s goes from 0.80 to 1.00 in steps of 0.05, parameter d goes from 1.5 to 3.5 in steps of 0.1, and parameter r goes from 0.00 to 0.50 in steps of 0.05 (1,155 iterations).

The iterations are run in Stata and all 1,560 iterations take approximately 20 hours to run on a normal PC.

Our estimation technique is similar to those employed by Bucklin (1971) and Huff (1962). The former minimizes the sum of the squared deviations between actual and estimated values. Huff (1962) minimizes correlation in his original study.

Our estimation technique is not able to statistically validate our estimated parameters. This is a common problem in applying Huff models which we elaborate on in the robustness discussion in section 5.3.

The parameters of the Tailored Huff Model are estimated, and the results are discussed, in section 5. The volume estimates computed with the estimated parameters are then, indirectly, evaluated in section 6 where we compare the cash flow forecasts by the CFF-CB model to actual cash flows. As the estimated cash flows by the CFF-CB Model depend on the volume estimates from the Tailored Huff Model, the accuracy of our cash flow estimates is also an evaluation of the estimated parameters and the specification of the Tailored Huff Model.

4.2. Building the cash flow forecasting capital budgeting model (the CFF-CB Model)

The next step in building the CFF-CB Model is to integrate the volume-estimating Tailored Huff Model with the NPV capital budgeting appraisal technique. By performing this integration the CFF-BC Model is itself able to *compute* cash flows, basic NPV and comprehensive NPV for a proposed new store if that

store's coordinates are inputted into the model. In other words, the model does not rely on case-by-case, experience-based cash flow forecasts being inputted by management – which is the procedure in the Company today, and a procedure which to date has produced cash flow forecasts that are poor in accuracy and are optimistically biased.

A key feature of the CFF-CB Model is the ability to compute “comprehensive NPV” for a proposed new store. With comprehensive NPV we mean basic NPV of a proposed new store less any lost NPV of already existing stores due to cannibalization effects of introducing the proposed new store. This feature is made possible as the CFF-CB Model takes the entire competitive landscape into account in its capital budgeting for new stores.

In this section we explain the CFF-CB Model. First, we explain the integration of the Tailored Huff Model with the NPV capital budgeting appraisal technique. Second, we introduce the concept of comprehensive NPV and explain the method for calculating it. Third, we account for the data assumptions used in the CFF-CB Model to calculate cash flow and comprehensive NPV. Finally, we present how the CFF-CB Model is tested empirically.

4.2.1. Integrating the Tailored Huff Model with the NPV capital budgeting appraisal technique

The standard NPV capital budgeting appraisal technique is selected due to its acknowledgement theoretically and widespread use in practice (Brunzell *et al.*, 2011; Ryan, 2002; Graham and Harvey, 2002; Brounen *et al.*, 2004). Cash flow is calculated according to a standard method, as described by for example Shrieves and Wachowicz Jr. (2007) and we make no attempt to improve the technicalities of the NPV technique in this thesis.

The forecast period is set to three years, which is the time required for all stores in the Tailored Huff Model to become fully ramped-up and reach steady state. The CFF-CB Model uses the Tailored Huff Model to forecast volume, which is the main driver of the cash flow and NPV forecasts.

See Appendix A2 for a presentation of the equations used to calculate cash flow and NPV. Further, see section 4.2.3 below for an overview of the assumptions and data used in the CFF-CB Model.

4.2.2. Computation of comprehensive NPV for a proposed new store

Introducing the comprehensive NPV concept

When using the CFF-CB Model for capital budgeting for a proposed new store, comprehensive NPV is the key metric used for investment appraisal. To calculate comprehensive NPV, two supporting metrics are required: i) basic NPV of the proposed new store and ii) the NPV cannibalization on existing stores (due to the introduction of the proposed new store):

$$\begin{aligned} \text{Comprehensive NPV} = \\ & \text{Basic NPV of proposed new store} \\ & - \text{NPV cannibalization on existing stores} \end{aligned}$$

where:

$$\begin{aligned} \text{Basic NPV of proposed new store} = \\ \text{NPV of cash flows for proposed new store}^9 \end{aligned}$$

⁹ With basic NPV we mean NPV that does not take cannibalization effects into account. The term “basic” is used to distinguish it from comprehensive NPV (defined as NPV that does take cannibalization effects into account).

and:

$$\begin{aligned} \text{NPV cannibalization on existing stores} = \\ \text{NPV of existing stores including the proposed new store in the competitive landscape} \\ - \text{NPV of existing stores excluding the proposed new store in the competitive landscape} \end{aligned}$$

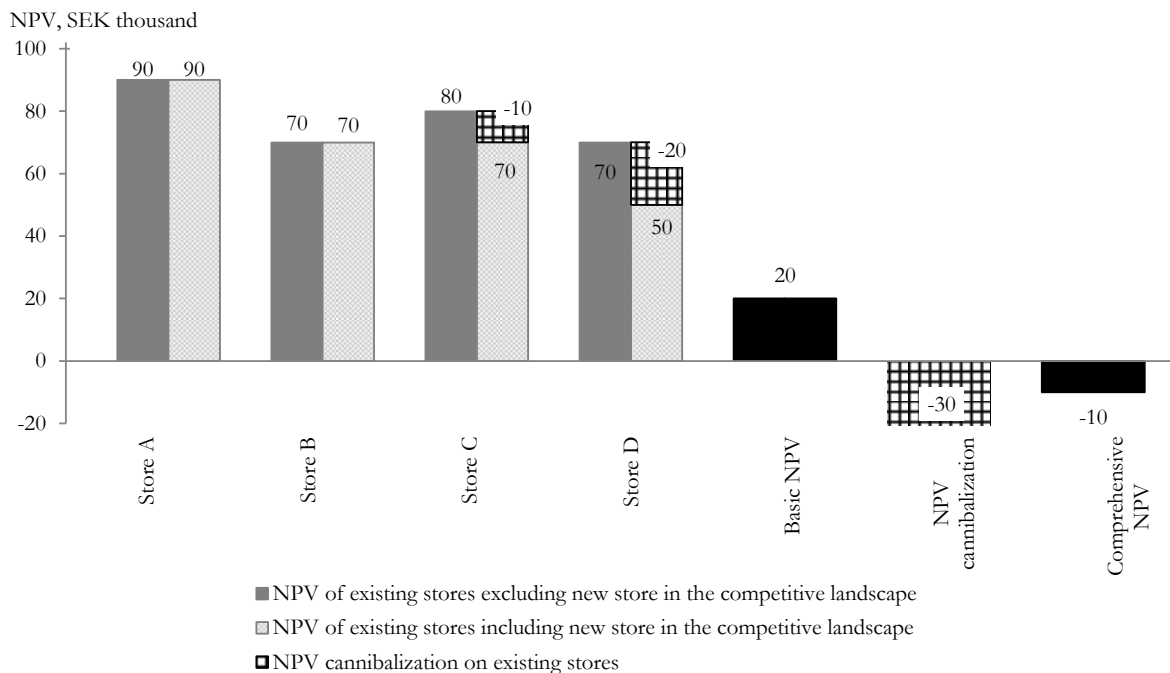
“Basic NPV of new store” is the traditional NPV only based on the cash flows related to the proposed new store. “NPV cannibalization on existing stores” occurs if the new store is located close to existing stores.

Cannibalization can be calculated as the Tailored Huff Model forecasts the volume for all stores in the market. By including a new store in the competitive landscape, the Tailored Huff Model’s volume forecasts – and cash flow and NPV forecasts in the CFF-CB Model – for nearby existing stores will decrease due to cannibalization.

Next, we will illustrate the intuition of the comprehensive NPV concept, and thereafter dig deeper into its technicalities.

Below is a graphical illustration of the comprehensive NPV concept. The example is based on arbitrary volume figures to arrive at cash flow and NPV, and not on actual forecasts from the CFF-CB Model. The example is based on a fictional company that has four existing stores and is contemplating to invest in a proposed new store.

Figure 4.8. Graphical illustration of the comprehensive NPV concept



In the above example, the NPVs of the existing stores A and B are unaffected by including the proposed new store in the competitive landscape as they are located far away from it. However, the combined NPV of stores C and D decreases by 30 as a result of NPV cannibalization from the new store. Therefore, while basic NPV of introducing the new store is positive at 20, comprehensive NPV is negative at -10 (20 – 30). The output from the CFF-CB Model reveals that the new store investment proposal should be accepted based on basic NPV, but rejected based on comprehensive NPV.

We will now use the above fictional example to explain the methodology of how the CFF-CB Model computes basic NPV and NPV cannibalization on existing stores to get to comprehensive NPV.

Computation of basic NPV for a proposed new store

The CFF-CB Model computes basic NPV by performing the cash flow and NPV calculations specified in Appendix A2. The volume estimates are attained from running the Tailored Huff Model with a proposed new store added to the competitive landscape for years 1–3. The other input data is attained as specified in table 4.2.

Below is an illustration of the basic NPV calculation for the same proposed new store as illustrated in figure 4.8.

Table 4.2. Basic NPV output for the fictional new store in figure 4.8, as calculated by the CFF-CB Model

<i>SEK thousand unless otherwise stated</i>	Year 0	Year 1	Year 2	Year 3
Volume from the Tailored Huff Model		60	89	105
Volume (product A), % of total volume		73%	73%	73%
Total volume		82	122	143
Average price per total volume, SEK		290	290	290
Sales		23.8	35.3	41.5
Personnel expenses (50% of sales)		-11.9	-17.6	-20.7
Non- Personnel expenses (20% of sales)		-4.8	-7.1	-8.3
EBITDA		7.1	10.6	12.4
Depreciation, (10% of sales)		-2.4	-3.5	-4.1
EBIT		4.8	7.1	8.3
Tax on EBIT (26.3%)		-1.3	-1.9	-2.2
+ Interest tax shield		0.0	0.0	0.0
+ Depreciation (10% of sales)		2.4	3.5	4.1
- ΔNWC		0.0	0.0	0.0
- CAPEX (10% of sales)		-2.4	-3.5	-4.1
- Initial store CAPEX	-50.6	-	-	-
Cash flow to equity	-50.6	3.5	5.2	6.1
Cost of equity		10%	10%	10%
NPV of cash flow to equity, forecast period	-50.6	3.2	4.3	4.6
NPV of cash flow to equity, terminal value	58.5			
Basic NPV	20.0			

Computation of NPV cannibalization on existing stores

We defined the “NPV cannibalization on existing stores” as:

NPV cannibalization on existing stores = NPV of existing stores *including* the new store in the competitive landscape – NPV of existing stores *excluding* the new store in the competitive landscape

To calculate NPV cannibalization on existing stores the CFF-CB Model performs two computations. i) A valuation of all existing stores in which the proposed new store is *excluded* from the competitive landscape; and, ii) a valuation of all existing stores in which the proposed new store is *included* in the competitive landscape. The difference between them is NPV cannibalization effect on existing stores.

The CFF-CB Model forecasts volume for the three-year explicit forecast period by coupling a store’s *actual* TTM volume from store P&Ls, with the *forecasted* volume growth from the Tailored Huff Model:

$$\text{Estimated volume Year 1} = \text{Actual volume Year 0} \times (1 + \text{estimated volume growth Year 1})$$

$$\text{Estimated volume Year 2} = \text{Estimated volume Year 1} \times (1 + \text{estimated volume growth Year 2})$$

$$\text{Estimated volume Year 3} = \text{Estimated volume Year 2} \times (1 + \text{estimated volume growth Year 3})$$

Why are we using actual volumes and forecasted volume growths? Clearly we would not want to use the volume forecasts from the Tailored Huff Model when we have useful, actual accounting data for *actual* TTM volume. But for future volume development we do not have any useful accounting data. To extrapolate the historical growth rate is not relevant if we are opening up a new store close by. For the future volume growth, then, we rely on the Tailored Huff Model.¹⁰

This approach for estimating volume is illustrated below in table 4.3 for an existing store that suffers from cannibalization when a new store is opened. The store represents fictional store C in figure 4.8. The growth rates in the table have been set arbitrarily to illustrate the calculation and are thus not based on any real forecasts from the CFF-CB Model.

Table 4.3. Example of volume estimates used in the CFF-CB Model to calculate NPV cannibalization on example store C in figure 4.8

<i>SEK thousand unless otherwise stated</i>	CFF-CB Model without new store				CFF-CB Model with new store			
	TTM	Year 1	Year 2	Year 3	TTM	Year 1	Year 2	Year 3
Accounting data from P&Ls								
Actual TTM volume	108	n.a.	n.a.	n.a.	108	n.a.	n.a.	n.a.
Growth	n.m.	n.a.	n.a.	n.a.	n.m.	n.a.	n.a.	n.a.
Tailored Huff Model								
Volume	95	97	99	101	95	89	86	87
Growth	n.m.	2.0%	2.0%	2.0%	n.m.	-6.0%	-4.0%	2.0%
Volume used in the CFF-CB Model								
Volume forecasts	108	110	112	115	108	102	97	99
Tailored Huff Model volume growth	n.m.	2.0%	2.0%	2.0%	n.m.	-6.0%	-4.0%	2.0%
CFF-CB Model NPV								
CFE		6.4	6.6	6.7		5.9	5.7	5.8
Growth		2.0%	2.0%	2.0%		-6.0%	-4.0%	2.0%
NPV	80				70			
NPV cannibalization on example store C	-10	(70 - 80)						

Note that excluding the proposed new store, the Tailored Huff Model estimates a volume growth of 2 percent, which is the aggregated market volume growth we use. If the new store is included, the existing store is estimated to have a negative volume growth during the ramp-up period of the new store. The growth rates from the Tailored Huff Model are applied to the base year's (TTM) actual volume of 108 to get the volume forecasts used by the CFF-CB Model to calculate cash flow and NPV.

Cash flow and NPV is then calculated according to the same equations and data assumptions as presented in Appendix A2. In the example above, NPV for the existing store including the proposed new store in the competitive landscape is 80, and NPV excluding the proposed new store in the competitive landscape is 70. The difference of -10 is the NPV cannibalization on the existing store.

The calculation in table 4.3 is then conducted for all existing Company stores. Returning once again to our fictional example in figure 4.8, the NPV cannibalization calculation is therefore also conducted for store A, B and D. The aggregate NPV cannibalization on existing stores is just a summary of the NPV cannibalization on each separate store as described in the below table.

¹⁰ This approach applies only to stores which have been opened for at least twelve months. For stores which have been opened less than twelve months we use the absolute volume estimates from the Tailored Huff Model, because we deem actual volume to be unreliable for these newly opened stores.

Table 4.4. Summarizing NPV cannibalization on each existing store, from figure 4.8, into NPV cannibalization on all existing stores

<i>SEK thousand</i>	NPV without new store	NPV with new store	NPV cannibalization on existing stores
Store A	90	90	0
Store B	70	70	0
Store C	80	70	-10
Store D	70	50	-20
NPV cannibalization on existing stores	310	250	-30

The aggregate NPV cannibalization on existing stores is finally deducted from the proposed new store's basic NPV in order to get comprehensive NPV.

4.2.3. Data assumptions for cash flow and NPV calculations

The below table summarizes the input variables, input values and data sources used in the CFF-CB Model to compute cash flows and ultimately comprehensive NPV. (See Appendix A2 for formulas).

Table 4.5. Key input variables, values and data sources in the CFF-CB Model

Input variable	Input value used	Data source
Volume (product A) for existing stores	Combination of Tailored Huff Model estimates and P&L data (explained in section 4.1)	Store P&L and Tailored Huff Model estimates ¹¹
Volume (product A) for proposed new store	Estimated via the Tailored Huff Model	Tailored Huff Model estimates ¹¹
Volume (product A), % of total volume	73%	Company data
Price per total volume	290	Company data
Personnel expense, % of sales	50%	Store P&L ¹²
Non-personnel expense, % of sales	20%	Store P&L ¹²
Depreciation, % of sales	10%	Store P&L ¹²
Interest payment on debt	0	Arbitrarily set
Change in net working capital	0	Arbitrarily set
CAPEX, % of sales	10%	Arbitrarily set
Initial CAPEX for hypothetical new store	SEK 6 million	Arbitrarily set
Tax	26.3%	Swedish statutory tax rate
Price growth in forecast period	0%	Arbitrarily set
Steady state growth	2%	Arbitrarily set
Net debt	0	Arbitrarily set
Cost of equity	10%	Arbitrarily set

For brevity and focus we use the same input values for all stores, with the exception of volume, which is forecasted by the Tailored Huff Model.

¹¹ The data sources for the estimates of the Huff Model are specified in section 4.1.

¹² Input values from store P&Ls represent average data for fully ramped-up stores.

4.2.4. Method for testing and exploring the CFF-CB Model

Reconnecting to the *case* purpose of improving the Company's cash flow forecasts in its capital budgeting with a comprehensive model, we perform two small-scale empirical back-tests of the CFF-CB Model. First, we test the model's cash flow forecasts at a point in time for investigating accuracy and bias in estimates. Second, we test its cash flow growth forecast accuracy in a longitudinal six-month test. Supported by the results from these two empirical tests, we finally explore the comprehensive NPV concept illustratively.

Back-test 1: Evaluating the cash flow forecasts at a point in time – accuracy and bias of forecasts

First we back-test the CFF-CB Model by comparing its cash flow forecasts against actual cash flow and the Company's *ex ante* cash flow forecasts. The back-test comprises the twelve-month period between November 1, 2011 and October 31, 2012. The purpose is to evaluate and discuss the accuracy and bias of the model's estimates.

To be able to compare the CFF-CB Model's cash flow forecasts, we need the actual cash flows and Company's *ex ante* cash flow forecasts. Actual cash flow is calculated by taking actual volume, and the same cash flow assumptions and methodology as for the CFF-CB Model (see table 4.5). Similarly, Company cash flow forecasts is calculated by taking *ex ante* Company volume forecasts, and the same cash flow assumptions and methodology as for the CFF-CB Model (see table 4.5). Hence, if the CFF-CB Model's cash flow forecast error compared to actual cash flow is X percent for a store, its volume forecast error would be X as well. The same relationship holds for the Company's cash flow forecasts.

The above approach is taken for two reasons. First, because most stores were only recently opened, their historical ratios are low and volatile and probably not meaningful indicators of future ratios. Second, we want to isolate the differences between actual and estimated volume as this is the key driver of cash flow and value in the industry.

Back-test 2: Evaluating the forecasting of growth in cash flow – development of forecasts over time in the changing competitive landscape

In a second back-test we compare the growth in actual cash flows with the growth of the CFF-CB Model's forecasted cash flows, from April 30, 2012 to October 31, 2012. The test is done for Company stores that had been opened at least twelve months as of April 30, 2012. The purpose is to evaluate how the CFF-CB Model performs in taking the rapidly changing landscape into account when forecasting the growth rates of stores.

Reliable growth forecasts are particularly important for computing comprehensive NPV as the Tailored Huff Model's volume growth forecasts are used to calculate NPV cannibalization, illustrated in table 4.3.

Exploring the comprehensive NPV concept

Finally, we explore and discuss the comprehensive NPV concept by applying the CFF-CB Model in capital budgeting for a hypothetical new Company store. We place the hypothetical new store in close proximity to another Company store in a mid-sized Swedish city. In addition, two Market Leader stores are located in the same city.

5. Results from estimating the parameters of the Tailored Huff Model

As described in section 4.1.8 and table 4.1, we perform two rounds of iterations to find the approximate minimum total absolute volume differences (TAVD) between the estimated volume and the actual volume for 23 of the Company's stores. First an iteration round with broad ranges for the parameter values to roughly ring fence the minimum. Then, a second round with narrow ranges for the parameter values to confine the minimum further. The ranges for the parameter values in the second round are set around those parameter values which produced the minimum TAVD in the first round.

5.1. The first, broad, iteration round

In the first iteration round parameter s goes from 0.0 to 2.4 in steps of 0.3, parameter d goes from 1.0 to 5.0 in steps of 0.5, and parameter r goes from 0.00 to 1.00 in steps of 0.25 (405 iterations).

The minimum TAVD is 27,215, found for $s = 0.9$, $d = 2.0$, $r = 0.25$. This minimum is displayed in figure 5.2. Figure 5.1 and 5.3 show the lowest TAVDs for $r = 0.0$ and $r = 0.5$, and we notice that these values are larger than for $r = 0.25$. The lowest TAVD values are retrieved tightly around $s = 0.9$, but for a broader range for d and r . These observations served as a basis for setting up the second iteration round.

Figure 5.1. The lowest TAVD for $r = 0.00$ is obtained for $s = 0.9$ and $d = 3.5$

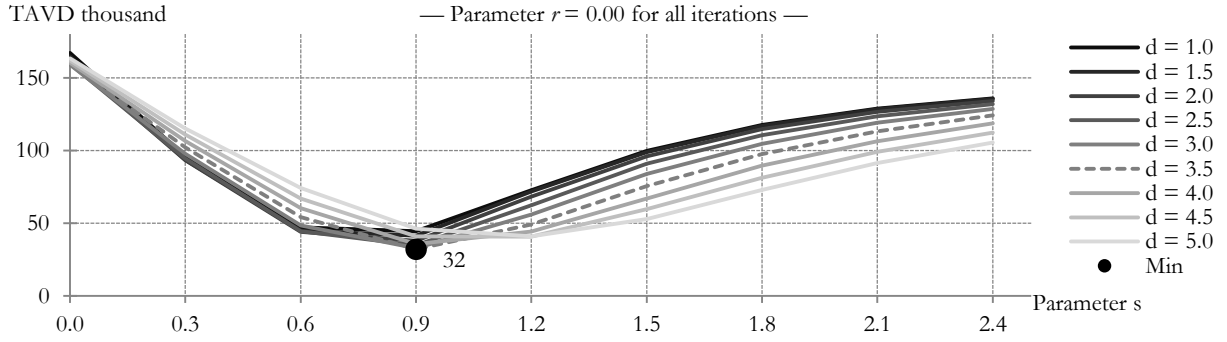


Figure 5.2. The minimum TAVD for the first iteration round is obtained at $s = 0.9$, $d = 2.0$ $r = 0.25$

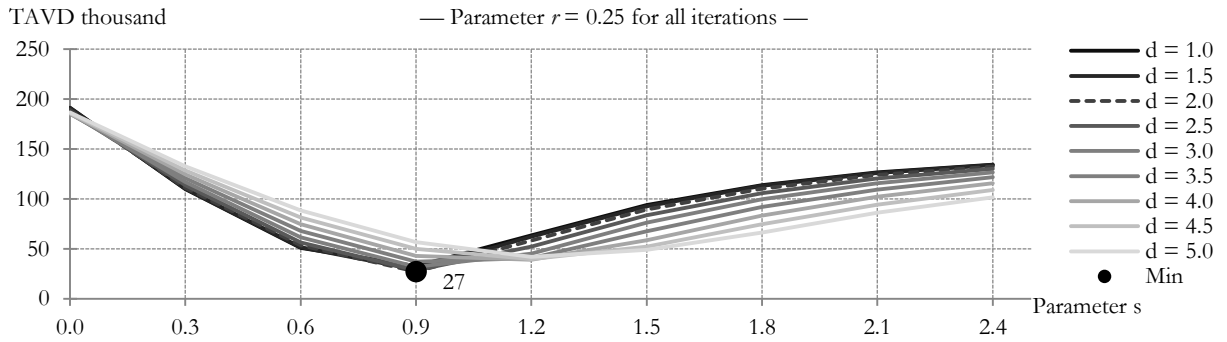
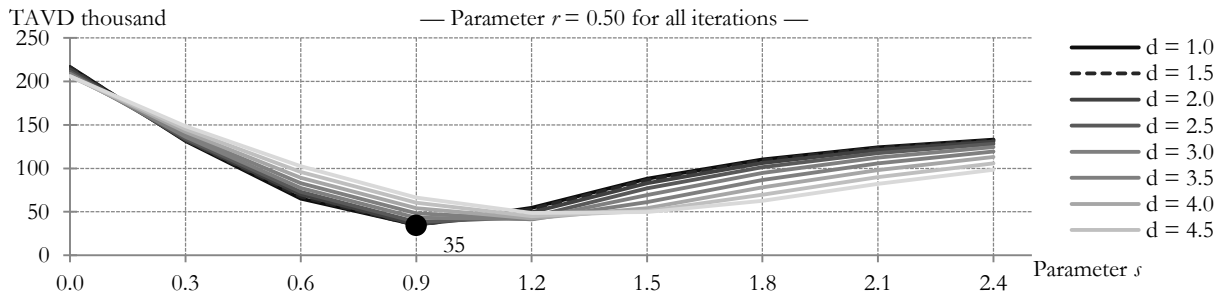


Figure 5.3. The lowest TAVD for $r = 0.50$ is obtained for $s = 0.9$ and $d = 1.5$

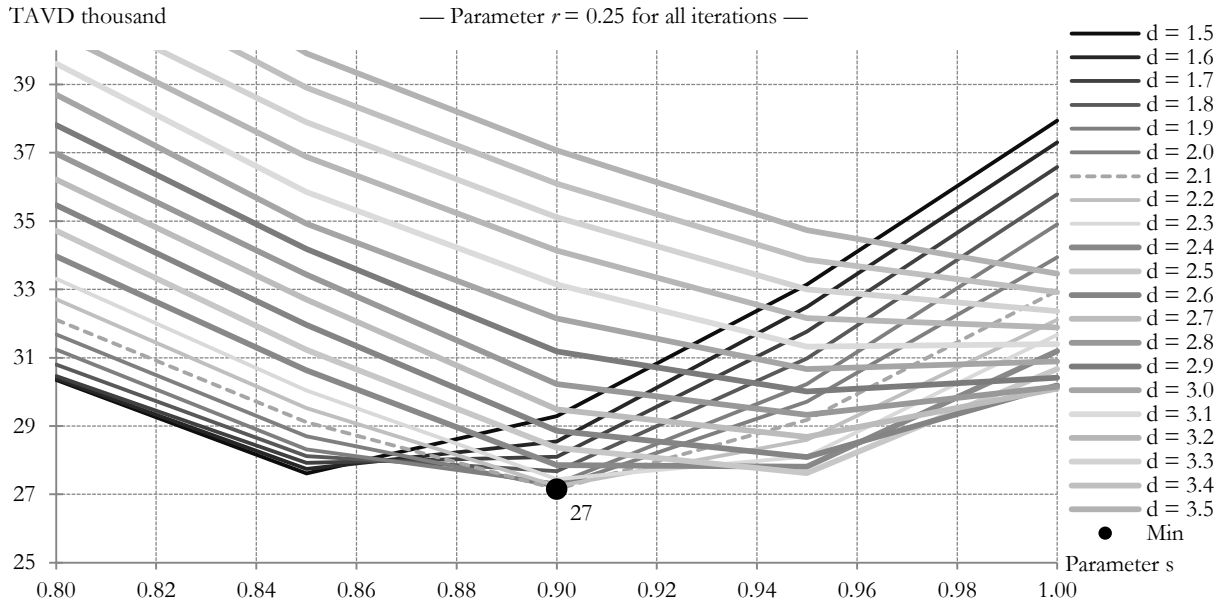


5.2. The second, narrow, iteration round

In the second iteration round parameter s goes from 0.80 to 1.00 in steps of 0.05, parameter d goes from 1.5 to 3.5 in steps of 0.1, and parameter r goes from 0.00 to 0.50 in steps of 0.05 (1,155 iterations).

The minimum TAVD is 27,150, found for $s = 0.9$, $d = 2.1$, $r = 0.25$. This minimum is only slightly lower than the minimum found in the first iteration round and d is the only parameter that is refined, from 2.0 to 2.1. The minimization is illustrated below.

Figure 5.4. The minimum TAVD for the second iteration round is obtained at $r = 0.25$, $s = 0.9$ and $d = 2.1$



The estimation of parameter s at 0.9 indicates a considerable old-industry-structure effect, that is, customers are more likely to patronize an old Market Leader stores than a newly opened store, all else equal. This makes logical sense as customers are aware of the existence of these stores, used to visiting them for years, and in general satisfied with the service provided by the Market Leader in the past.

The magnitude of the old-structure-stickiness effect may erode over time and the s parameter should be re-estimated periodically to accommodate for this.

The estimation of parameter d at 2.1 indicates, not surprisingly, that driving time is a very important factor for predicting consumer geographic behavior in the case industry. For example, if stores A and B are 20 and 30 minutes, respectively, away from a postal code in terms of driving time – all else equal and no other competing stores existing – stores A receives 70.1 percent of that postal code's volume compared to 29.9 percent for store B.

The 2.1 value for the d parameter is roughly in line with other research. In Huff's (1962) original paper, he estimated his similar λ parameter for travel distance at between 2.1 and 3.8, for different neighborhoods and different commodities. Others have set similar parameters to 2.0, based on economic reasoning (Xu and Liu, 2004; Huff 1962).

Then, economic reasoning and prior research indicates that our 2.1 value for parameter d is reasonable.

The estimation of parameter r at 0.25 indicates a considerable ramp-up effect in the case industry where a store has an average TTM ramp-up level of 25 percent on its opening day and then increases its ramp-up level gradually over 30 months to 100 percent.

5.3. Robustness discussion

Our estimation technique is not able to statistically validate our estimated parameters with, for example, t-statistics for the parameters or an adjusted-R² value for the model. This is a widespread problem in applying the Huff model or variations of it. Initially there was not even a theoretical statistical method available for estimating the parameters, due to the nonlinear properties of the model (Huff, 1962). This problem was later solved by transforming the model into log form (Nakanishi and Cooper, 1974; Nakanishi and Cooper, 1982). Since then, the problem of statistical validation has instead consisted of the practicalities of handling the vast amount of data necessary. This practicality problem is gradually being overcome with improvements in statistical and other software, but statistical validation is still very rare (Huff, 2008).

Another issue is the reliability of the data underlying the parameter estimation of the Tailored Huff Model. Much of the data we have used have a high degree of reliability, such as coordinates of stores and postal codes, driving times, opening dates of stores and Company P&L data. Other data are less reliable, such as the number of x-attributes of old Market Leader stores and the distribution of actual volume onto postal codes as described in section 4.1.7.

A final issue is that the values of the estimated parameters are likely to vary between different geographical regions. For example, old-structure stickiness may be stronger in areas where consumers are more favorable to the Market Leader and thus the s parameter should be higher there. The magnitude of the distance-decay parameter d , is likely to differ between cities and less densely populated areas – an issue raised by for example Bucklin (1971) and Fotheringham (2010). The estimated parameters of the Tailored Huff Model are, therefore, best seen as approximations of the averages for the case industry as a whole.

6. Results from applying the CFF-CB Model

6.1. Back-test 1: Evaluating the cash flow forecasts at a point in time – accuracy and bias of the forecasts

Table 6.1 presents the results from the point-in-time back-testing of the CFF-CB Model for the TTM ending on October 31, 2012 (i.e. November 1, 2011 to October 31, 2012). The table shows the CFF-CB Model's *ex post* cash flow forecasts, the Company's *ex ante* cash flow forecasts, and the actual cash flows – and below we discuss the differences between these.¹³

6.1.1. Forecast accuracy

For stores opened more than twelve months, the average absolute forecast error (an estimate's absolute percentage deviation against actual) of the CFF-CB Model is 20 percent compared to an average absolute forecast error of 78 percent for the Company. The average absolute forecast error of the CFF-CB Model appears to be within reasonable accuracy "on its own", but particularly when considering that it is 58 percentage points better than that of the Company.

For stores opened less than twelve months, the average absolute forecast error of the CFF-CB Model is 12 percent. If we assume that this is a reasonable absolute forecast error, it would lend support to the ways the ramp-up and incomplete-TTM variables are constructed as newly opened stores can be handled by the model with decent accuracy. Because the Company forecasts cash flows annually, no meaningful comparison with the Company is possible for stores opened less than 12 months.

For all stores, the average absolute forecast error of the CFF-CB Model is 17 percent.

The correlation between forecasted and actual cash flow is 0.55 for the CFF-CB Model and 0.37 for the Company. This indicates that the relative ranking of the estimated cash flows are positively correlated to the actual relative ranking – even though this positive correlation is rather weak.

Outliers

For stores opened more than twelve months, the four largest forecast errors of the CFF-CB Model are all under-estimations – stores C, F, H, and R. Three of these stores – stores F, H, and R – are the best performing stores of the Company. This indicates that there may be upside-potential factors that are important for volume and cash flow, that are not included in the Tailored Huff Model. These factors could be pricing of close-by competitors, visibility and accessibility of a store or above-normal marketing efforts. The outliers highlight that the CFF-CB Model should not be applied without judgment and consideration of additional factors to those included in the Tailored Huff Model.

6.1.2. Bias of estimates

The CFF-CB Model's average non-absolute forecast error (an estimate's, non-absolute, percentage deviation against actual) is –7 percent for all stores, –6 percent for stores opened more than 12 months, and –8 percent for stores opened less than 12 months. Forecasts are thus slightly negatively biased, but close to unbiased. Given that the Tailored Huff Model distributes the aggregate market volume onto all stores in the market, close to unbiased cash flow estimates are expected and this is one of the CFF-CB Model's key benefits.

¹³ Recall that actual TTM cash flow and Company TTM cash flow forecasts are calculated according to the same assumptions and methodology as for the CFF-CB Model's cash flow. Cash flow is thus calculated according to the same assumptions, as presented in table 4.5. Based on these assumptions, if the Company or the CFF-CB Model cash flow forecast error compared to actual cash flow is X percent for a store, the percentage difference in volume will be the same.

The Company's average forecast error, for stores opened more than 12 months, is 78 percent – which is equal to the Company's average absolute forecast error. This implies that the Company is consistently optimistically biased in its cash flow forecasting, in line with previous research on capital budgeting cash flow forecasting (Pruitt and Gitman, 1987; Statman and Tyebjee, 1985; Flyvbjerg *et al.*, 2007).

Optimistically biased cash flow forecasts are less of an issue if most potential new stores are NPV-positive investments, which may be the case in the case industry currently as the reregulation occurred recently. However, as the industry matures, potential new stores with positive NPV will become more difficult to find. Optimistically biased forecasts can then lead to overinvestment and misallocation of capital. A positive feature of the CFF-CB Model is thus that its relatively unbiased forecasts reduce the risk of investing in negative NPV projects.

Table 6.1. Empirical results from the point-in-time back-test of the CFF-CB Model

The table presents the empirical results from running a point-in-time back-test of the CFF-CB model between November 1, 2011 and October 31, 2012. “% Error compared to actual” is calculated as the cash flow forecast (by the CFF-CB Model *ex post* or by the Company *ex ante*) less actual cash flow divided by the actual cash flow. “% Absolute error compared to actual” is the absolute value of “% Error compared to actual”.

To be able to compare the CFF-CB Model’s cash flow forecasts, we compute actual cash flows and Company cash flow forecasts based on the following assumptions. Actual cash flow is calculated by taking actual volume, and the same cash flow assumptions and methodology as for the CFF-CB Model (see table 4.5). The Company’s *ex ante* cash flow forecasts are calculated by taking the Company’s volume forecasts, and the same cash flow assumptions and methodology as for the CFF-CB Model (see table 4.5). The assumptions for calculating CFF-CB Model, actual and Company forecasted cash flows are linearly proportional to volume. Hence, if the CFF-CB Model’s cash flow forecast error compared to actual cash flow is X percent for a store, its volume forecast error would be X as well. The same relationship holds for the Company’s cash flow forecasts.

Further note that only annual Company cash flow forecasts are available. Company cash flow forecasts for stores opened less than 12 months are therefore not meaningful (n.m.).

Store	Store opened	TTM cash flows in SEK thousand			% Error compared to actual		% Absolute error compared to actual	
		Actual	CFF-CB Model	Company forecasts	CFF-CB Model	Company forecasts	CFF-CB Model	Company forecasts
Store S	≥ 12 months	591	582	807	-1%	37%	1%	37%
Store M	≥ 12 months	389	395	824	2%	112%	2%	112%
Store T	≥ 12 months	545	565	681	4%	25%	4%	25%
Store J	≥ 12 months	449	475	742	6%	65%	6%	65%
Store B	≥ 12 months	305	345	818	13%	169%	13%	169%
Store D	≥ 12 months	633	546	1,065	-14%	68%	14%	68%
Store E	≥ 12 months	388	453	810	17%	109%	17%	109%
Store K	≥ 12 months	302	377	803	25%	166%	25%	166%
Store I	≥ 12 months	417	537	797	29%	91%	29%	91%
Store R	≥ 12 months	640	437	805	-32%	26%	32%	26%
Store F	≥ 12 months	802	531	890	-34%	11%	34%	11%
Store H	≥ 12 months	775	509	922	-34%	19%	34%	19%
Store C	≥ 12 months	423	201	935	-53%	121%	53%	121%
Sum	≥ 12 months	6,658	5,952	10,898	-73%	1,018%	263%	1,018%
Average	≥ 12 months	512	458	838	-6%	78%	20%	78%
Store V	< 12 months	183	183	n.m.	0%	n.m.	0%	n.m.
Store O	< 12 months	241	237	n.m.	-1%	n.m.	1%	n.m.
Store W	< 12 months	274	282	n.m.	3%	n.m.	3%	n.m.
Store P	< 12 months	279	270	n.m.	-3%	n.m.	3%	n.m.
Store N	< 12 months	144	139	n.m.	-3%	n.m.	3%	n.m.
Store Q	< 12 months	138	125	n.m.	-9%	n.m.	9%	n.m.
Store A	< 12 months	157	181	n.m.	15%	n.m.	15%	n.m.
Store G	< 12 months	266	211	n.m.	-21%	n.m.	21%	n.m.
Store L	< 12 months	85	66	n.m.	-23%	n.m.	23%	n.m.
Store U	< 12 months	90	54	n.m.	-40%	n.m.	40%	n.m.
Sum	< 12 months	1,857	1,748	n.m.	-83%	n.m.	120%	n.m.
Average	< 12 months	186	175	n.m.	-8%	n.m.	12%	n.m.
Sum	All	8,515	7,701	n.m.	-156%	n.m.	382%	n.m.
Average	All	370	335	n.m.	-7%	n.m.	17%	n.m.
Correlation between CFF-CB Model and actual cash flows, for stores opened more than 12 months								0.55
Correlation between Company forecasts and actual cash flows, for stores opened more than 12 months								0.37

6.2. Back-test 2: Evaluating the forecasting of growth in cash flow – development of forecasts over time in the changing competitive landscape

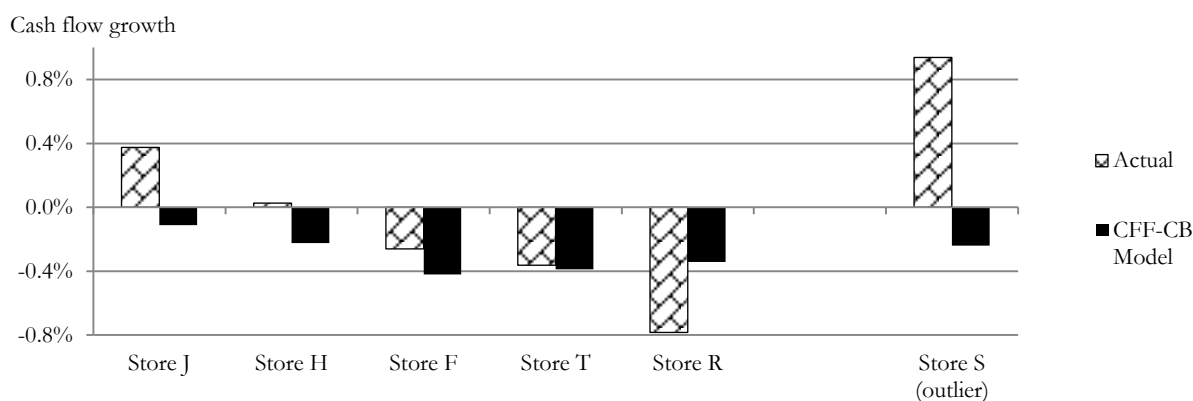
The table and figure below depict growth in actual cash flow and the CFF-CB Model's cash flow forecasts for the six-month period between April 30, 2012 and October 31, 2012. The stores included in the test had been opened for at least twelve months at April 30, 2012.

Table 6.2. Actual and estimated growth in cash flow April 30, 2012–October 31, 2012

To be able to compare the CFF-CB Model's cash flow forecast growth with actual, we compute actual cash flow based on the same assumptions as in table 4.5. Note that CFF-CB Model and actual cash flows for the TTM ending October 31, 2012 are the same as in table 4.5. The forecast error is calculated “CFF-CB Model cash flow growth” less “actual cash flow growth” divided by the “actual cash flow growth”.

<i>SEK</i> <i>thousand</i>	TTM cash flow ending Apr 30, 2012		TTM cash flow ending Oct 31, 2012		Growth Apr 30, 2012–Oct 31, 2012		
Store ID	Actual	CFF-CB Model	Actual	CFF-CB Model	Actual	CFF-CB Model	Forecast error
Store J	433	480	449	475	3.8%	-1.1%	-4.9%
Store H	773	520	775	509	0.3%	-2.2%	-2.5%
Store F	824	554	802	531	-2.6%	-4.2%	-1.6%
Store T	565	588	545	565	-3.6%	-3.9%	-0.3%
Store R	694	452	640	437	-7.8%	-3.4%	4.4%
Assumed outlier							
Store S (outlier)	9,245	10,214	10,114	9,968	9.4%	-2.4%	-11.8%
Average growth (including outlier):					-0.1%	-2.9%	
Average growth (excluding outlier):					-2.0%	-3.0%	
Correlation between actual and CFF-CB Model growth in cash flow (including outlier):							0.64
Correlation between actual and CFF-CB Model growth in cash flow (excluding outlier):							0.78

Figure 6.1. Actual and CFF-CB Model growth in cash flow during the six-month period April 30, 2012–October 31, 2012



The CFF-CB Model estimates have been calculated with a market volume growth of 2.0 percent per annum, and the fact that not all of the stores exhibit an estimated 1.0 percent growth in cash flow during the six-month review period is explained by the changes in the competitive landscape. The decline in estimated TTM cash flow has either gone to stores established during the review period, or stores established before the review period but whose ramped-up level increased. The estimated *average* cash flow growth for the stores, excluding one outlier, was estimated at –3.0 percent, compared to the actual of –2.0 percent.

The absolute deviation of forecasted growth in cash flows from actual growth is rather large, on average 2.7¹⁴ percentage points. But it appears to be a relationship between forecasted growth and actual growth in terms of the relative ranking of the stores' growth rates, as the correlation is 0.78¹⁴. This relationship is graphically illustrated in figure 6.1.

The growth-rate estimates for stores F, H, R deviates 2.8 percent from actual, slightly below average. Note that these were the same outliers – with an absolute forecast error of over 30 percent – we discussed under “Outliers” above. This indicates that even if a store's forecasted cash flow deviates greatly from actual, its forecasted cash flow *growth* could be fairly close to the actual growth rate. This lends support to our methodology for calculating comprehensive NPV, as cannibalization by a potential new store on existing stores is entirely dependent on estimated changes in existing stores cash flow growth.¹⁵

6.3. Hypothetical new store – exploring the comprehensive NPV concept

The empirical tests of the CFF-CB Model indicate a reasonable point-in-time and growth-rate accuracy of the cash flow forecasts – thus adding comfort to the method for calculating comprehensive NPV. Next we explore the comprehensive NPV concept for a hypothetical new store to illustrate the uses of the CFF-CB Model. The hypothetical new store is placed in a mid-sized Swedish city, close to the existing Company store F. In addition, the Market Leader operates two stores in the same city.

Figure 6.2. Illustration of comprehensive NPV for the hypothetical new store and its NPV cannibalization on four selected Company stores

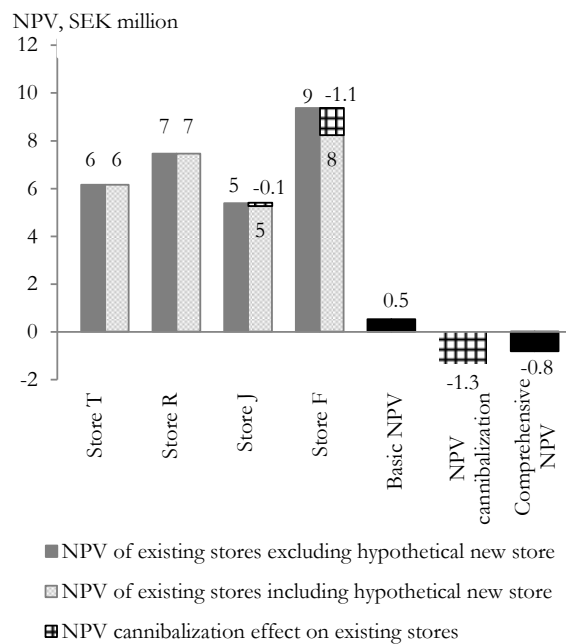


Figure 6.3. Development of cash flow forecasts for store F, and the hypothetical new store, during the three-year explicit forecast period

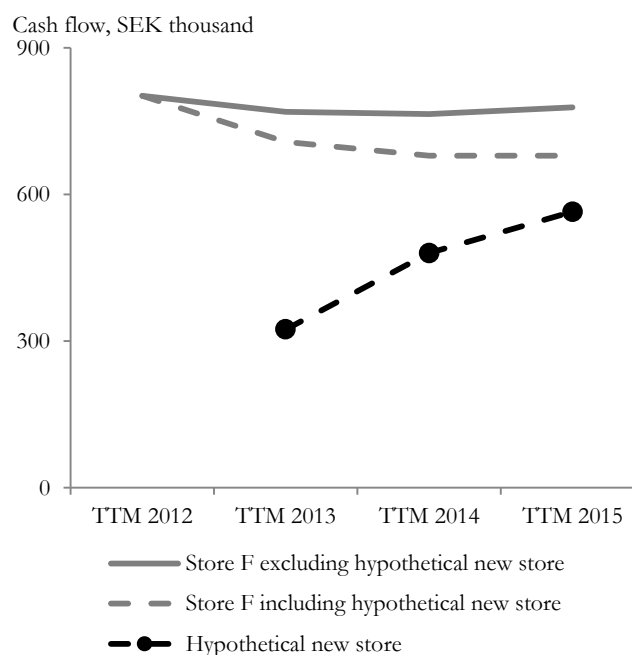


Figure 6.2 illustrates how the NPV forecasts for existing store J and F decrease as the hypothetical new store is added to the competitive landscape.¹⁶ The NPV cannibalization effect from the hypothetical new store on store F is further understood by looking at figure 6.3. Store F's cash flow forecasts decrease as the hypothetical new store ramps up. Meanwhile, NPV is unchanged for store T and R – all located over 5 hours in driving time from the hypothetical new store. The aggregate estimate for the hypothetical new

¹⁴ The value is calculated excluding store S which is assumed to be an outlier.

¹⁵ Recall formula: "Estimated volume Year 1 = Actual volume Year 0 * (1 + Estimated growth Year 1)".

¹⁶ The NPVs of store B, C, H, P, U, and S are also affected by NPV cannibalization of the hypothetical new store. The total NPV cannibalization on those stores is estimated at a negligible SEK -69 thousand.

store is a basic NPV of SEK 0.5 million, which less NPV cannibalization of SEK –1.3 million leads to a comprehensive NPV of SEK –0.8 million. Further see Appendix A2 for a spreadsheet of the comprehensive NPV calculation for the hypothetical new store as computed by the CFF-CB Model.

The hypothetical new store example serves as illustration of the benefits and uses of the CFF-CB Model. As basic NPV can be positive when comprehensive NPV is negative, the CFF-CB Model is useful for investment appraisal of proposed new stores. NPV cannibalization is less of an issue at the moment, when new stores still are opened rapidly, presumably with positive NPVs. However, as the market matures, positive NPV locations will become scarcer. Comprehensive NPV is then a good metric for avoiding overinvestment. The CFF-CB Model and comprehensive NPV is thus theoretically appealing as it takes the entire competitive landscape into account in the capital budgeting for proposed new stores.

6.4. Robustness discussion

We see two main issues to be addressed with the CFF-CB Model. The first relates to the evaluation method of empirically back-testing the model and the second relates to excluded variables.

The two back-tests were designed as small-scale empirical tests of the CFF-CB Model's ability to forecast cash flow. The CFF-CB Model's average point-in-time forecast error, for stores opened more than 12 months, is 58 percentage points better than the Company's forecasts. This figure appears to be considerably better. However, the CFF-CB Model is evaluated on *ex post* information, whereas the Company's forecasts are evaluated on *ex ante* information. The results are therefore not entirely comparable and can only provide an indication of the CFF-CB Model's forecasting ability. It would be valuable to also test the CFF-CB Model *ex ante*. We did not perform an *ex ante* test as we regard our dataset to cover too short of a time span.

As noted in section 4.1.6, the Tailored Huff Model excludes variables that are likely to have an impact on store volume. These variables include for example price, visibility and accessibility of a store, and a store's proximity to car clusters such as industrial areas. The exclusion of such variables probably has a meaningful negative impact on the Tailored Huff Model's ability to project volume estimates (and hence in the CFF-CB Model's ability to project cash flow forecasts). And as discussed under "Outliers" in section 6.1.1, we identified four stores as outliers, probably relating to excluded variables. However, measuring the intensity and weight of all excluded variables is difficult as useful data on most of these variables is difficult to obtain. In line with Huff's reasoning (1962) we believe that the variables included in the Tailored Huff Model should be able to capture the most important factors for enabling a reasonable accuracy and consistency of the forecasts by the CFF-CB Model. Then a good deal of management judgment should be applied to these forecasts in order to include additional qualitative factors.

Finally, we stress the fact that the concept of comprehensive NPV has only been illustrated and not tested in this study. The concept was not tested because of the previously mentioned problem of a too longitudinally limited dataset.

7. Conclusions

In this case study of a Company in a Swedish retail industry, we have built and tested the CFF-CB Model. This capital budgeting model itself computes forecasts of cash flow, basic NPV and comprehensive NPV for a proposed new store, if the store's coordinates are inputted into the model.

Our case purpose was to investigate whether the CFF-CB Model could improve the Company's case-by-case, experience-based cash flow forecasting in its capital budgeting – on three accounts: i) improve the accuracy of the forecasts, ii) mitigate the optimistic bias of the forecasts, and iii) induce a more holistic approach to capital budgeting through the concept of comprehensive NPV.

We investigated the accuracy and bias of the CFF-CB Model in two small-scale empirical back-tests. The CFF-CB Model's cash flow forecasts were compared to actual *ex post* cash flows, and to the Company's *ex ante* forecasts. The forecasts of the model are considerably more accurate than the Company's and do not exhibit any optimistic bias.

The comprehensive NPV concept was explored by letting the CFF-CB Model compute capital budgeting for a hypothetical new Company store within the actual competitive landscape of a mid-sized Swedish city. We showed that even though the hypothetical new store's basic NPV was positive, its comprehensive NPV was negative as it cannibalized on one of the Company's existing stores.

A primary cause for concern is that the Tailored Huff Model, which is used by the CFF-CB Model for forecasting volume, excludes variables that probably have a considerable impact on store volume.

In aggregate, our research has demonstrated a significant potential of using models for forecasting cash flow in capital budgeting. In particular, models have the potential to mitigate the well-documented problem of poor accuracy and optimistic bias in cash flow forecasts (Baker and English, 2011; Pruitt and Gitman, 1987; Statman and Tyebjee, 1985; Flyvbjerg *et al.*, 2007). In addition, we have introduced the theoretically appealing concept of comprehensive NPV. Hopefully, our study will inspire more research into this neglected field within financial accounting.

Appendix A. Supportive data and calculations

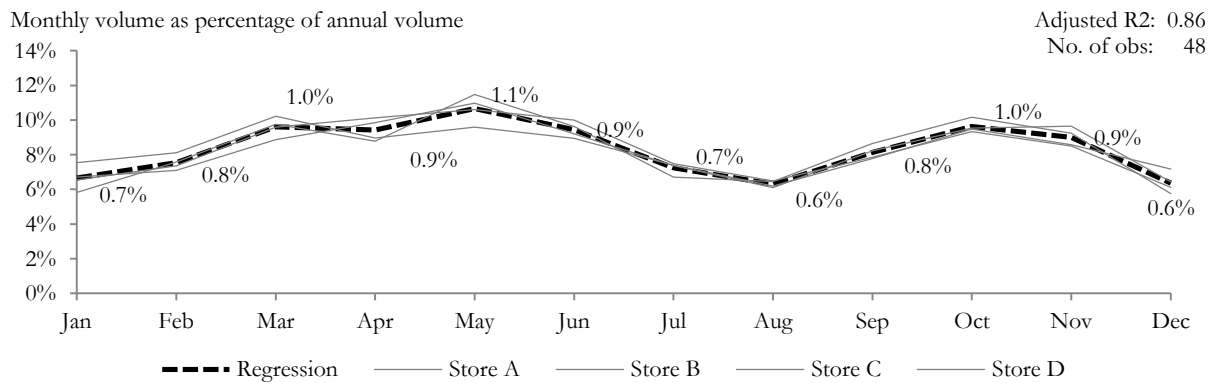
A1. Seasonality calculation for the incomplete TTM variable in the Tailored Huff Model

While adjusting for if a store has been opened for less than twelve months of the TTM, the incomplete-TTM variable takes the industry's seasonality effect into account.

The cumulative volume of the number of months a store has been opened during the TTM, represents a certain percentage of annual volume – the incomplete-TTM variable takes on this value. This value should be approximately 25 percent (three twelfth) if a store has been opened for three months in the TTM. However, we adjust for a rather distinctive seasonality trend in the industry. Each month's percentage of annual volume was determined by regressing month dummy variables on monthly volumes as percentages of annual volume, for four of the Company's fully ramped up stores (Wooldridge 2009, p. 368–369).

The regressed values for each month are used in the Tailored Huff Model for the incomplete-TTM variable. To exemplify, let's say that the Tailored Huff Model were to be run on October 31, 2012. A store which had been opened for three months at this point receive an incomplete-TTM value of October + September + August = 6.3 + 8.1 + 9.6 percent = 24.0 percent. Thus, if the store had received a volume estimate of 8,000 had it been opened during the entire TTM, it will now receive a volume estimate of 1,920 all else equal.

Figure A.1. Seasonality in the case industry



A2. Equations and data assumptions for the cash flow and NPV calculations in the CFF-CB Model

NPV is obtained from the below formula.

$$NPV_0 = Inv_0 + \frac{CFE_1}{(1+r_e)^1} + \frac{CFE_2}{(1+r_e)^2} + \frac{CFE_3}{(1+r_e)^3} + \frac{CFE_3*(1+g)}{(r_e-g)}$$

Equation A2.1

Where subscript (0, 1, 2 and 3) denotes year, and

Where:

Inv_0	=	Initial store capital expenditure
CFE_t	=	Cash flow to equity
r_e	=	Cost of equity
g	=	Steady-state CFE growth rate

Cash flow to equity (CFE) is calculated based on the below formula (see e.g. Shrieves and Wachowicz Jr., 2007):

$$CFE_t = \{[(S_t - Per_t - Nonper_t - Depr_t - Int_t) * (1 - Tax_t)] + Depr_t\} - \{\Delta Capex_t + \Delta WC_t\} + \Delta B_t$$

$$Cash\ flow\ to\ equity_t = \{Net\ income\ after\ tax_t + Depreciation_t\} - \{Total\ net\ investment_t\} + net\ debt\ issuance_t$$

Equation A2.2

Where subscript (t) denotes year and

S_t	=	Sales
Per_t	=	Personnel expenses
$Nonper_t$	=	Non-personnel expenses
$Depr_t$	=	Depreciation
Int_t	=	Interest payments on debt, less any interest income
Tax_t	=	Tax rate
$\Delta Capex_t$	=	Net investment in non-current assets, (i.e., net of asset sales)
ΔWC_t	=	Net investment in working capital
ΔB_t	=	Net debt issuance (i.e., new borrowing net of repayment)

Sales (S_t), is further calculated according to the following formula.

$$S_t = \left[\frac{Vol_t}{\%Totvol_t} \right] * Price_t$$

$$Sales_t = Total\ volume_t * Price_t$$

Equation A2.3

Where subscript (t) denotes year and

Vol_t	=	Volume (product A) from the Tailored Huff Model
$\%Totvol_t$	=	Volume (product A) percentage of total volume
$Price_t$	=	Average price per total volume

The below table summarizes the input variables, input values and data sources used in the CFF-CB Model to compute cash flows and ultimately comprehensive NPV.

Table A1. Key input variables, values and data sources in the CFF-CB Model

Input variable	Acronym in equation A1, A2 or A3	Input value used	Data source
Volume (product A) for existing stores	Vol_t	Combination of Tailored Huff Model estimates and P&L data (explained in section 4.1)	Store P&L and Tailored Huff Model estimates ¹⁷
Volume (product A) for proposed new store	Vol_t	Estimated via the Tailored Huff Model	Tailored Huff Model estimates ¹⁷
Volume (product A), % of total volume	$Totvol_t$	73%	Company data
Price per total volume	$Price_t$	290	Company data
Personnel expense, % of sales	Per_t	50%	Store P&L ¹⁸
Non-personnel expense, % of sales	$Nonper_t$	20%	Store P&L ¹²
Depreciation, % of sales	$Depr_t$	10%	Store P&L ¹²
Interest payment on debt	Int_t	0	Arbitrarily set
Change in net working capital	ΔWC_t	0	Arbitrarily set
CAPEX, % of sales	$\Delta Capex_t$	10%	Arbitrarily set
Initial CAPEX for hypothetical new store	Inv_0	SEK 6 million	Arbitrarily set
Tax	Tax_t	26.3%	Swedish statutory tax rate
Steady-state growth	g	2%	Arbitrarily set
Net debt	ΔB_t	0	Arbitrarily set
Cost of equity	r_e	10%	Arbitrarily set

¹⁷ The data sources for the estimates of the Huff Model are specified in section 4.2.

¹⁸ Input value from store P&Ls represent average data for fully ramped-up stores.

Table A2. Spreadsheet output illustrating the CFF-CB Model's forecasts of comprehensive NPV for the hypothetical new store

<i>SEK thousand unless otherwise stated</i>	CFF-CB Model forecasts			
	2012 Year 0	2013f Year 1	2014f Year 2	2015f Year 3
Volume (product A) from the Tailored Huff Model		5,544	8,217	9,664
Volume growth		<i>n.a.</i>	48%	18%
Volume (product A), % of total volume		73%	73%	73%
Total volume		7,578	11,232	13,211
Average price per total volume, SEK		290	290	290
Sales		2,198	3,257	3,831
Personnel expenses		-1,099	-1,629	-1,916
Non- Personnel expenses		-440	-651	-766
EBITDA		659	977	1,149
Depreciation		-220	-326	-383
EBIT		440	651	766
Tax on EBIT		-116	-171	-202
+ Interest tax shield		0	0	0
+ Depreciation		220	326	383
- ΔNWC		0	0	0
- CAPEX		-220	-326	-383
- Initial store CAPEX	-6,000			
Cash flow to equity	-6,000	324	480	565
NPV free cash flow, forecast period	-6,000	294	397	424
NPV free cash flow, terminal value	5,409			
Basic NPV	525			
<i>List of existing Company stores affected by NPV cannibalization</i>				
Store F	-1,144			
Store J	-131			
Store U	-17			
Store H	-16			
Store B	-13			
Store S	-13			
Store P	-6			
Store C	-5			
NPV cannibalization on existing Company stores	-1,344			
Comprehensive NPV	-819			

Bibliography

- Baker, Kent H., and Philip English. *Capital Budgeting Valuation : Financial Analysis for Today's Investment Projects*. Hoboken, New Jersey: Wiley & Sons, 2011.
- Biondi, Yuri, and Giuseppe Marzo. "Chapter 23 Decision Making Using Behavioral Finance for Capital." In *Capital Budgeting Valuation: Financial Analysis for Today's Investment Decisions*, by Kent H. Baker and Phillip English, 421-444. Hoboken New Jersey: Wiley & Sons, 2011.
- Brounen, Dirk, Abe de Jong, and Kees Koedijk. "Corporate Finance in Europe: Confronting Theory with Practice." *Financial Management* 33, no. 4 (2004): 77-101.
- Brunzell, Tor, Eva Liljeblom, and Mika Vaihekoski. "Determinants of capital budgeting methods and hurdle rates in Nordic firms." *Accounting and Finance*, 2011: 1-26.
- Bucklin, Louis P. "Retail Gravity Models and Consumer Choice: A Theoretical and Empirical Critique." *Economic Geography* 47, no. 4 (1971): 489-497.
- Craig, C. Samuel, Avijit Ghosh, and Sara McLafferty. "Models of the Retail Location Process: A Review." *Journal of Retailing* 60, no. 1 (1984): 5-36.
- Drezner, Tammy, and Zvi Drezner. "Validating the Gravity-Based Competitive Location Model Using Inferred Attractiveness." *Annals of Operations Research* 111 (2002): 227-237.
- Flyvbjerg, Bent, Mette Skamris, and Soren Buhl. "Underestimating Costs in Public Works Projects: Error or Lie?" *Journal of the American Planning Association* 68, no. 3 (2002): 279-295.
- Fotheringham, Stewart. "Spatial Structure and Distance-Decay Parameters." *Annals of the Association of American Geographers* 71, no. 3 (2010): 425-436.
- Gautschi, David A. "Specification of Patronage Models for Retail Center Choice." *Journal of Marketing Research* 18, no. 2 (1981): 162-174.
- Giuseppe, Bruno, and Gennaro Improta. "Using gravity models for the evaluation of new university site locations: A case study." *Computers & Operations Research* 35 (2006): 436-444.
- Graham, John, and Harvey Campbell. "The Theory and Practice of Corporate Finance: Evidence from the Field." *Journal of Financial Economics* 60 (2001): 187-243.
- Huff, David L. *Determination of Intra-Urban Retail Trade Areas*. Los Angeles: University of California, 1962.
- Huff, David L. *Calibrating the Huff Model Using ArcGIS Business Analyst*. White Paper, Redlands: ESRI, 2008.
- Joseph, Lawrence, and Michael Kuby. "Gravity Modeling and its Impact on Location Analysis." In *Foundations of Location Analysis*, by H.A. Eiselt and Vladimir Marianov. New York: Springer Science+ Business Media, 2011.
- Koontz, Christine M. "Retail Location Theory: Can it Help Solve the Public Library Location Dilemma." In *Planning for a New Generation of Public Library Buildings*, by Gerard B. McCabe, 171-180. Greenwood Publishing Group, 2000.
- Mccarty, Teresa M., Donna F. Davis, Susan L. Golicic, and John T. Mentzer. "The Evolution of Sales Forecasting Management: A 20-Year Longitudinal Study of Forecasting Practices." *Journal of Forecasting* 25 (2006): 303-324.

- McLafferty, Sara. "Predicting the Effect of Hospital Closure on Hospital Utilization Patterns." *Social Science and Medicine* 27, no. 3 (1988): 255-262.
- Nakanishi, Masao, and Lee G. Cooper. "Parameter Estimation for a Multiplicative Competitive Interaction Model: Least Squares." *Journal of Marketing Research* 11, no. 3 (1974): 303-311.
- Nakanishi, Masao, and Lee G. Cooper. "Simplified Estimation Procedures for MCI Models." *Marketing Science* 1, no. 3 (1982): 314-322.
- Okabe, Atsuyuki, and Kei-ichi Okunuki. "A Computational Method for Estimating the Demand of Retail Stores on a Street Network and its Implementation in GIS." *Transactions in GIS* 5, no. 3 (2001): 209-220.
- Ozimek, Adam, and Daniel Miles. "Stata utilities for geocoding and generating travel time and travel distance information." *The Stata Journal*, no. 1 (2011): 106-119.
- Pruitt, Stephen W., and Lawrence J. Gitman. "Capital Budgeting Forecast Biases: Evidence from the Fortune 500." *Financial Management* 16, no. 1 (1987): 46-51.
- Ryan, Patricia A. "Capital Budgeting Practices of the Fortune 1000: How Have Things Changed?" *Journal of Business and Management* 8, no. 4 (2002): 355-364.
- Shrieves, Ronald E., and John M. Wachowicz Jr. "Free Cash Flow (FCF), Economic Value Added (EVA™), and Net Present Value (NPV): A Reconciliation of Variations of Discounted-Cash Flow (DCF) Valuation." *The Engineering Economist: A Journal Devoted to the Problems of Capital Investment* 46, no. 1 (2007): 33-52.
- Statman, Meir, and Tyzoon T. Tyebjee. "Optimistic Capital Budgeting Forecasts: An Experiment." *Financial Management* 14, no. 3 (1985): 27-33.
- Turner, Michael J., and Chris Guilding. "Factors affecting biasing of capital budgeting cash flow forecasts: evidence from the hotel industry." *Accounting and Business Research* 42, no. 5 (2012): 519-545.
- Wooldridge, Jeffrey M. *Introductory Econometrics*. Vol. 4. Canada: South-Western, Cengage Learning, 2009.
- Xu, Yingying, and Lin Liu. "GIS Based Analysis of Store Closure: A Case Study of an Office Depot Store in Cincinnati." *Proc. 12th Int. Conf. on Geoinformatics – Geospatial Information Research: Bridging the Pacific and Atlantic*. Gävle: Geoinformatics, 2004. 533-540.

Additional data sources

Data sources used for building and testing the CFF-CB Model

Data:	Available:
<ul style="list-style-type: none"> 9,675 Swedish postal codes with address information and population 	www.scb.se/Grupp/Totalbef_postnr.xls
<ul style="list-style-type: none"> Addresses for all stores in Sweden 	Concealed
<ul style="list-style-type: none"> Coordinates for 9,675 postal codes Coordinates for all stores Driving times for 224,555 store–postal-code pairs 	www.maps.google.com (via Stata commands “geocode” and “traveltime”)
<ul style="list-style-type: none"> The number of x-attributes for old Market Leader stores 	www.google.com/streetview and the Company’s investment memoranda
<ul style="list-style-type: none"> Accounting data (volume data, income statements and CAPEX per store) Opening dates for the Company’s stores 	Monthly P&Ls for 20–50 Company stores, from the Company
<ul style="list-style-type: none"> The Company’s <i>ex ante</i> volume and cash flow forecasts 	Investment memoranda for 20–50 Company stores, from the Company
<ul style="list-style-type: none"> Opening dates for Market Leader stores 	Annual and interim financial reports for the Market Leader available at its website
<ul style="list-style-type: none"> Volume for the Market Leader per county 	Available at the Market Leader’s website
<ul style="list-style-type: none"> Population per Swedish postal code 	www.scb.se/Grupp/Totalbef_postnr.xls
<ul style="list-style-type: none"> Postal code-to-municipality links for all 9,675 postal codes (retrieved by an Excel macro, coded by us). Municipality per postal code was not provided by Statistiska centralbyrån 	http://www.postnummerservice.se/adress_oekning

Summary of interviews with the Company's CFO

Interview date	Interview type	Interview duration
May 09, 2012	Telephone conference	20 minutes
May 12, 2012	Telephone conference	30 minutes
June 03, 2012	Physical interview	90 minutes
September 07, 2012	Physical interview	90 minutes
September 18, 2012	Telephone conference	15 minutes
September 26, 2012	Telephone conference	15 minutes
September 27, 2012	Telephone conference	15 minutes
October 16, 2012	Physical interview	120 minutes
November 6, 2012	Telephone conference	30 minutes

In addition, we had ongoing e-mail correspondence with the CFO. The issues covered in these e-mails were data requests, discussion around preliminary results and general guidance given to us by the CFO. This correspondence included more than 100 e-mails sent (counting both e-mails sent to the CFO and received from the CFO).