U.PORTO
FEP FACULDADE DE ECONOMIA
UNIVERSIDADE DO PORTO

A workforce size optimization framework for retailers

Patrícia Gomes da Conceição

Dissertation
Master in Economics

Supervised by
Dalila Martins Fontes
Natércia Fortuna

2019

Acknowledgments

I would like to express my sincere gratitude to all the people who contributed to the realization of this dissertation and who accompanied me throughout my academic journey and personal life.

First and foremost, I would like to thank my supervisors Professor Dalila Martins Fontes and Professor Natércia Fortuna, who were always available to help me and who have enriched this work with their exceptional knowledge and valuable advice. I was fortunate to find such good supervisors, and without whom concluding this work would have been more difficult.

To my Master's tutor, Professor Vítor Carvalho, I thank for his guidance and availability during this academic year.

Also, I want to thank the company *NOS Comunicações, S.A.*, especially the Franchising department, for providing the necessary data to achieve the goals proposed for this dissertation.

I thank my family for all the support, understanding, affection, and motivation they have been giving me throughout my life. Most of all, to my parents, I am eternally grateful for all the sacrifices they have made to invest in my education and for giving me the opportunities that lead to my success. To my sister, I thank her for the patience, useful advice and honest opinions provided in the course of this work.

I thank my closest friends, for always being on my side and for inspiring me to want to do more and better.

Last, but not least, the support of ERDF – European Regional Development Fund through the Operational Programme for Competitiveness and Internationalisation - COMPETE 2020 Programme and the Portuguese funding agency, FCT – Fundação para a Ciência e a Tecnologia within project POCI-01-0145-FEDER-031821 is acknowledge.

Resumo

A decisão relativa à dimensão da equipa de colaboradores tem vindo a ganhar importância na literatura relativa à gestão do capital humano no sector do retalho, dado o grande impacto que exerce sobre a capacidade e qualidade de atendimento ao cliente, bem como sobre a performance financeira das lojas de retalho. Este trabalho tem como principal objetivo melhorar o processo de planeamento e dimensionamento da equipa de colaboradores, do ponto de vista dos retalhistas, e, consequentemente, diminuir os efeitos nefastos associados a um número insuficiente e/ou excessivo de colaboradores. Para este efeito, sugere-se uma metodologia composta por duas fases que permite obter um número ótimo mensal de colaboradores para uma amostra de lojas franchisadas. Em primeiro lugar, estimou-se um modelo de comportamento das vendas mensais para cada uma das lojas individuais, utilizando dados históricos das vendas, do nível de colaboradores e do tráfego. De seguida, utiliza-se um método iterativo que permite obter o maior número de colaboradores a alocar a cada loja que garante valor acrescentado para a empresa, em termos de acréscimo de vendas. O método proposto é aplicado e validado na empresa de telecomunicações NOS Comunicações, S.A., como caso de estudo. De acordo com os nossos resultados, o número ótimo mensal de colaboradores necessário durante o período em análise (março 2018 – janeiro 2019), é mais elevado do que o número adoptado pela empresa, o que indica que o retalhista tinha um número insuficiente de funcionários. Consequentemente, houve, no referido período, um volume de vendas inferior ao potencial. Mais especificamente, a alocação do número ótimo de colaboradores permitiria um acréscimo de 5% do volume de vendas, face ao realizado.

Códigos JEL: J21; J22; J23; C61

Palavras-chave: desempenho de vendas; atendimento ao cliente; gestão da equipa de colaboradores; decisão laboral; dados em painel

Abstract

The staffing level decisions have been gaining importance in retail workforce management literature, given its significant impact on customer service and overall store financial performance. The main objective of this research is to improve the workforce planning process for retailers and consequently diminish the costs associated with both over or understaffed stores and lost sales. For this purpose, a two-stage process that returns an optimal monthly staffing level for a sample of franchised stores is developed. First, a sales response model is estimated for each store, using monthly historical data on sales, actual staffing level, and traffic. Then, an iterative method is used to find the largest staffing level that ensures added value brought to the company in terms of increased sales. The proposed method is applied and validated in the telecommunications company *NOS Comunicações, S.A.*, as a case study. According to our findings, on average the optimal staffing level is higher than the company's actual staffing level, which suggests that the retailer was understaffing their stores during the period under analysis (March 2018 to January 2019. Adopting the proposed optimal staffing levels would result in a 5% sales lift for the retailer.

JEL codes: J21; J22; J23; C61

Keywords: sales performance; customer service; sales personnel management; staffing decisions; panel data

Contents

List of Tables

List of Figures

List of Annexes

1. Introduction

This study introduces and analyses a staffing plan to improve labour planning for retailers. Furthermore, the main objective is to develop a solution approach that allows for determining the monthly workforce size[1] for each store of a retail chain.

The motivation for this topic arises from a challenge proposed by *NOS Comunicações, S.A.*[2], to solve a real personnel assignment problem it faces. Therefore, the validation of the approach proposed will resort to a comparison between the volume of sales achieved by the determined staffing levels and those of the resulting staffing level of the retailer's current practice.

This particular company is continuously suffering from staff breakdowns, associated with sporadic and structural factors, such as vacations, absences, high traffic peaks, sickness, and exiting from the company. In practice, whenever a store experiences a shortage in sales collaborators, workers from nearby stores are relocated. Although such relocations obviate the understaffing issue, they still lead to negative impacts in the store performance and service quality, since the collaborators brought in is unfamiliar with the working methods of that particular store.

The problem discussed here is not specific to this particular company; evidence suggesting that several retailers have not been making the best labour planning decisions. Thus calling for the need to improve the staffing level decisions. In the hope of saving up on personnel costs, retailers are, in the majority of the cases, understaffing their stores (Chuang, Oliva, & Perdikaki, 2016; Mani, Kesavan, & Swaminathan, 2015; Netessine, Fisher, & Krishnan, 2010; Ton, 2009). This tendency may be related to the fact that retail managers do not consider labour as a source of profits, but rather only as a source of costs. As such, their focus's on not increasing operational costs, associated with additional staffing (Ton, 2009). So, in order to improve the labour planning decisions, Chuang *et al.* (2016) advise retailers to make staffing decisions with the aim of maximizing sales and profits, as an alternative to minimizing costs.

[1] Hereafter, we will use workforce size and staffing level indistinguishably.

[2] Portuguese telecommunications operator, launched in 2014, that provides a full range of communication and entertainment services (Internet, telephone, and audiovisual products). NOS is listed in the Portuguese Stock Index (PSI-20) and is one of the most valuable and share turnover companies in the general stock market of the Lisbon exchange.

Ensuring the right staffing level available at a particular store is vital to guarantee successful store financial results, while still maintaining customer satisfaction and loyalty. Verde Group & Baker Retailing Initiative (2007) concluded that the lack of employees was the most frequent complaint amongst 1000 American shoppers that participated in the study, with 6% of customers leaving the store dissatisfied because they were unable to find the assistance they needed.

Therefore, this research is essential for retail management in general, as its outcomes will provide retailers with an efficient solution approach to set in-store labour supply proportional to customer demand; this way improving customer experience while optimizing employee productivity and maximizing sales. So, labour decisions instead of relying solo on sales forecasts, as store managers traditionally do (Fisher, Gallino, & Netessine, 2017; Lam, Vandenbosch, & Pearce, 1998), should take advantage of historical traffic data to determine the most appropriate staffing level.

To some extent, the optimal staffing level, achieved through this framework, can indirectly contribute to some or all of the following: shorter waiting time, higher customer and employee satisfaction, increased customer conversion rate[3], better work environment, lower turnover rate, higher volume of sales and profits, and improved store reputation (Quan, 2004).

The rest of this report is organized as follows. Chapter 2 reviews the primary literature concerning workforce planning decisions. In particular, this Chapter focus on the importance of staffing level on the overall retail chain performance (Section 2.1), as well as on the typical process when deploying labour force (Section 2.2). In Chapter 3, the labour planning methodology followed is described, giving some insights into the variables to be used to support this procedure. Chapter 4 presents the case study, giving detailed information regarding the research setting, the data used, uncovering and evaluating the results obtained, and, finally, discussing some of the managerial implications of our findings. Finally, Chapter 5 presents a summary of the main conclusions drawn, as well as some suggestions for future work.

[3] Ratio between the number of customer transactions and the total number of customers entering a store.

2. Literature Review

Cotten (2007) defines workforce planning as a strategic decision which consists of "ensuring that an organization currently has and will continue to have the right people with the right skills in the right job at the right time performing their assignments efficiently and effectively" (p. 6). Inefficient labour planning decisions in a business organization can be translated into either understaffing, whenever the store labour employed is not enough to meet the customer requirements, or overstaffing whenever an excessive staffing level is available (Ogumeyo, 2010; Quan, 2004).

Particularly in the retail environment, proper labour planning decisions must be a priority for store managers (Ton, 2009), given the great proximity between customers and suppliers which turns employees into a precious resource to deliver exceptional customer experiences and motivate purchase decisions (Prickett, 2018). Also, adequate workforce planning can enhance a retailer's competitive advantage over competitors (Quan, 2004).

2.1. Importance of staffing level on retail chain performance

Recently, staffing level optimization has been gaining importance in the literature as several papers analyzing the importance of staffing level on customer service and on overall retail store performance have been published (see, e.g., the reviews by Mani *et al.* (2015) and by Perdikaki, Kesavan, & Swaminathan (2012)).

Using data from forty-one retail stores, Mani *et al.* (2015) show that short-staffed stores damage retailer's revenue the most, being responsible for a negative impact of 5.74% on store profitability, while overstaffing has a smaller negative impact of 2.04%.

Fisher *et al.* (2017) studied the impact on revenue by adding labour in a study involving more than 700 stores of a specialized retailer. A positive impact was observed in 168 stores with the following characteristics: "highest potential demand, as indicated by average basket size, several households and household growth, the greatest competition, and the most experienced store managers." (p. 17). According to Ton (2009), the increased profitability is explained directly by a higher conformance quality – capacity of employees to efficiently execute their work - derived by a higher level of staffing. However, concerning flexible labour resources, Kesavan, Staats, & Gilland (2014b) emphasized that this positive relationship does not hold for all staffing levels; when the staffing level reaches a certain saturation point, it is no longer advantageous to increase any further the level of staffing as it will induce lower sales. As potential causes for this relationship, Heath & Staudenmayer (2000), Raman, DeHoratius, & Ton (2001) and Staats, Milkman, & Fox (2012) claimed the increased distractions, the difficulties in coordinating tasks, and the disturbances in store operations as the staffing level reaches higher numbers.

Some researchers have stressed the importance of having optimal staffing levels by quantifying their benefits in the overall store performance. In particular, Chuang *et al.* (2016) showed that optimal staffing was responsible for a 10% increase on sales performance in the fashion retail stores analyzed in their study. Mani *et al.* (2015) estimated store profitability to increase by 3.8% to 5.9% due to an ideal workforce size.

In terms of customer satisfaction, Fisher, Krishnan, & Netessine (2006), Perdikaki *et al.* (2012), Mani *et al.* (2015) pointed out the critical role that sales collaborators availability play in the process of converting shoppers into potential buyers. Sales collaborators are facilitators of the purchase process since they can provide the help that customers need, give vital information about a product, find the products more quickly, and even suggest other products that the customer may be interested in. Hence, sub-optimal staffing levels are incapable of answering all the previously stated customer desires, affecting both service quality and customer satisfaction (Mani *et al.*, 2015) negatively. As a result, shoppers begin to lose their trust in the enterprise and eventually start to purchase in a competitor capable of delivering a better service (Chuang *et al.*, 2016). In addition, unhappy customers have the power to generate a harmful word-of-mouth[4] effect by sharing their bad shopping experiences, which damages the retailer's reputation (Harris & Arendt, 1998). To understand the magnitude of the word-of-mouth effect on consumer behaviour, Verde Group & Baker Retailing Initiative (2007) surveyed 1000 American customers and concluded that 50% of the analyzed sample gave up visiting a particular store because someone else shared a negative experience.

Moreover, inadequate staffing levels can jeopardize employees' happiness in their work and consequently lead to increased voluntary job turnover[5] (Mosadeghrad, Ferlie, & Rosenberg, 2011). For instance, whenever a store is understaffed, it is more likely that their employees are over-worked and feeling emotionally stressed, which affects their well-being levels; this way, increasing the incidence of mistakes and poor performance in executing their daily

---

[4] Transmission of information from person to person by verbal communication.

[5] Whenever an employee decides that he/she does not want to work in that particular company any more.

tasks (Oliva & Sterman, 2001). On the other hand, highly staffed stores can also affect employee's motivation to work, since they end up working fewer hours than expected and do not feel as productive as they could be (Thompson, 1998). Besides, retail managers are incurring in higher salary costs in order to pay for those extra-workers, thus affecting their profitability profoundly (Mani *et al.*, 2015).

In brief, both over – and understaffed stores are harmful to the retailer's success, causing lower levels of labour productivity and damaging the retailer's financial status (Ganster & Dwyer, 1995). To sum up, researchers debating on this topic show a unanimous conclusion that better labour planning has the power of improving retailers' financial performance, while still providing a high-quality service to their customers.

2.2. The retail labour planning process

Usually, the workforce planning procedure comprises two main steps: forecasting staffing requirements and workforce scheduling, based on the amount of labour required and employee availability (Perdikaki *et al.*, 2012).

Next, a more detailed explanation about these steps is presented; however, particular emphasis is given to the first step since it is the one addressed in this research.

2.2.1. Labour requirements forecasting

The determination of the workforce size requires either forecasting sales or traffic, which are reliable indicators of the labour needed to face customer demand (Netessine *et al.*, 2010).

From the literature, it can be seen that two main frameworks are used to determine the staffing requirements, namely: matching labour to sales and matching labour to traffic.

The traditional labour practice, which links the sales forecasts with a planned staffing level, is supported by Kabak, Ülengin, Aktaş, Önsel, & Topcu (2008). Also, it is the most commonly used planning method amongst retailers (see, e.g., the retailers used in the studies by Fisher *et al.* (2006), Kesavan *et al.* (2014b), and Netessine *et al.* (2010)).

In the first stage of their optimization framework, Kabak *et al.* (2008) determined the hourly staffing needs during a workweek based on forecasted sales revenue. The sales forecast was obtained using data on the number of customers, time period (hours), product promotions, and scheduled labour. The relationship between these variables and the volume of sales is expressed by adopting the store sales function proposed by Lam *et al.* (1998).

In addition to the sales revenue, past studies also incorporate labour costs in their labour planning decision and set the profit-maximizing staffing levels as the ideal number of sales collaborators, see, e.g., Ahipasaoglu, Erkip, & Karasan (2019), Chapados, Joliveau, L'Ecuyer, & Rousseau (2014), and Mani *et al.* (2015).

This view is often criticised because it neglects the critical role of traffic on store sales and, thus, fails to transmit to retail managers the real sales potential of their stores (Lam *et al.*, 1998; Netessine *et al.*, 2010; Perdikaki *et al.*, 2012). If the labour planning decisions depend only on actual sales, then the customers who visited the store but did not make any purchase (perhaps due to lack of assistance) are not taken into account (Chuang *et al.*, 2016). Moreover, the strong causal relationship between store labour and sales is an obstacle to the implementation of this method (Fisher *et al.*, 2017; Perdikaki *et al.*, 2012).

In contrast, some researchers advocate the traffic-based labour approach as the best way to improve labour planning (Chuang *et al.*, 2016; Lam *et al.*, 1998; Perdikaki *et al.*, 2012). According to Perdikaki *et al.* (2012), switching from the first framework to the second one lead to an increase of 1.4% in the volume of sales in a large retail chain; in addition, this change is associated with higher customer basket values[6], thus improving the retailer's revenue (Netessine *et al.*, 2010). Also, Chuang *et al.* (2016) have shown, for a specific apparel retailer, that a 10% increase in sales performance is accomplished, merely, by following the proposed traffic approach. The inclusion of a retail foot traffic metric induces a better utilization of workers and makes it easier to identify peak periods, which require greater staff availability (Perdikaki *et al.*, 2012). In contrast to the traditional approach, there is no need to consider the correlation between store traffic and workforce size, since, in the short term, it is unlikely that employee unavailability will have a significant impact in customer flow (Netessine *et al.*, 2010).

Even though few studies have developed a traffic-based method to determine optimal staffing levels. Chuang *et al.* (2016) proposed a staffing heuristic which uses historical traffic information to determine weekly labour requirements; on the other hand, Lam *et al.* (1998) relied on traffic forecasts to obtain an appropriate workforce size in a single store.

[6] Average total amount spent by each customer over a given period of time.

However, care must be taken when determining labour needs through incoming traffic and a match should be avoided, since not all customer traffic leads to sales. Also, there is heterogeneity between customer traffic behaviour across stores within the same retail chain. Thus, retailers should not use the exact same model to determine labour needs for all stores. Furthermore, Ryski (2005) advises retailers to account for "traffic velocity" (ratio between total traffic and number of operating hours), since the impact of traffic volume depends substantially on store's working hours.

A drawback of such an approach is the high cost of traffic counters (also known as "customer counters", "footfall counters" or "people counters")[7], which are needed to obtain accurate traffic information. In addition to the monetary cost, the lack of precision of these devices is another factor discouraging retailers from installing the counters in their stores.

Finally, high traffic peaks may not always be associated with higher productive periods, as customers may be window-shopping and thus, not augmenting store conversion rate, which may render traffic based labour investments unprofitable. Therefore, the usage of a sales-driven labour forecast might be a better choice in order to maximize labour productivity.

2.2.2. Labour scheduling optimization

For Chapados *et al.* (2014), the computation of the optimal staffing must be followed by a workforce scheduling phase, where retailers allocate those employees to working shifts after accounting for organizational and legal regulations (e.g., working time regulations, rest times, lunch break regulations, and minimum rest period between successive shifts), employees' skills and preferences, and others (Quan, 2004). Also, Chapados *et al.* (2014) suggested the inclusion of operational constraints concerning each store, such as store working hours, a mix of full-time and part-time workers, and a minimum staffing level necessary to operate a particular store.

In this allocation process, some authors use pre-existing shifts (see, e.g., Kabak *et al.*, 2008), while others also construct the shifts, i.e., schedules to be assigned to the employees (Chapados *et al.*, 2014; Talarico & Duque, 2015). The most commonly used approach to design staff schedules is mixed integer-programming models, see, e.g., Azmat, Hürlimann,

[7] Sensors installed in stores to monitor store traffic.

& Widmer (2004), Corominas, Lusa, & Pastor (2002), and Hertz, Lahrichi, & Widmer (2010).

This research adds up to the scarce literature that takes full advantage of customer traffic data to improve staffing level decisions (Chuang *et al.*, 2016; Kabak *et al.*, 2008; Lam *et al.*, 1998), by proposing and developing a methodology to determine adequate staffing levels (see, e.g., Fisher *et al.*, 2017; Kabak *et al.*, 2008; Tan & Netessine, 2014).

Regardless of some different features, the solution approach presented here is based on the retail labour planning framework proposed by Kabak *et al.* (2008). However, we add to it by including labour adequacy, measured by the ratio of staffing level to traffic. The labour-to-traffic ratio has been shown to be a determinant factor in determining staffing levels (Chuang *et al.*, 2016). According to them, "what matters is not labour available per se, but how labour compares to traffic" (Chuang *et al.*, 2016; p. 9); thus, its use is expected to improve in-store customer service and prevent overworked employees as this way the employee customer service rate is used. Secondly, we use historical traffic data, rather than traffic forecasting, which is highly volatile (Chuang *et al.*, 2016), Another important improvement is the incorporation of store location factors, such as competitors and buying power, as crucial determinants of the store performance. Nevertheless, most studies on staffing level optimization disregard these factors. Although we do not address the workforce scheduling problem here, essential labour regulations (e.g., the maximum number of working hours allowed by law, and store operating hours) are included in the process of determining an optimal staffing level.

3. Methodology

3.1. Specification of the labour-planning framework

The retail labour-planning framework chosen to improve a retailer's workforce decision is inspired by that of Kabak *et al.* (2008) and involves a two-stage process.

The literature reports two main labour-planning approaches, namely: traffic-based approach and sales-based approach. The methodological framework chosen for this study is mainly sales driven, because ultimately, what matters the most for retailers in labour planning is guaranteeing a staffing level that will enhance the store financial results. Explicitly, the staffing decision is based on the forecasted sales volume associated with each staffing level. Furthermore, the workforce size is determined through a sales volume optimization setting that ensures retailers a staffing level that enhances the volume of sales, while minimizing staffing requirements. Even though, the usage of store traffic information has not been completely discarded, as the solution approach proposed here uses historical traffic data to support the staffing decision and highlights the labour-to-traffic ratio as a critical driver of store sales performance. Also, as suggested before, special care is taken to ensure that the model retrieves a staffing level capable of serving the expected customer traffic, based on historical maximum active service.

In the first stage, a sales response model is estimated using monthly historical data from each store of a retail chain; the estimation is performed for each individual store and month. Then, the estimated model is evaluated, through a comparison between the forecasted sales using real staffing levels and the actual sales; naturally, using a different data set. The validation set is vital to ensure that the estimated model appropriately weights all the variables and thus provides a good representation of the store sales measurement. In the case of significant differences between the forecasted and the real data, a recalibration of the model must be done to improve the robustness of the model.

In the second stage, an iterative procedure is used to obtain an optimal monthly workforce size for each store. The main idea behind this method is to keep increasing the staffing level while the added value brought to the retail operation is deemed of interest. The termination criterion chosen requires a particular percentage increase in the volume of sales resulting from a higher staffing level.

Finally, the results obtained by the proposed framework are evaluated through a comparison between the forecasted volume of sales with the staffing level in the second stage and the forecasted volume of sales with the current staffing level. This way, it is possible to check whether the workforce size suggested by the model is capable of delivering higher sales performance for the retailer.

3.1.1. Sales Response Model specification

From the sales estimation models used in previous studies (see, e.g., Kabak *et al.*, 2008; Lam *et al.*, 1998; Mani *et al.*, 2015), the one proposed by Chuang *et al.* (2016) formulation was chosen, due to the following reasons: introduction of labour-to-traffic ratio (termed *labour adequacy*) as an explanatory variable, instead of solo relying on labour availability and formulation allowing for the use of panel data[8]. Nevertheless, we extended Chuang *et al.* (2016)' sales function to account for location-specific attributes influencing store sales.

Therefore, in addition to store traffic and *labour adequacy*, the model proposed here considers the level of competition, purchasing power per capita, and the possession of a mobile service in the store location as explanatory variables of store sales. The inclusion of these additional variables was motivated by previous empirical researches that have studied the influence of some of these factors on the sales performance.

Regarding the competitive factor, an early study by Hise, Kelly, Gable, & McDonald (1983) evaluated the impact of both primary competitors (similar retail stores in the same mall) and secondary competitors (similar departments in stores in the same mall) on the performance of stores located in malls; later on, Arnold, Palmatier, Grewal, & Sharma (2009) examined the influence of store manager behaviours on store performance for different levels of competition; Perdikaki *et al.* (2012) concluded that the competition environment, used as a control variable in their analysis, adequately explains the store performance; and, finally, Fisher *et al.* (2017) used it as a store specific attribute and demonstrated that the stores with the most significant competition experience a higher sales increase from additional staffing.

The inclusion of the buying power was inspired by Perdikaki *et al.* (2012), who investigated the impact of the locations' per capita income in the stores performance and were able to

[8] Multidimensional data that involves measurements over time for the same group of individuals, units or entities.

conclude that stores located in neighborhoods with higher per capita income have higher conversion rate.

Lastly, as previous researches suggest (e.g., Cullen, Robinson, Dolphin, Kubitschke, & Clarkin, 1997; McBurney, Parsons, & Green, 2002), defining all potential users of telecommunications services is crucial to forecast the demand for this market; that said, the mobile service rate in the store location was introduced since it reflects the potential customer demand for telecommunications stores, and thus impacts differently their sales performance. From a total of four telecommunications services (Internet, television, mobile, and landline phone), the mobile service was chosen given that it was the service mostly owned by the respondents of a survey carried out by the company used in our work.

Similarly to Chuang *et al.* (2016), control variables are also included in the regression to capture some time-invariant store characteristics that may affect the outcomes of this study. While these authors controlled individually for store location and store size, this research controls simultaneously both store size and type. Note that store type was also used as a control variable in the study by Luceri, Latusi, Vergura, D. T., & Lugli (2014). Authors other than Chuang *et al.* (2016) have used store size in their store performance analysis, see, e.g., Hise *et al.* (1983) and Lusch, Serpkenci, & Orvis (2015).

The monthly fixed effects sales model we propose is given in Equation (1).

$$\boldsymbol{S_{it} = e^{\gamma_t} N_{it}^{\beta_1} e^{\beta_2 \frac{1}{LAit} + \beta_3 PP_i + \beta_4 LC_{it} + \beta_5 MS_{it} + \sum_{j=M,S,G} \propto_j Dj_i A_i + \mu_{it}}} \tag{1}$$

In summary, our sales response model includes the following variables:

- S_{it} – sales volume for store *i* in month *t;*
- N_{it} – store traffic for store *i* in month *t*;
- LA_{it} – labour adequacy for store *i* in month *t*, given by the ratio between the staffing level (*L*) and the store traffic (*N*);
- PP_i – purchasing power per capita in the neighborhood of store *i*;
- LC_{it}– level of competition faced by store *i* in month *t;*
- $\mathrm{MS}_{\mathrm{it}}$ – possession of a mobile service in the neighborhood of store *i* in month *t*;

- DM_i – dummy variable set to one if store *i* is a shopping mall store, and zero for a street store, gallery store and citizen store;
- DS_i – dummy variable set to one if store *i* is a street store, and zero for a shopping mall store, gallery store and citizen store;
- DG_i – dummy variable set to one if store *i* is a gallery store, and zero for a street store, shopping mall store and citizen store.
- A_i – selling area of store *i;*
- $\alpha_{M,S,G}$ – regression coefficients which represent the response of sales to the selling area of a shopping mall store, street store, and gallery store, respectively, relatively to the same impact exerted by the selling area of a citizen store (reference and omitted category).
- $\beta_{1,2,3,4,5}$ – regression coefficients to be estimated;
- μ_{it} – perturbation term;
- γ_t – monthly fixed effects estimate.

The relations between the variables implied by the proposed model are validated theoretically and empirically in the literature. Queueing theory[9] suggests that larger staffing levels (*L*) are associated with higher sales volume (*S*) since larger staffing levels lead to a decrease in the waiting time in lines, which turn decreases the number of customers abandoning the store without making any purchase; Chuang *et al.* (2016) and Hise *et al.* (1983) sustained this positive relationship empirically. Regarding store traffic, Chuang *et al.* (2016) evidenced a positive impact on overall sales performance, *ceteris paribus*.

However, there are diminishing returns to scale associated with these relations, since both staffing level and customer traffic increase sales at a diminishing rate. Indeed, in respect to the former case, several authors claimed that as the store becomes more staffed, each additional collaborator brings a lower positive impact to the sales revenue (see, e.g., theoretical research by Hopp, Iravani, & Yuen (2007) and empirical researches by Fisher *et al.* (2006), Mani *et al.* (2015), and Perdikaki *et al.* (2012)). For the latter case, while Perdikaki *et al.* (2012) provide empirical evidence of this assumption, in theory, this can be explained

[9] Mathematical study of waiting in lines.

by a larger fraction of customers who are just window-shopping and do not intend to make any purchase and also by the adverse effects of crowded stores on customer purchasing behaviour (Eroglu & Machleit, 1990; Harrell, Hutt, & Anderson, 1980; Kesavan *et al.*, 2014a), and having an inadequate staffing level to handle a higher customer demand (Grewal, Baker, Levy, & Voss, 2003). Concerning the crowding effect, it is expected that a higher customer density increases the waiting time in line and deteriorates the customer service, which ultimately can push customers away from making any purchase (Kabak *et al.*, 2008; Perdikaki *et al.*, 2012). Therefore, we anticipate the parameter associated with the store traffic to be positive, but less than 1, that is, $0<\widehat{\beta_1}<1$. As for the sales response to the *labour adequacy*, we can only anticipate a negative impact, i.e., $\widehat{\beta_2}<0$, given the results obtained by Chuang *et al.* (2016).

The purchasing power per capita in the store location will have a positive influence on the sales volume. This prediction is confirmed empirically in the research of Perdikaki *et al.* (2012), whereas an increase in the location's per capita income gave rise to higher sales, *ceteris paribus*.

According to the findings of Hise *et al.* (1983) and Perdikaki *et al.* (2012), we expect the level of competition to have a negative impact on the sales volume, i.e., $\widehat{\beta_4}<0$.

Regarding the control variables, following Haans & Gijsbrechts (2011) and Arnold *et al.* (2009), it is expected that larger stores, of any store type, experience higher sales volumes, i.e., $\widehat{\alpha_M}, \widehat{\alpha_S}, \widehat{\alpha_G} >0$. In fact, with a larger selling area, sales personnel are capable of serving more customers, the adverse effects of crowding are mitigated, the waiting time in line decreases, and the range of products available to the customers expands; these indirectly lead to a better customer experience, and consequently to a sales improvement.

Finally, as referred to by Cullen *et al.* (1997) and McBurney *et al.* (2002), a higher percentage of people owning a mobile service in a given region signals an increase in the demand for closely-located stores offering such services; thereby translating into more sales opportunities. In this context, a positive coefficient is expected for the referred to variable.

Since the sales model given by Equation (1) is non-linear, before applying the different estimation methods, we linearized it by taking the natural log. The log-linearized sales model is given in Equation (2).

$$\ln S_{it} = \gamma_t + \beta_1 \ln N_{it} + \beta_2 \frac{1}{LA_{it}} + \beta_3 PP_i + \beta_4 LC_{it} + \beta_5 MS_{it} + \sum_{j=M,S,G} \propto_j Dj_i A_i + \mu_{it} \qquad (2)$$

3.1.2. Panel data estimation

This section describes and compares the estimation techniques used to obtain the parameter estimates of the log-linearized sales-model. Also, the econometric tests conducted to decide the correct specification for our model are presented.

Initially, the model was estimated without controlling store characteristics and period heterogeneity, i.e., pooled Ordinary Least Squares (OLS) estimation. Then, we applied panel data estimation methods which recognize the panel structure of our data, such as fixed effects (FE) and random effects (RE). Finally, in order to deal with potential endogeneity bias, we used the Generalized Method of Moments (GMM).

The pooled OLS model combines all cross-sectional and time-series data into one dataset, neglecting their differences. If the model does not contain an unobserved effect, then this estimation method is efficient, and the conclusions retrieved by the statistical inference are valid (Park, 2011; Wooldridge, 2010).

However, in the presence of an individual effect (cross-sectional or time-period effect), it's not recommended to use a pooled OLS regression since its estimation might lead to omitted-variable bias[10] and inconsistent estimators (Wooldridge, 2010). Therefore, in these cases, the alternative is to use panel data techniques, such as fixed effects and random effects, which include variables to capture the unobserved heterogeneity among different cross-sections. Additionally, the usage of panel data methods is associated with several advantages over the traditional OLS specification, such as higher level of information, more data variation, less collinearity, and more degrees of freedom from combining the two-dimensional information; more adequate models to study the dynamics of change and sophisticated behavioral methods; minimization of the bias of aggregating cases into broader categories; better at identifying and computing the unobserved effects in either cross-section or time-dimension data (Baltagi, 2008; Gujarati, 2009).

In terms of specification, the fixed effects model includes the effects of the individual characteristics as part of the intercept, allowing for each cross-section to have its intercept; whereas the random effects model captures this individual heterogeneity in the error term

[10] Problem of over-estimating or under-estimating the causal effect of the included independent variables, which results from the omission of a significant independent variable that is correlated with both dependent variable and one or more other independent variables

(García, 2014). Basically, the main difference between these two models is explained by the properties of the individual effect. The fixed effects approach allows the time-invariant individual effect to be arbitrary correlated with the observed independent variables; while a random effects model considers these differences among individuals to be stochastic and neither correlated with any regressor nor with the idiosyncratic error (Croissant & Millo, 2008).

The Generalized Method of Moments (GMM) technique was performed to account for the potential endogeneity between the staffing level and sales volume (Chuang *et al.* (2016), Perdikaki *et al.* (2012), and Mani, Kesavan, & Swaminathan (2011) also use this technique for the same reasons). The endogeneity issue, which translates into biased coefficient estimates, arises due to the following reasons. Firstly, retailers usually decide the staffing level based on the expected demand (sales or traffic); thus, not controlling for expected demand, which may result in biased estimates. Secondly, not incorporating into the estimation procedure the unobserved store characteristics that drive both staffing level and sales volume may lead to biased estimates of the coefficients due to stores unique behaviours. Lastly, the possible dual causality between sales and staffing level may induce the endogeneity bias; in fact, not only the deployed staffing level is able to influence sales, but also the current sales performance can make retailers alter their staffing levels (Lam *et al.*, 1998; Kesavan *et al.*, 2014b).

However, some of the following particularities used in our framework may have mitigated the endogeneity issue: the usage of actual store traffic information, which controls for the expected demand, as suggested by Chuang *et al.* (2016), Mani *et al.* (2011), and Perdikaki *et al.* (2012); and, the inclusion of store size and store type to control for time-invariant store characteristics. Moreover, Chuang *et al.* (2016) realized through interviews that retailers did not change that often the staffing level decision; this finding might obviate the reverse causal relationship mentioned previously. Nevertheless, we decided to resort to the GMM method to test the robustness of our estimates.

Whenever there is a suspicion that one or more explanatory variables are correlated with the error term of the regression equation, that is, to be endogenous, it is preferable to use the GMM estimation method, which consists of "replacing" the explanatory variable whose exogeneity is dubious by one or more instrumental variables. The instrumental variables need to fulfil two essential conditions: play the same role in the model as the explanatory variable(s) they are replacing, i.e., should be correlated with the endogenous regressor; and being uncorrelated with the errors (Wooldrige, 2010).

Following Bloom & Van Reenen (2007), Chuang *et al.* (2016), Kesavan *et al.* (2014b), Siebert & Zubanov (2010), and Tan & Netessine (2014), we used the lagged values of the staffing level as instrumental variables.

After obtaining the parameter estimates of each model, three statistical tests were conducted to verify the appropriate model for our data. First, the Redundant Fixed Effects Test (or F-test) and the Breusch and Pagan's (1980) Lagrange Multiplier (LM) test were performed to compare the pooled OLS against the fixed effects and the random effects, respectively. The F-test evaluates the null hypothesis that all time-specific intercepts are equal to zero (pooled OLS), alongside the alternative hypothesis that at least one of those components is not zero (fixed effects). The LM test assesses if the random variance of the individual heterogeneity is equal to zero, which is equivalent to the absence of an unobserved effect. In both tests, the non-rejection of the null-hypotheses indicates that the pooled OLS model is a better specification than the panel data models. If the null hypothesis of the F-test is rejected, then it means that there is a significant fixed effect in our model and thus a fixed effects model is more suitable than a pooled OLS one. On the other hand, the rejection of the null hypothesis of the LM-test suggests that there are substantial differences across cases to be accounted for, thus, favouring a random effects model (Wooldridge, 2010).

Lastly, we performed the Hausman test (Hausman, 1978) to help us decide between the estimates in the fixed effects and random effects models. This test evaluates the null hypothesis of no statistical difference between the two estimates, against the alternative hypothesis of a considerable distance among the models. The similarity between the two estimates happens when both models are consistent, i.e., when individual effects are not correlated with an independent variable. In these situations, it is better to choose the random effects model because it is more efficient than fixed effects. On the other hand, when the estimates statistically deviate from each other it is because one model is consistent and the other is inconsistent; this occurs when the individual effects are correlated with the other regressors, which turns a random effects model inconsistent. Therefore, in these cases, the consistent estimates delivered by a fixed effects model are preferred.

3.1.3. Determination of the workforce size

This section details the iterative procedure used to find an optimal workforce size every month. The baseline of this procedure is the robust sales estimation model obtained in the first stage of the proposed framework.

The flowchart shown in Figure 3.1 illustrates the steps involved in determining the monthly workforce size for each store.

First of all, our framework needs to ensure some minimum staffing levels to sustain the store operation and the customer experience. Therefore, the starting point of this procedure is the minimum staffing level (*minl*) necessary to (i) have the store open and (ii) satisfy the expected number of customers, based on the historical maximum active service. Details on the computation of this variable are presented later on in this study.

Above the minimum staffing level, the procedure iteratively increases the size of the sales force while the ratio between the estimated sales growth rate and labour growth rate is above a pre-defined threshold (δ); consequently, optimal staffing levels will depend mostly on this threshold. The procedure starts by checking whether the estimated discrete sales growth rate from stage 1 (L=*minl*) to stage 2 (L=*minl*+1) is at least δ times higher than the labour growth rate. If that is the case, then the workforce size is incremented by 1 and it proceeds to the next iteration; otherwise, the procedure ends and the workforce size is set to the optimal staffing level. By incorporating the labour growth rate instead of solo relying on the absolute staffing level increase, we have in consideration the current workforce size.

As Fisher *et al.* (2017) point out, the effect of increasing workforce size in the sales volume is expectantly different across stores; thus, this procedure is performed individually for each store (i) and each month (t). This way, we can obtain workforce size recommendations over time for each store, which is the primary purpose of this study.

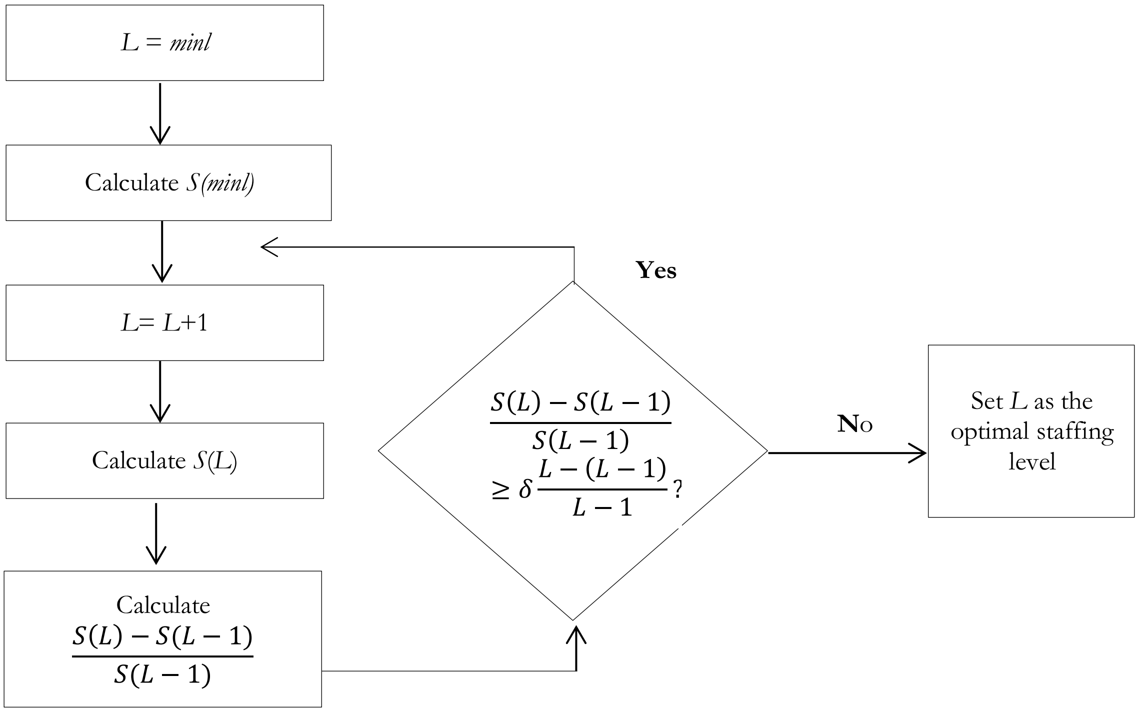


Figure 3.1. *Flowchart of the iterative procedure used to determine the workforce size.*

4. Case Study

4.1. Data

The proposed labour-planning framework was applied in the telecommunications company *NOS Comunicações, S.A.*, as a case study.

4.1.1. Research Setting and Data Collection

NOS Comunicações offers its products and services in physical stores, i.e., retailers, and also through door-to-door sales. The retail service operates through multiple store channels, such as own stores[11], franchised stores[12], and shop-in-shop stores[13]. Our analysis was restricted to the franchised stores since these were the stores for which the company requested a labour planning approach and provided the necessary data to test our methodology.

In January 2019, the company had 115 stores operating under a franchising regime. These franchised stores were distributed among five franchisees: Agent A, with 32 stores; Agent B with 27 stores; Agent C, with 25 stores; Agent D, with 31 stores. In total, there were 321 sales collaborators equipping the stores; Agents B and D were the franchisees with the largest and smallest workforce size, respectively. Regarding the geographical distribution of the labour resources, Figure 4.1 shows that they were widely spread over the national territory. It is possible to see that Porto and Braga were the districts with the largest total workforce size; however, this might be associated with the highest store density observed in these locations, that can be observed by the larger circles.

---

[11] Stores that only sell products and services from the operator, and thus, its sales revenue is wholly directed to NOS.

[12] Like own stores, franchised stores sell products exclusively from the principal, i.e., *NOS*. However, these stores are owned by a franchisee agent.

[13] Stores that sell products of any operator, and thus, the obtained sales revenue is shared amongst the different competitors.

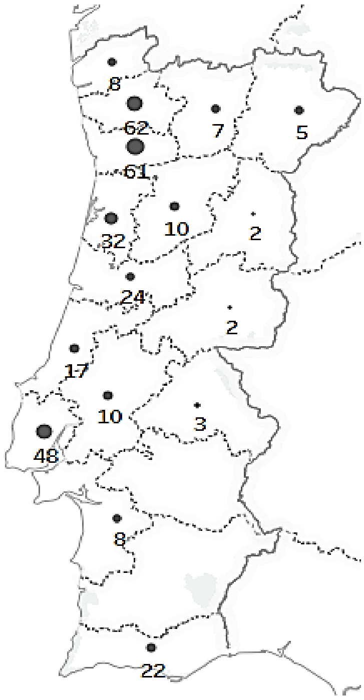


Figure 4.1. *Geographical distribution of the labour resources and stores throughout the Portuguese districts. The reported values correspond to the number of sales collaborators, and the size of the filled circles represents the number of stores in each district. Data are referring to January 2019.*

Besides, the store chain comprises four different store types: street stores, gallery stores, shopping mall stores, and citizen stores. In January 2019, street stores and shopping mall stores were the most predominant, representing 69% and 17% of the whole store sample, respectively. Gallery stores followed with a smaller percentage of 11% and citizen stores accounted for the remaining 3%.

We obtained the following monthly data from the retailer over the period of March 2017 – January 2019: (i) store sales volume; (ii) staffing level; (iii) store traffic data (i.e., number of customers entering the store and number of open transactions); (iv) adequate service data (i.e., number of customers served and number of closed transactions); (v) location-specific data (i.e., purchasing power per capita, number of competitor stores, and possession of a mobile service); (vi) store size; (vii) store type; (viii) weekly store operating hours (including weekends and excluding lunch breaks).

4.1.2. Data Description and Preprocessing

Table 4.1 describes in detail all the variables used in Eq. (1) along with the corresponding data sources.

Table 4.1. *Description of the variables used in the sales model.*

Variable	Description	Units	Source	Available Periodicity
S	Sales volume expressed as points achieved from the number of products and services sold.	Points	Company Reports	Monthly
N	Number of customers entering the store for stores with a ticket counter system; conversion of open transactions into customer volume, otherwise.	Count	Company Reports	Monthly
L	Number of active employees, i.e., number of employees available to work, excluding employee absenteeism.	Count	Company Reports	Monthly
PP	Purchasing power, in per capita terms, of the municipality where the store is located, using the national average (100) as a reference [b].	Index	*Pordata*	Biennial
LC	Number of competitor stores in the municipality where the store is located.	Count	Company Reports	April 2017; December 2017; July 2018; February 2019
MS	Percentage of respondents with mobile service in the region where the store is located.	Percentage	Company Reports	Quarterly
DM	Dummy set to 1 if the store is a shopping mall store; and 0 for a street store, gallery store, and citizen store.	Binary	Company Reports	Not applicable
DS	Dummy set 1 if the store is a street store; and 0 for a shopping mall store, gallery store, and citizen store.	Binary	Company Reports	Not applicable
DG	Dummy set to 1 if the store is a gallery store; and 0 for a shopping mall store, street store, and citizen store.	Binary	Company Reports	Not applicable
A	Storefront office area	Square meter	Company Reports	Not applicable

[a] The percentage of stores with a ticket counter system, and thus for which the real store traffic information was available, is around 19%. For the other stores, the conversion of open transactions into costumer volume was obtained by resorting to the relationship between the number of customers entering the store and open transactions in the group of stores with the system installed.

[b] This indicator compares the local purchasing power with national performance.

[c] Expect for the last trimester of 2018 and the first trimester of 2019.

As can be seen in the last column of Table 4.1, not all data were available monthly. Therefore, we had to assume the following period of reference: data from the most up-to-date year, i.e., 2015, as a *proxy* of the buying power achieved monthly by each municipality; Regarding to the possession of a mobile service (MS), from March 2017 to September 2018 (inclusive), the moment of reference is the correspondent trimester and, from October 2018 to January 2019 (inclusive), it is the 3rd trimester of 2018; the monthly level of competition (*LC*) refers to the nearest month for which this information was available (for instance, from May 2017 to December 2017, the reference period is December 2017).

While collecting the data, special care had to be taken because of store migration cases to a different franchisee and one consolidation case where two stores have been merged. For the migration cases, we assume the complete store history regardless of the franchise agent to which it belonged; while for the latter case, before the merger, we assume the combined data of the two individual stores, and afterwards, we gathered the consolidated store data.

The time series used, from March 2017 to January 2019, was divided into two subsamples: the fit sample (from March 2017 to February 2018) and the test sample (from March 2018 to January 2019). The fit sample was used to estimate the sales model, in the first stage of our framework, and to compute the maximum active service, required in the second phase. The test sample was used to evaluate the sales estimation model, as an input of the iterative procedure, and, also, to validate the results obtained by our solution approach.

Descriptive statistics of the main variables regarding the complete sample and each subsample are presented in tables 4.7, 4.8, and 4.9, in Annex A. Additionally, pairwise correlations relative to the complete sample and each subsample can be found in tables 4.10, 4.11, and 4.12, in Annex B.

4.2. Results

In this section, the results obtained by the proposed labour planning framework are presented and evaluated. Section 4.2.1 reports the first-stage results, i.e., the sales model regression results and the validation of the calibration obtained. Particular emphasis is given to the second-stage results, in Section 4.2.2, whereas the optimal staffing levels are shown and inferred for a higher sales performance.

4.2.1. First-stage results

Using the unbalanced fit sample (March 2017 – February 2018), we estimate the Pooled OLS; the monthly fixed effects; the monthly fixed effects using the first and eighth lags of labour as instrumental variables (IV); and the monthly random effects regression models.

4.2.1.1. Estimation results

Table 4.2 reports the parameter estimates of the different specifications. Notation (a) shows the estimates of a reduced model accounting only for traffic and labour adequacy as regressors. Notation (b) represents the complete model given in Eq. (2). Notation (c) represents an alternative monthly fixed effects model where we have applied the conversion of open transactions into customer volume for the entire store sample (*NM*), instead of just using for those stores without the ticket-counter. The figures provided in between brackets are the standard errors for which a consistent estimator for heteroscedasticity and autocorrelation was used.

Table 4.2. *Regression results of the sales estimation model.*

Variable	Pooled OLS		Monthly Fixed Effects			Monthly Fixed Effects IV		Monthly Random Effects	
	(a)	(b)	(a)	(b)	(c)	(a)	(b)	(a)	(b)
ln(N)	1.0012*** (0.017)	0.9741*** (0.022)	0.9917*** (0.017)	0.9611*** (0.022)		0.9721*** (0.017)	0.9469*** (0.030)	0.9985*** (0.017)	0.9741*** (0.022)
1/LA	-0.0005*** (0.000)	-0.0005*** (0.000)	-0.0006*** (0.000)	-0.0005*** (0.000)	-0.0005*** (0.000)	-0.0007*** (0.000)	-0.0006*** (0.000)	-0.0006*** (0.000)	-0.0005*** (0.000)
DMA		0.0060*** (0.001)		0.0060*** (0.001)	0.0033*** (0.001)		0.0051*** (0.002)		0.0060*** (0.001)
DSA		0.0075*** (0.001)		0.0074*** (0.001)	0.0044*** (0.001)		0.0076*** (0.001)		0.0075*** (0.001)
DGA		0.0035** (0.002)		0.0036*** (0.001)	0.0009 (0.001)		0.0024 (0.002)		0.0035** (0.002)
PP		0.0020** (0.001)		0.0020*** (0.001)	0.0022*** (0.001)		0.0020** (0.001)		0.0020** (0.001)
LC		-0.0082 (0.005)		-0.0075** (0.004)	-0.0057* (0.003)		-0.0028 (0.004)		-0.0082 (0.005)
MS		0.0144*** (0.004)		0.0133*** (0.005)	0.0081** (0.004)		0.0087** (0.004)		0.0144*** (0.004)
ln(NM)					0.9814*** (0.019)				
F-statistic	1792.731***	470.1524***	285.6285***	204.6582***	225.7884***			1786.632***	470.1524***
Adjusted R^2	0.7332	0.7493	0.7394	0.7550	0.7726	0.7271	0.7438	0.7325	0.7493
Observations	1305	1257	1305	1257	1258	428	412	1305	1257

Notes: *, **, *** Significant at 10%, 5% and 1%, respectively. In order to prevent losing information and biased coefficients, stores with some data missing (only in certain months) were kept in the estimation procedure.

Next, we detailed how the right specification was chosen. In the first place, we decide on the model to use, namely: the fixed effects model or the Pooled OLS regression, through the usage of a Redundant Fixed Effects Test (or F-Test). Given that the p-value obtained was 0.0000, which is lower than 5%, we rejected the null hypothesis and concluded that the fixed effects model is preferred over the Pooled OLS model. In other words, this means that we must incorporate the heterogeneity of each time unit in the estimation procedure. Also, for the purpose of our study, it is advantageous to use a panel data model rather than a simple Pooled OLS estimation, since the latter does not capture the evolution of the store characteristics over time and thus the estimates obtained might not represent accurately the relationships involved in this framework.

Afterwards, we performed a Hausman test to test the fixed effects (FE) model against the random effects (RE) model. With a p-value of 0.0002, we rejected the null hypothesis that the RE model is consistent, and thereby conclude that the FE model was a better fit.

Finally, a Hausman endogeneity test was performed to check the staffing level endogeneity, for which two instrumental variables are available: the first and eighth lags of labour. If there is no endogeneity issue, the IV estimates are similar to the FE estimates, and both estimates are consistent. By comparing both estimates, we can infer that instrumentation had a small impact on the FE estimates; thereby, confirming the robustness of our estimates. Even so, the test was performed to draw a more reliable conclusion. The null hypothesis that the staffing level is uncorrelated with the sales volume, i.e., exogenous, was not rejected; this way, eliminating the endogeneity concerns.

In sum, from these tests, we conclude that the FE model, as given in Eq. (2), was the most appropriate for our data and thus, the one to be used in the rest of our analysis, due to the existence of nonrandom individual differences across time over stores. Moreover, the FE model presents a higher adjusted coefficient of determination (adjusted R^2) than the Pooled OLS and RE models, which indicates an improvement in the goodness of fit of the model by using this specification.

By comparing the regression results of the FE model in column (a) with those in column (b), it can be seen that the inclusion of more regressors to the original model proposed by Chuang *et al.* (2016) improved significantly the explanatory power, which is measured by the adjusted R^2. This confirms that, for our data, the model proposed in this study is more reliable to explain and predict future sales outcomes. Also, the resulting parameter estimates from the estimation of the alternative specification, given in column (c), were further apart from our guideline estimation study by Chuang *et al.* (2016). Although this option would be more consistent, the usage of real traffic data leads to more accurate results; this way, justifying the preference for the estimates considering both direct and indirect measures of customer traffic, which are given in column (b).

As it can be seen, from the FE regression results in Table 4.2, the regression and all independent variables are statistically significant to explain the sales behaviour across stores and across time, i.e., they all exert a significant impact, different from zero, over the sales volume. Besides, these variables explain 75.50% (adjusted R^2) of the dependent variable, which is considered to be a significant explanatory power.

Moreover, the estimated parameters have the expected sign, which supports the accuracy of our results. In conformity with Perdikaki *et al.* (2012), competition influences the sales volume negatively, whereas an increase in one competitor store decreases sales volume by approximately 0.75%, *ceteris paribus*. As expected, the empirical evidence shows that store traffic, purchasing power per capita, and possession of a mobile service all have a positive impact on the sales volume. Indeed, the positive influence exerted by the store's location buying power is consistent with the findings of Perdikaki *et al.* (2012); although the authors have used the location's per capita income, it can be used as an indicator of the buying power. Although several studies have confirmed that higher store traffic generates more sales (see, e.g., Chuang *et al.*, 2016; Kabak *et al.*, 2008; Perdikaki *et al.*, 2012), our impact is much powerful, reaching almost unitary elasticity; meaning that, the majority of the customers entering in these stores effectively make a purchase.

Regards to the *labour adequacy*, our result confirms the findings of Chuang *et al.* (2016); however, in our estimation setting it has a much smaller impact than that of the results reported by Chuang *et al.* (2016).

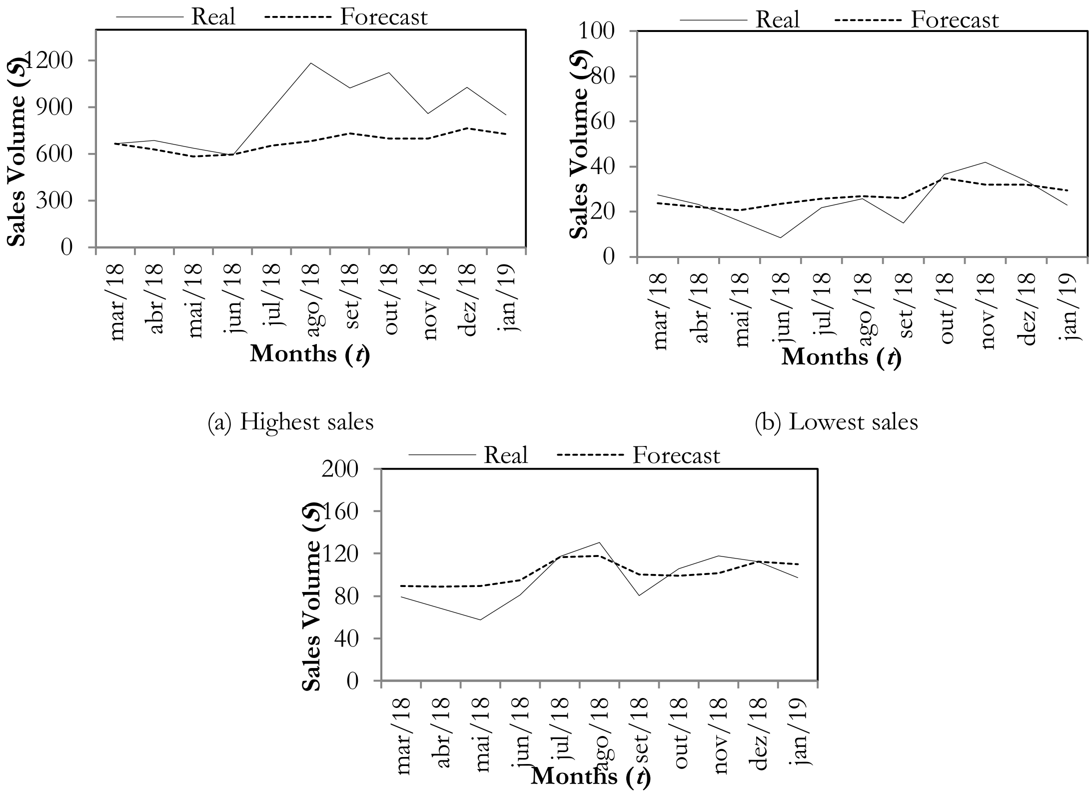
In line with Lusch *et al.* (2015), all the control variables influence the dependent variable positively. Nevertheless, the street store's size accounts for the stronger positive impact upon the volume of sales compared to the effect of the citizen store's size, our reference category.

4.2.1.2. Model Validation

In order to check the robustness of the estimated sales model, we used it to forecast the volume of sales in our test sample period (March 2018 – January 2019). By using the test sample, we were able to evaluate how well this model is likely to forecast on new data that was not used to calibrate the model.

The validation was conducted in two steps: first, we obtained the forecasted sales volume by substituting into Eq. (2) the parameter estimates of the FE complete model, reported in Table 4.2, and using the test period data; then, we compared the forecasts against actual sales, during the test period. Note that, due to some missing data, the model could not forecast sales for the entire store sample. More specifically, the forecasts refer to 111 stores out of the 115 existing ones.

A graphical representation of both the forecasted and the actual series is provided in Figure 4.2, concerning the store with the (a) highest, (b) lowest, and (c) intermediate total sales volume during the test period.



(a) Highest sales (b) Lowest sales

(c) Intermediate sales

Figure 4.2. *Comparison of the real and forecasted sales for the store with the (a) highest, (b) lowest, and (c) intermediate total sales volume.*

Regarding the store with the highest sales volume, Figure 4.2 (a), we infer that the sales forecasts obtained from our model were quite assertive from March 2018 until June 2018, as the lines correspondent to the real and forecasted values were almost perfectly aligned during this period. However, from July 2018 onwards, the forecasted sales started to diverge slightly from real sales. In particular, the most significant deviation occurred in August 2018, which can be explained by the sales effect of the operator network expansion in the store's location during that particular month; an event that was not controlled by our model because this information was given a posteriori. Still, we should not entirely rely on these results to draw our conclusions, since the store illustrated here represents a minority of the store sample, since few stores achieve such high sales volume. Therefore, in our opinion, it is more appropriate to evaluate the proposed model based on the forecasting performance in other stores, for which the predicted sales were over time very close to real sales, as it can be seen in Figure 4.2 (b) and (c).

To draw more accurate conclusions, we further evaluate the forecast ability of our model based on the computed values of commonly used accuracy measures, as such: Theil's Inequality Coefficient (U_2); Root Mean Squared Error (RMSE); Mean Absolute Error (MAE); and Mean Absolute Percentage Error (MAPE). All these metrics provide a measure of the forecast error, that is, the difference between the model forecasts and the actual values. Thus, they indicate how close the model's predictions are from reality. Further description of the forecasting techniques is given in Table 4.3. It should be noticed that the MAPE gives an accurate measure of the percentage error when dealing with strictly positive target forecast series (Chapados, Joliveau, & Rousseau, 2011); otherwise, alternative measures should be used.

Table 4.3. *Description of the forecast accuracy measures used in the model validation.*

Accuracy Measure	Description	Scale	Desirable Value
Theil Inequality Coefficient (U)	Root mean square error in relative terms	Scale-invariant	Closer to zero
Root Mean Squared Error (RMSE)	Standard deviation of the forecasted errors	Scale-dependent	Closer to zero
Mean Absolute Error (MAE)	Mean of the absolute forecasted errors	Scale-dependent	Closer to zero
Mean Absolute Percentage Error (MAPE)	The absolute forecast error (i.e., MAE) as a percentage of the actual values	Scale-invariant	Closer to zero

The Theil Inequality Coefficient varies between 0 and 1. A zero value is the best one and is associated with a perfect forecast, i.e., the predicted values are equal to the actual values.

On the other hand, if it approaches 1, then the forecasts are quite inaccurate (Bliemel, 1973). This statistical measure can be decomposed into three proportions, namely: bias proportion, variance proportion, and covariance proportion. The bias proportion indicates the systematic error of the forecast (over or under prediction); in other words, it measures the difference between the mean of the forecasted series and original series. The variance proportion measures how well the predicted values replicate the variability of the original series; that is, it shows how far the variation of the forecast is from the variation of the original series. The covariance proportion indicates the unsystematic forecasting errors. Ideally, the covariance proportion should be the highest proportion among all, and, in addition, the bias and variance proportions should assume small values.

Since both RMSE and MAE are scale-dependent measures, i.e., their magnitude depends on the magnitude of the dependent variable, there is no absolute threshold value to evaluate the forecast accuracy; thus, these evaluation measures are only useful to compare forecasts on the same scale across different models. Still, it is commonly accepted that the lower these values are, the better is the predictive performance. In the presence of data sets with different scales, a scale-independent measure, such as MAPE, must be used. Although a value of 0 indicates a perfect fit, there is also no upper limit for the percentage error.

The forecast accuracy measures are reported in Table 4.4.

Table 4.4. *Results of the forecast accuracy metrics.*

Accuracy Measure	Result
U - Theil Inequality Coefficient	0.0405
Bias Proportion	0.0023
Variance Proportion	0.0650
Covariance Proportion	0.9328
MAE - Mean Absolute Error	0.2880
RMSE - Root Mean Squared Error	0.3761
MAPE - Mean Absolute Percentage Error	6.7102

The low assessed values for these measures suggest that our model was well calibrated to measure the store sales volume. In fact, the Theil Inequality Coefficient is only 0.0405, suggesting that the predictive performance of our model is almost perfect; as desired, the covariance proportion assumes the highest value of the three proportions; the bias and variance proportion are small, suggesting that the forecasted sales volume replicate well the mean and variance of the actual sales volume. In addition, the MAE is 0.2880, and the RMSE is 0.3761, which support the conclusion of a good forecast ability achieved by our model. We have obtained a relatively high value for MAPE (6.7102%). Nevertheless, this can be explained by the presence of outliers, that interfere with the calculation of this measure; mainly, because of the target forecast series, i.e., sales volume, which assumes values close to zero (as suggested by Chen, Twycross, & Garibaldi (2017) and Kim & Kim (2016)).

In sum, our model has been shown to be not only suitable, but also robust to forecast store sales volumes.

4.2.2. Second-stage results

After concluding, in the first phase of the methodology, that our econometric model was well specified and calibrated for the target series, we were able to proceed to the second phase, i.e., using the proposed iterative procedure to determine the workforce size.

At this point, an important determinant was the minimum monthly staffing level (*minl)* necessary for each store. We divided the weekly store operating hours by the allowable number of weekly working hours of an employee (i.e., 40h) to obtain the store personnel necessary to fulfil condition (i), that is, to have the store open. Next, condition (ii), that is, to serve the expected customer traffic, is satisfied with a staffing level correspondent to the ratio of the test period store traffic (*N*), and the maximum active service achieved in the previous period, i.e., fit sample period. The maximum active service denotes the greatest number of customers served by each sales collaborator. Since we could not observe the number of customers served for the entire store sample, we followed the procedure used for the store traffic variable (*N*). Thus, we assume the active service to be the: real number of customers served for the stores with the ticket counter installed and the conversion of closed transactions into customers served for the others.

Note that the maximum active service should not be interpreted as the sales personnel' service capacity since it does not include the fraction of customers that each collaborator would be capable of serving if the store traffic was higher. In other words, the service capacity is at least the maximum active service, and thus, the usage of the latter might be overestimating the minimum staffing levels.

Naturally, whenever condition (ii) implies a higher staffing level than condition (i), what prevails as minimum staffing level is the first requirement, since our labour-planning framework needs to ensure a staffing level capable of having the store in operation.

4.2.2.1. Determination of the sales increment parameter

This section starts by explaining how the sales increment parameter value - a required input to solve the iterative procedure – is obtained. Then, in the following sections, we present and validate the results obtained for the retail franchisee. Given the substantial amount of derived outcomes, only the main illustrative results are presented.

Since the studied company did not provide a specific value for the required ratio between the percentage sales gain and labour increase (δ), we decided to set it to 22%, based on the stores' behaviour in prior periods and through a series of comparisons between simulations using two different thresholds. The company requires an average labour productivity level (volume of sales per collaborator) of 40 points per month, and thus our decision was grounded as well in the corresponding levels achieved by each threshold. Also, a simulation using a decreasing threshold in the course of the iterative procedure was performed to analyze and compare with the results obtained using a constant threshold. The decreasing threshold is in line with the diminishing returns to scale associated with the staffing level and store sales (see further explanation in Section 3.1.1). However, we found that a variable threshold would make little difference to the labour productivity levels achieved by the optimal staffing levels, and thus decided on a constant threshold, as proposed by the company.

We started by computing the real value of this parameter, i.e., the ratio between the sales growth rate and labour growth rate, for some stores using the estimated model. As the values found were within the range 20-30%, we run the iterative procedure considering values between 15% and 30%. Then, according to the results obtained, we narrow those values down to 20-30% (excluding 23%[14]). The range of values considered ensures the company required productivity levels; while also ensuring a normally distributed series at the significance level of 5%.

Given the considerable amount of suitable parameter values, under the company's terms, the final decision was assigned to the company. Therefore, after observing the possible outcomes obtained by each threshold, the percentage of the increment that the company considered more appropriate was 22%, whose ideal staffing levels would contribute to an average labour productivity of 43 points. Since the *Jarque-Bera* statistics has a p-value equal

[14] The parameter value 23% was excluded due to the non-normality of the resulting labour productivity series at the significance level of 5% and 10%.

to 0.1218, this threshold guarantees the normality of the labour productivity series at 1%, 5% and 10% significance levels.

Furthermore, we check whether the company's choice had a significant impact on the final results through a sensitivity analysis using the set of acceptable parameter values; see the results obtained for a representative store in Figure 4.3.

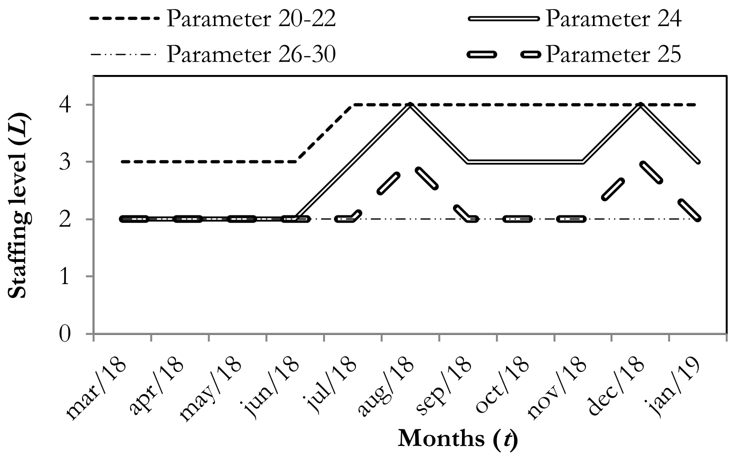


Figure 4.3. *Results obtained under different parameter values for a representative store of the actual staffing level during the test period. Same results for parameter values between 20% and 22%, and from 26% to 30%.*

As it can be seen, the optimal staffing levels would be similar regardless of the company's decision, which demonstrates the robustness of our results. Note that, the magnitude of the difference between the obtained workforce sizes for different parameter values varies between 0 and 2.

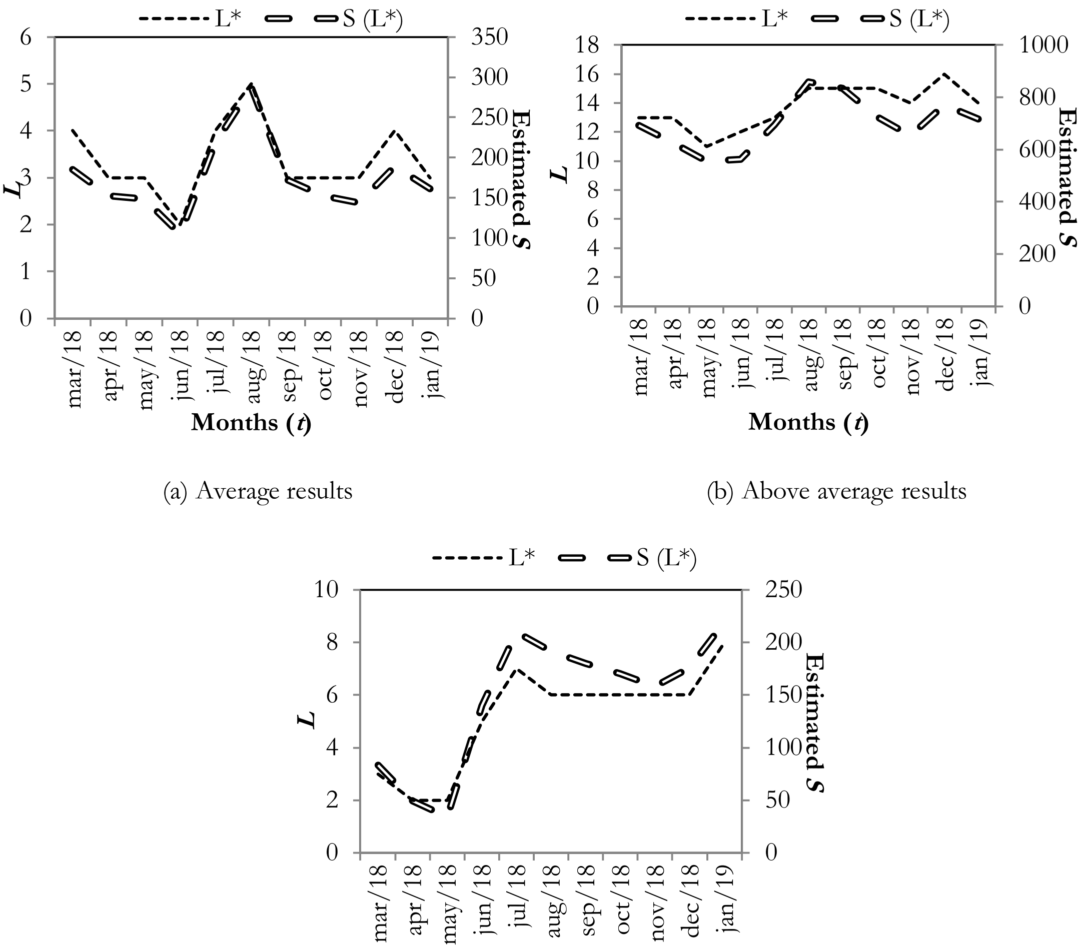
4.2.3. Optimal staffing levels

In this section, we present a summary of the outcomes obtained from the iterative procedure, i.e., an optimal monthly staffing level for each store of the retail franchisee. It is noteworthy that, due to the absence of data, nine of the 115 franchised stores were excluded from this procedure.

The implementation of the iterative procedure resulted, for some cases, in a staffing level larger than the minimum required; however, the number of collaborators was never more than 3 above that number. In fact, none of the stores experienced a sales growth rate of at least 22% when considering the additional fourth collaborator, and thus the software ended this procedure in the fourth iteration.

The average total optimal monthly staffing level, for the whole retail chain during March 2018 – January 2019, amounts to 328 active collaborators[15]. Regarding the franchisee agents, Agent A contributed the most, requiring 101 collaborators; Agents B and C followed with 94 and 75, respectively; and Agent D with 59 collaborators. At the store level, our approach yields an average monthly staffing level of 4 active collaborators over the test period.

For the sake of simplicity, Figure 4.4 reports the results obtained for three individual stores. The stores chosen are associated with average, above average, and below average optimal staffing levels (*L**). The specific outcomes for each store and month are given in Annex C (tables 4.13 to 4.16).



(a) Average results

(b) Above average results

(c) Below average results

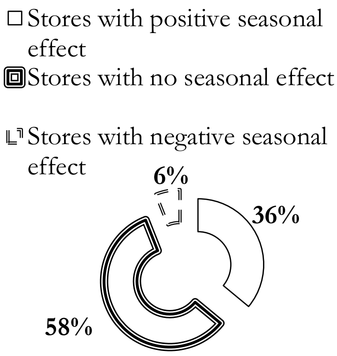
Figure 4.4. *Illustrative results of the optimal staffing levels and corresponding sales volume estimate, for an (a) average, (b) above average, and (c) below average store.*

[15] Number of employees available to work and that effectively contribute to the volume of sales, excluding employee absenteeism.

From Figure 4.4, it can be inferred that staffing level and sales volume have a similar pattern, over time, which shows the impact of sales collaborators on the sales performance.

Overall, the seasonality effect was well captured by our framework. While during the low season an average of 3 collaborators is required per month, in the high season (July, August, and December) this number rises to 4. This statement is also clearly evident in Figure 4.4, with the highest volatility in the optimal staffing levels coinciding with the summer and Christmas periods. These are not surprising results since the peak seasons are typically characterized by a higher volume of customers, which in turn requires a higher number of collaborators, so that the increased customer demand may be satisfied.

Furthermore, the seasonal differences were analyzed for each individual store, and the main conclusions are illustrated in Figure 4.5. The seasonal effect was measured as the difference between the average optimal monthly staffing levels required during the high season and the low season.



Store	# Sales personnel			
	Low season	High season		
		Total	Summer	Christmas
A13	8	10	10	9
A14	5	7	7	6
A29	6	8	8	9
B15	3	5	5	4
D1	3	5	5	4

Note: The reported values denote the average optimal staffing levels during each season. Total corresponds to July, August, and December; Summer to July and August, and Christmas to December.

Figure 4.5. *Seasonal differences in the staffing requirements: frequency and the most severe cases.*

As it can be seen, for the majority of the stores, the seasonal effect did not impact the staffing requirements; still, 36% of the store sample needs to be prepared with additional labour resources for the high season; and, surprisingly, 6% of the stores require one fewer sales collaborator during the high season. Within the stores with a positive seasonal effect most are street stores, which is the store format with the most significant store traffic increase from low season to high season (see the detailed data in Annex D, Table 4.17). On the other hand, mall stores are the ones in which the highest staffing level increase happens from November to December.

Regarding the labour productivity levels, as stated before, if the adopted staffing levels were the optimal allocation suggested by our approach, each sales collaborator would contribute with an average of 43 points per month, which is higher than the level requested by the company. Table 4.5 shows the top ten stores regarding with average labour productivity levels. In relative terms, 71% of the stores would be able to achieve at least the required labour productivity, while the remaining 29% would be under the desired value. Therefore, the majority of the stores would comply with the productivity desired by the company.

Table 4.5. *Stores with the highest average labour productivity levels yield by the optimal* L.

Stores	Average labour productivity
B1	64
A15	61
C12	59
B10	59
B20	58
D2	58
D17	57
C5	57
C10	56
B13	56

Note: The labour productivity levels were determined as the ratio between the estimated sales volume and the optimal staffing levels over the test period, approximated to units. Recall that company policy requires 43 points.

4.3. Validation of the results

In this section, a validation of the ideal staffing levels recommended by our labour-planning framework is performed. For that purpose, and using the estimated sales with the retailer's current staffing levels as a reference, we check whether the determined workforce size would contribute to a higher volume of sales throughout the test period.

Note that all the reported sales volume correspond to the sum of the monthly sales volume estimated by our model using both real and optimal staffing levels, over the test sample period, approximate to units; the sales growth rates stand for the discrete rate of change between the optimal and actual sales; and, finally, the labour increase is the difference between the average monthly staffing levels determined in this study and the actual levels.

First of all, this evaluation was performed individually for each franchisee agent and the retail chain as a whole, as given in Figure 4.6.

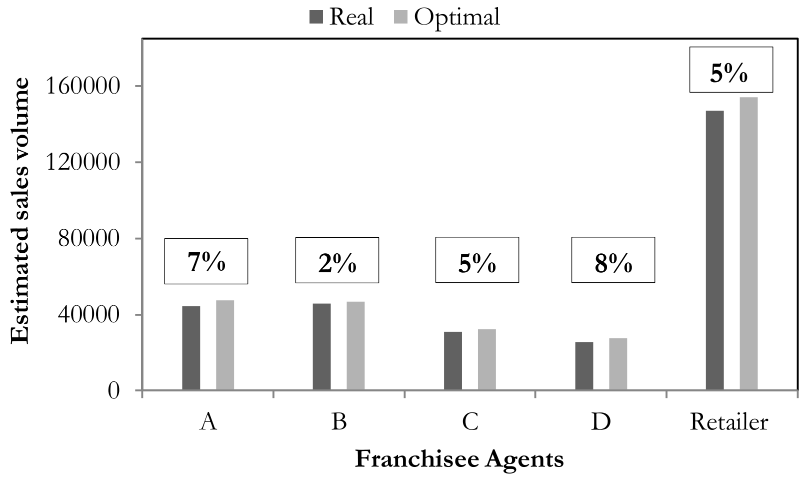


Figure 4.6. *Estimated sales volume with the real and the proposed workforce sizes for the test period along with the percentage sales increase displayed above the two series.*

From the analysis above, we conclude that the proposed staffing levels would be preferred over the retailer's current staffing plan, as they would generate a higher total sales volume. Indeed, we have estimated a sales increase of approximately 5%, which means 7000 additional points for the studied retailer. At the same time, all the agents involved would benefit as well from following our results, with Agent D achieving the highest sales growth rate and Agent B the lowest one.

Furthermore, the results were validated at the store level. The ten stores that would experience the most significant sales increase from implementing our staffing levels are displayed, in descending order of sales increase, in Figure 4.7, alongside with the corresponding sales evaluation. On the right panel of the figure, it has provided as well the implied monthly labour changes, in an average number of sales collaborators, to achieve these significant sales improvements.

Stores	Estimated Sales Volume	
	Real L	Optimal L
A26	911	1561
C9	848	1144
A18	1555	1965
D16	1543	1920
C5	1110	1378
D1	1573	1914
D18	648	780
D14	512	609
C16	724	848
D13	711	833

Note: The sales volume has been rounded to the nearest integer.

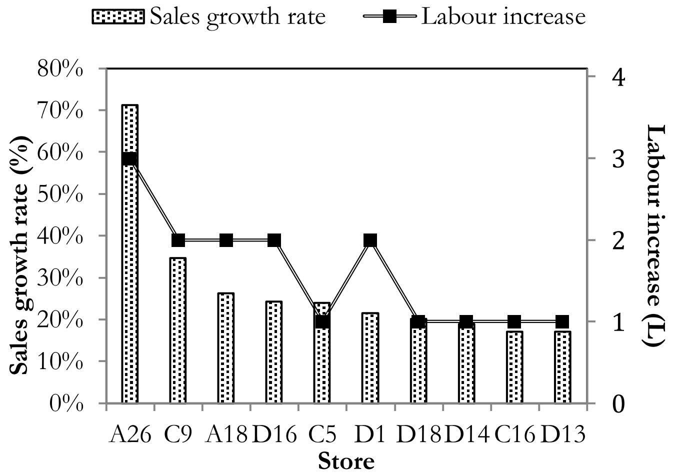


Figure 4.7. *Sales growth evaluation for the stores benefiting the most from our recommendation.*

From the results reported, it can be seen that store A26 would experience the most significant sales increase, thereby supporting the extreme average staffing increase of 3 sales collaborators advised by our approach. Although not so expressive, the growth rates of the remaining stores (C9, A18, D16, C5, D18, D14, C16, and D13) are still very large (between 17% and 35%) and require just one or two additional collaborators.

Regarding the stores whose staffing level was further apart from the optimal allocation, either due to insufficient or excessive staffing, the majority of them would still experience higher sales volumes by following our recommendations (see Figure 4.8); this way, supporting the labour changes proposed for these stores.

Stores	Estimated Sales Volume	
	Real L	Optimal L
A29	1287	1463
A26	911	1561
D1	1573	1914
A1	349	396
C4	2285	2481
A18	1555	1965
C2	2657	2941
B9	3659	3596
B21	1695	1677

Note: The sales volume has been rounded to the nearest integer..

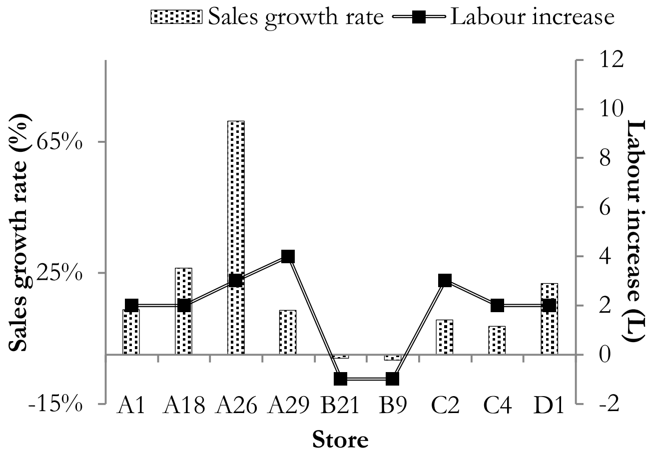


Figure 4.8. *Sales growth evaluation for the stores that were the furthest away from the optimal plan.*

Stores *B9* and *B21* would have a decrease in the sales volume (2% and 1%, respectively). However, such a decrease would allow to reduce the number of collaborators that could be allocated to another and more efficient stores. In addition, those two stores would become more efficient since the sales per collaborator would now be, respectively, 43 and 40 points; while before these figures were 39 and 38 points, respectively.

4.4. Implications of the findings

This section discusses some of the main managerial implications of the staffing level recommendations made in this study.

Figure 4.9 compares the actual average monthly staffing levels with the corresponding optimal levels retrieved by this study.

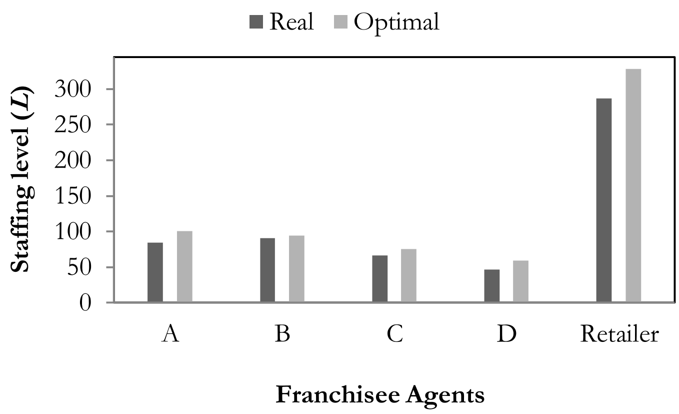


Figure 4.9. *Real and optimal average monthly staffing levels over the period March 2018 – January 2019.*

As mentioned before, the results of this study suggest that the retailer should have had on average, 328 active sales collaborators per month during the test period. Since the actual average staffing level is 287, the company was understaffing its stores by a total of 41 collaborators. As for the franchisee agents, although all of them were below the average recommendations, Agents A and D were the furthest away from the ideal number.

Table 4.6 reports the proportion of understaffed, overstaffed, and ideally staffed stores for each franchisee agent and for the whole store chain. The reported percentages were determined based on the optimal average monthly staffing level and the respective actual level for each store. We do not quantify how overstaffed or understaffed the store is, but rather only if it is or not.

Table 4.6. *Frequency of overstaffed, understaffed, and ideally staffed stores.*

Franchisee Agents	Overstaffing	Understaffing	Ideal
A	3%	47%	50%
B	29%	29%	42%
C	8%	36%	56%
D	7%	37%	56%
Retailer	11%	38%	51%

From the results above, it is possible to observe that about half of the stores were already employing the ideal staffing levels. However, it is also possible to observe that from the half that is not ideally staffed most were understaffed.

According to the company's perspective, the lack of sales collaborators was more evident in their low-staffed stores. This argument was proven empirically through our results since we found that 70% of the 40 understaffed stores were low-staffed[16]. Hence, this suggests that the retailer needs to plan more cautiously their staffing levels in these stores, in order to prevent the negative consequences associated with understaffing. The overstaffed stores were not targeted in this analysis because they do not represent a significant issue for the retailer (see Table 4.6).

[16] We considered a store to be low-staffed if its actual average staffing level is not above neither equal to the average staffing level, over all stores during the test period.

5. Final Remarks

This chapter concludes this research with a summary of the primary outcomes in Section 5.1 and suggestions for future work in Section 5.2.

5.1. Conclusions

In this study, we have proposed an efficient two-stage sales-based labour planning methodology to find an optimal monthly staffing level for each store of a retail chain. Firstly, the sales volume is forecasted for each individual store and month. Then, in the second phase, we establish, through several simulations, adequate staffing levels which ensure the best sales-to-labour relationship for a retailer. It is worth noticing that the outcomes of this approach are one possible set of ideal staffing levels, as it depends on the value chosen for the ratio between the percentage sales gain and labour increase.

The proposed framework is applied and validated in the franchised stores of a telecommunications company, as a case study. To our knowledge, this is the first empirical research that optimizes the workforce planning decision in the telecommunications sector; thereby, contributing significantly to the academic literature in this topic.

The volume of sales is initially forecasted using innumerous factors, including customer traffic, staffing level, location-specific attributes (competitors, buying power, and possession of a mobile service), selling area, and store type. This adds to our guideline estimation study by Chuang *et al.* (2016) in that they only rely on the first two variables. In addition, all variables were found to be statistically significant and more powerful to explain the sales behaviour of the analyzed stores in comparison to Chuang *et al.* (2016).

The proposed staffing plan ensures at least the minimum labour productivity levels required by the company for the majority of the stores; this way, presenting the implied staffing changes as profitable investments. All in one, we demonstrated the proposed staffing levels to be a practical and preferred fit for the retail chain compared to their actual staffing choices observed during the test period.

By following the optimal labour plan, we have projected for the retailer a sales growth rate of 5%, during the period March 2018 – January 2019; for each franchisee agent, the range of sales improvement would be 2% to 8%.

Lastly, the studied company is considering a pilot implementation of our results in the stores with fewer sales collaborators, which are the ones with the enormous urge to solve the staffing shortage.

5.2. Future work

A vital constraint when staffing is the labour cost associated with the proposed workforce size. Therefore, future work should augment our framework and incorporate the store profitability, defined as the difference between the sales revenue and store expenses, as the primary determinant in the staffing level decision.

Since the implementation and validation of the proposed approach are restricted to a single retail chain, we suggest upcoming studies to test it in other companies and retail settings as well to improve the generalizability of our methodology. In different background settings, it might be necessary to make minor changes in the specification of the sales model; for example, the possession of a mobile service is important in a telecommunications context, however, it might be irrelevant in others.

To simplify the labour planning process and provide retailers with broader staffing suggestions, forthcoming works can perform cluster analysis in the estimation procedure, i.e., group stores based on a specific criterion (such as the conversion rate, the actual staffing level, and the geographical area).

In terms of the specification of the proposed sales model, additional research may include the following aspects which were unavailable to us: labour heterogeneity (full-time, part-time, and temporary personnel); further information regarding the personnel and store manager (such as their experience and knowledge); monthly marketing actions (such as price promotions and free gift cards); and, finally, the distinction, in the store traffic variable, between the customers who effectively bought from those who only visited the store. All these factors are likely to have an impact on the sales performance and consequently on the outcomes of this approach as well.

Finally, future work can use the information on the specific recruitment period to set more accurately the instrumental variables correlated with the staffing level. Unfortunately, this was not possible to include in this study, since the studied company stated not having a fixed period to do so.

6. References

Ahipasaoglu, O., Erkip, N., & Karasan, O. E. (2019). The venue management problem: setting staffing levels, shifts and shift schedules at concession stands. *Journal of Scheduling*, 22(1), 69-83.

Arnold, T. J., Palmatier, R. W., Grewal, D., & Sharma, A. (2009). Understanding retail managers' role in the sales of products and services. *Journal of Retailing*, 85(2), 129-144.

Azmat, C. S., Hürlimann, T., & Widmer, M. (2004). Mixed integer programming to schedule a single-shift workforce under annualized hours. *Annals of Operations Research*, 128(1-4), 199-215.

Baltagi, B. (2008). *Econometric analysis of panel data*. New York: John Wiley & Sons.

Bliemel, F. (1973). Theil's forecast accuracy coefficient: A clarification. *Journal of Marketing Research*, 10, 444-446.

Bloom, N., & Van Reenen, J. (2007). Measuring and explaining management practices across firms and countries. *The quarterly Journal of Economics*, 122(4), 1351-1408.

Cotten, A. (2007). *Seven steps of effective workforce planning*. Washington, D.C.: *IBM Center for the Business of Government*.

Cullen, K., Robinson, S., Dolphin, C., Kubitschke, L., & Clarkin, N. (1997). *Telecommunications for older people and disabled people in Europe: preparing for the information society*. Amsterdam: IOS Press.

Chuang, H. H. C., Oliva, R., & Perdikaki, O. (2016). Traffic-based labor planning in retail stores. *Production and Operations Management*, 25(1), 96-113.

Chen, C., Twycross, J., & Garibaldi, J. M. (2017). A new accuracy measure based on bounded relative error for time series forecasting. *PloS one*, 12(3), e0174202.

Chapados, N., Joliveau, M., & Rousseau, L. M. (2011). Retail store workforce scheduling by expected operating income maximization. In *International Conference on AI and OR Techniques in Constriant Programming for Combinatorial Optimization Problems* (pp. 53-58). Springer, Berlin, Heidelberg.

Chapados, N., Joliveau, M., L'Ecuyer, P., & Rousseau, L. M. (2014). Retail store scheduling for profit. *European Journal of Operational Research*, 239(3), 609-624.

Croissant, Y., & Millo, G. (2008). Panel data econometrics in R: The plm package. *Journal of Statistical Software*, 27(2), 1-43.

Corominas, A., Lusa, A., & Pastor, R. (2002). Using MILP to plan annualised working hours. *Journal of the Operational Research Society*, 53(10), 1101-1108.

Eroglu, S. A., & Machleit, K. A. (1990). An empirical study of retail crowding: antecedents and consequences. *Journal of Retailing*, 66(2), 201.

Fisher, M., Krishnan, J., & Netessine, S. (2006). *Retail store execution: An empirical study*. University of Pennsylvania Working Paper.

Fisher, M., Gallino, S., & Netessine, S. (2017). *Setting Retail Staffing Levels: A Methodology Validated with Implementation*. Wharton Working Paper.

García, J. R. G. (2014). Using panel data techniques for social science research: an illustrative case and some guidelines. *CIENCIA ergo-sum*, 21(3), 203-216.

Ganster, D. C., & Dwyer, D. J. (1995). The effects of understaffing on individual and group performance in professional and trade occupations. *Journal of Management*, 21(2), 175-190.

Grewal, D., Baker, J., Levy, M., & Voss, G. B. (2003). The effects of wait expectations and store atmosphere evaluations on patronage intentions in service-intensive retail stores. *Journal of Retailing*, 79(4), 259-268.

Gujarati, D. N. (2009). *Basic econometrics* (5th ed.). New York: McGraw-Hill.

Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica: Journal of the econometric society*, 1251-1271.

Haans, H., & Gijsbrechts, E. (2011). "One-deal-fits-all?" On category sales promotion effectiveness in smaller versus larger supermarkets. *Journal of Retailing*, 87(4), 427-443.

Harris, J. H., & Arendt, L. A. (1998). Stress reduction and the small business: Increasing employee and customer satisfaction. *SAM Advanced Management Journal*, 63(1), 27.

Harrell, G. D., Hutt, M. D., & Anderson, J. C. (1980). Path analysis of buyer behavior under conditions of crowding. *Journal of Marketing Research*, 17(1), 45-51.

Heath, C., & Staudenmayer, N. (2000). Coordination neglect: How lay theories of organizing complicate coordination in organizations. *Research in Organizational Behavior*, 22, 153-191.

Hertz, A., Lahrichi, N., & Widmer, M. (2010). A flexible MILP model for multiple-shift workforce planning under annualized hours. *European Journal of Operational Research*, 200(3), 860-873.

Hise, R. T., Kelly, J. P., Gable, M., & McDonald, J. B. (1983). Factors Affecting the Performance of Individual Chain Store Units-An Empirical-Analysis. *Journal of Retailing*, 59(2), 22-39.

Hopp, W. J., Iravani, S. M., & Yuen, G. Y. (2007). Operations systems with discretionary task completion. *Management Science*, 53(1), 61-77.

Kabak, Ö., Ülengin, F., Aktaş, E., Önsel, Ş., & Topcu, Y. I. (2008). Efficient shift scheduling in the retail sector through two-stage optimization. *European Journal of Operational Research*, 184(1), 76-90.

Kesavan, S., Deshpande, V., & Lee, H. S. (2014a). Increasing sales by managing congestion in self-service environments: Evidence from a field experiment. *Available at SSRN 2523680.*

Kesavan, S., Staats, B. R., & Gilland, W. (2014b). Volume flexibility in services: The costs and benefits of flexible labor resources. *Management Science,* 60(8), 1884-1906.

Kim, S., & Kim, H. (2016). A new metric of absolute percentage error for intermittent demand forecasts. *International Journal of Forecasting*, 32(3), 669-679.

Lam, S., Vandenbosch, M., & Pearce, M. (1998). Retail sales force scheduling based on store traffic forecasting. *Journal of Retailing*, 74(1), 61-88.

Luceri, B., Latusi, S., Vergura, D. T., & Lugli, G. (2014). The impact of store flyers on store performance: a format and customer related approach. *International Journal of Retail & Distribution Management*, 42(3), 219-234.

Lusch, R. F., Serpkenci, R. R., & Orvis, B. T. (2015). Determinants of retail store performance: a partial examination of selected elements of retailer conduct. In *Proceedings of the 1995 World Marketing Congress* (pp. 495-504). Springer, Cham.

Mani, V., Kesavan, S., & Swaminathan, J. M. (2011). *Understaffing in retail stores: Drivers and consequences*. University of Pennsylvania Working Paper.

Mani, V., Kesavan, S., & Swaminathan, J. M. (2015). Estimating the impact of understaffing on sales and profitability in retail stores. *Production and Operations Management*, 24(2), 201-218.

McBurney, P., Parsons, S., & Green, J. (2002). Forecasting market demand for new telecommunications services: an introduction. *Telematics and Informatics*, 19(3), 225-249.

Mosadeghrad, A. M., Ferlie, E., & Rosenberg, D. (2011). A study of relationship between job stress, quality of working life and turnover intention among hospital employees. *Health Services Management Research*, 24(4), 170-181.

Netessine, S., Fisher, M., & Krishnan, J. (2010). *Labor planning. execution. and retail store performance: An exploratory investigation*. University of Pennsylvania Working Paper.

Ogumeyo, S. (2010). Optimum workforce-size model using dynamic programming approach. *Global Journal of Mathematical Sciences*, 9(1), 11.

Oliva, R., & Sterman, J. D. (2001). Cutting corners and working overtime: Quality erosion in the service industry. *Management Science*, 47(7), 894-914.

Park, H. M. (2011). Practical guides to panel data modeling: A step by step analysis using Stata. *Public Management and Policy Analysis Program, Graduate School of International Relations, International University of Japan*, 1-52.

Perdikaki, O., Kesavan, S., & Swaminathan, J. M. (2012). Effect of traffic on sales and conversion rates of retail stores. *Manufacturing & Service Operations Management*, 14(1), 145-162.

Prickett, J. (2018, March 29). Employee retention in retail is more important than ever. [Web log post]. Retrieved from https://www.tellermate.com/news-and-resources/employee-retention-retail-important-ever/

Quan, V. (2004, December). Retail Labor Scheduling. *OR/MS Today*. Retrieved from https://www.informs.org/ORMS-Today/Archived-Issues/2004/orms-12-04/Retail-Labor-Scheduling

Raman, A., DeHoratius, N., & Ton, Z. (2001). Execution: The missing link in retail operations. *California Management Review*, 43(3), 136-152.

Ryski, M. (2005). *When Retail Customers Count*. Bloomington: AuthorHouse.

Staats, B. R., Milkman, K. L., & Fox, C. R. (2012). The team scaling fallacy: Underestimating the declining efficiency of larger teams. *Organizational Behavior and Human Decision Processes*, 118(2), 132-142.

Siebert, W. S., & Zubanov, N. (2010). Management economics in a large retail company. *Management Science*, 56(8), 1398-1414.

Tan, F., & Netessine, S. (2014). When does the devil make work? An empirical study of the impact of workload on worker productivity. *Management Science*, 60(6), 1574-1593.

Talarico, L., & Duque, P. A. M. (2015). An optimization algorithm for the workforce management in a retail chain. *Computers & Industrial Engineering*, 82, 65-77.

Ton, Z. (2009). *The effect of labor on profitability: the role of quality*. Harvard Business School Working Paper 09-040.

Thompson, G. M. (1998). Labor scheduling, part 1: Forecasting demand. *Cornell Hotel and Restaurant Administration Quarterly*, 39(5), 22-31. doi: 10.1016/S0010-8804(98)89034-4

Verde Group & Baker Retailing Initiative (2007). Shoppers at Risk: An annual study of retail dissatisfaction. Retrieved from https://www.verdegroup.com/wp-content/uploads/2012/10/Shoppers-At-Risk-Exec-Summary_07.pdf

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. Cambridge: MIT press.

Annexes

Annex A: *Descriptive Statistics*

Table 4.7. *Descriptive statistics of the variables used in the complete sample.*

	Variable					
Statistics	*S*	*N*	*LA*	*PP*	*LC*	*MS*
Mean	130.8201	1228.4380	0.0029	93.4049	3.9645	90.0116
Median	98.5000	966.6688	0.0022	88.2000	2.0000	90.9359
Maximum	1185.0000	8538.0000	0.9739	214.5000	30.0000	95.0518
Minimum	0.2500	1.0268	0.0005	60.1000	0.0000	82.3702
Standard Deviation	118.9222	931.4845	0.0193	26.4424	5.0634	2.9861
Skewness	3.7478	2.5616	50.0666	2.8754	3.5063	-0.7061
Kurtosis	23.3323	12.6692	2524.4470	12.9732	17.1903	2.9888

Table 4.8. *Descriptive statistics of the variables used in the fit sample.*

	Variable					
Statistics	*S*	*N*	*LA*	*PP*	*LC*	*MS*
Mean	132.4177	1258.6300	0.0032	93.5616	4.0214	89.9176
Median	99.7500	985.7490	0.0021	88.2000	2.0000	90.7830
Maximum	1101.0000	8538.0000	0.9739	214.5000	30.0000	94.5663
Minimum	0.2500	1.0268	0.0005	60.1000	0.0000	82.3702
Standard Deviation	124.7916	977.9634	0.0269	26.5946	5.2175	2.8543
Skewness	3.6808	2.6869	35.8921	2.8783	3.5046	-0.7234
Kurtosis	21.0949	13.5319	1293.7659	12.9061	17.0485	2.9343

Table 4.9. *Descriptive statistics of the variables used in the test sample.*

	Variable					
Statistics	*S*	*N*	*LA*	*PP*	*LC*	*MS*
Mean	129.1589	1197.0939	0.0025	93.2420	3.9054	90.1092
Median	98.1250	951.4367	0.0023	88.2000	2.0000	90.9359
Maximum	1185.0000	6621.0000	0.0287	214.5000	29.0000	95.0518
Minimum	7.2500	126.2709	0.0006	60.1000	0.0000	82.3702
SD	112.5207	879.9337	0.0016	26.2928	4.8996	3.1154
Skewness	3.7993	2.3423	6.7852	2.8719	3.4936	-0.7057
Kurtosis	26.1701	10.8075	90.9155	13.0416	17.1954	3.0081

Annex B: *Correlation Coefficients*

Table 4.10. *Correlation Coefficients between the variables for the complete sample.*

	N	*LA*	*PP*	*LC*	*MS*	*A*
N	1.0000					
LA	-0.0462	1.0000				
PP	0.3223	-0.0037	1.0000			
LC	0.3505	-0.0077	0.9076	1.0000		
MS	0.1422	0.0025	0.2838	0.2948	1.0000	
A	0.3523	-0.0173	0.1242	0.1914	-0.0314	1.0000

Table 4.11. *Correlation Coefficients between the variables for the fit sample.*

	N	*LA*	*PP*	*LC*	*MS*	*A*
N	1.0000					
LA	-0.0499	1.0000				
PP	0.3143	-0.0060	1.0000			
LC	0.3594	-0.0141	0.9081	1.0000		
MS	0.1182	-0.0001	0.2696	0.2964	1.0000	
A	0.3568	-0.0264	0.1291	0.2028	-0.0373	1.0000

Table 4.12. *Correlation Coefficients between the variables for the test sample.*

	N	*LA*	*PP*	*LC*	*MS*	*A*
N	1.0000					
LA	-0.2505	1.0000				
PP	0.3323	0.0121	1.0000			
LC	0.3394	0.0623	0.9077	1.0000		
MS	0.1713	0.0755	0.2989	0.2959	1.0000	
A	0.3483	0.0398	0.1190	0.1788	-0.0257	1.0000

Annex C: *Optimal staffing levels*

Table 4.13. *Agent A' optimal staffing levels resultant from the iterative procedure.*

Stores	Months										
	Mar	Apr	May	Jun	Jul	Ago	Set	Out	Nov	Dez	Jan
A2	2	2	2	2	3	3	2	2	2	3	4
A3	1	1	1	1	1	2	2	2	2	2	2
A4	1	1	1	1	2	2	2	2	2	2	2
A5	5	5	4	5	6	5	6	5	5	6	6
A6	2	2	2	2	2	2	2	3	2	3	3
A7	3	3	3	4	5	5	4	4	3	4	3
A8	3	3	3	3	4	4	4	3	3	5	4
A9	2	2	2	2	3	3	2	2	3	2	3
A10	3	3	3	3	3	3	3	3	3	3	3
A11	3	3	3	3	4	5	4	4	4	4	4
A12	5	5	5	5	6	7	7	6	5	7	7
A13	7	7	7	6	9	10	8	8	8	9	8
A14	3	2	2	5	7	6	6	6	6	6	8
A1				2	2	2	2	2	2	2	3
A15	2	2	2	2	3	4	3	3	3	4	4
A16	3	3	3	2	3	3	3	2	2	3	2
A17	2	2	2	1	2	2	1	2	1	1	1
A18	4	3	4	4	4	4	4	4	4	4	3
A19	2	2	2	2	2	2	2	2	2	2	2
A20	2	2	2	2	2	2	2	2	2	2	2
A21	5	4	4	3	5	5	5	4	4	4	4
A22	5	4	4	4	4	5	6	4	4	5	5
A23	2	2	2	2	2	2	2	2	2	2	2
A24	2	2	2	2	2	2	2	2	2	2	2
A25	2	2	2	2	3	2	2	2	2	2	2
A26	3	3	3	3	4	4	3	3	3	3	3
A27	4	4	4	4	4	4	4	4	4	3	4
A28	6	6	6	5	6	7	6	6	5	5	6
A29	5	5	5	6	7	8	6	6	7	9	7
A30	1	1	2	1	2	2	2	1	1	2	2
Total	90	86	87	89	112	117	107	101	98	111	111

Table 4.14. *Agent B' optimal staffing levels resultant from the iterative procedure.*

Stores	Months										
	Mar	Apr	May	Jun	Jul	Ago	Set	Out	Nov	Dez	Jan
B1	2	2	2	2	2	2	2	2	2	2	2
B2	2	2	2	2	2	2	2	2	2	2	2
B3	4	4	3	4	4	4	4	3	4	5	4
B4	2	2	2	2	2	2	2	2	2	2	2
B5	2	2	2	2	2	2	2	2	2	2	2
B6	4	3	4	4	5	5	4	4	5	6	5
B7	11	10	10	10	11	10	12	12	10	12	11
B8	2	2	2	2	2	2	2	2	2	2	2
B9	7	7	6	6	8	9	7	9	8	9	9
B10	3	3	3	3	3	3	3	3	3	4	3
B11	13	13	11	12	13	15	15	15	14	16	14
B12	3	2	2	2	2	3	2	2	2	3	2
B13	8	6	8	8	9	8	9	9	9	9	9
B16	2	2	2	2	2	3	2	2	2	2	2
B17	3	3	2	2	3	4	3	3	3	3	3
B18	3	3	3	3	3	3	4	4	3	3	4
B19	2	2	2	2	3	3	3	3	2	2	3
B21	4	3	3	3	4	4	4	4	4	5	4
B22	4	3	3	3	4	4	4	4	4	5	4
B14	2	2	2	2	3	4	2	2	2	2	2
B15	3	3	3	3	4	5	3	3	3	4	3
B24	1	1	2	2	2	2	2	2	2	2	2
B20	1	1	1	1	2	2	1	1	1	1	1
B23	2	2	2	2	2	3	2	2	2	3	3
Total	90	83	82	84	97	104	96	97	93	106	98

Table 4.15. *Agent C' optimal staffing levels resultant from the iterative procedure.*

Stores	Months										
	Mar	Apr	May	Jun	Jul	Ago	Set	Out	Nov	Dez	Jan
C1	3	3	3	3	3	4	4	4	4	5	4
C2	7	7	7	7	9	8	8	7	8	9	10
C3	1	1	1	1	2	2	1	1	1	1	1
C4	4	3	4	4	5	6	6	6	6	7	5
C5	2	2	2	2	3	3	2	2	2	2	2
C6	7	7	7	8	9	9	8	8	8	7	8
C7	2	2	2	2	2	3	2	2	2	3	3
C8	2	2	2	2	2	2	2	2	2	2	2
C9	2	2	2	2	3	3	2	2	2	2	2
C10	3	2	3	2	3	3	3	3	3	4	3
C11	2	3	3	3	3	4	3	3	3	3	3
C12	2	2	2	2	2	2	2	3	3	2	3
C13	2	2	2	2	2	2	2	2	2	2	2
C14	2	2	2	2	2	2	2	2	2	2	2
C15	4	4	4	5	5	5	4	4	4	4	5
C16	2	2	2	2	2	2	2	2	2	2	2
C17	2	2	2	2	5	4	3	3	3	3	3
C18	4	4	3	3	3	3	3	3	2	3	3
C19	2	2	2	2	2	2	2	2	2	2	2
C20	3	3	3	2	2	2	2	2	2	2	2
C21	2	2	2	2	2	2	2	2	2	2	2
C22	2	2	2	2	3	3	3	3	2	2	3
C23	2	3	2	2	2	3	2	2	2	3	2
C24	2	2	2	2	2	2	2	2	2	2	2
C25	3	3	3	3	3	4	3	3	3	3	3
Total	69	69	69	69	81	85	75	75	74	79	79

Table 4.16. *Agent D' optimal staffing levels resultant from the iterative procedure.*

Stores	Months										
	Mar	Apr	May	Jun	Jul	Ago	Set	Out	Nov	Dez	Jan
D15	2	2	2	2	2	2	2	2	2	2	2
D1	4	3	3	2	4	5	3	3	3	4	3
D16	4	3	4	4	4	4	4	5	4	3	5
D3	1	1	1	1	1	2	1	1	1	1	1
D17	1	1	1	1	1	1	1	1	1	2	2
D14	2	2	2	2	2	2	2	2	2	2	2
D18	2	2	2	2	2	2	2	2	2	2	2
D2	1	1	1	1	1	1	1	1	1	1	1
D19	1	1	1	1	2	2	1	2	1	1	2
D20	2	2	2	2	2	3	2	2	2	3	2
D21	2	2	2	2	2	2	2	2	2	2	2
D22	2	2	2	2	2	2	2	2	2	2	2
D23	2	2	2	2	2	2	2	2	2	2	2
D4	2	2	2	2	2	2	2	2	2	2	2
D5	1	1	1	1	1	1	1	1	1	1	1
D6	2	2	2	2	2	2	2	2	2	2	2
D7	2	1	2	1	2	2	2	1	2	2	1
D8	6	5	5	5	5	7	6	5	5	7	6
D9	2	2	2	2	3	4	2	3	3	4	3
D10	2	2	3	3	3	4	3	3	3	4	4
D11	2	2	2	2	3	3	2	2	2	3	3
D24	2	2	2	2	2	2	2	2	2	2	2
D25	2	1	1	2	2	2	1	2	2	2	2
D12	2	2	2	2	2	2	2	2	2	2	2
D26	2	2	2	2	2	2	2	2	2	2	2
D13	2	2	2	2	2	2	2	2	2	2	2
D27	2	2	2	2	2	2	2	2	2	2	2
Total	57	52	55	54	60	67	56	58	57	64	62

Annex D: *Store traffic during high season and low season months for each store type*

Table 4.17. *Comparison between the average monthly store traffic during the high season and low season for each store format.*

Format	N_H	N_L	$\triangle$
Shopping mall	139761	45191	94570
Citizen store	184337	4906	179431
Gallery	152194	17496	134699
Street	476292	65950	410342

Notes: N_H denotes the average monthly store traffic during the high season (July, August, and December), and N_L represents the same but for the low season. The reported values have been rounded up to the nearest integer. $\triangle$ stands for the store traffic difference between the high season (N_H) and the low season (N_L).