Ca' Foscari
University
of Venice

Master Degree programme

in Data Analytics
for Business and
Society

Final Thesis

# Customer, Product, and Stock Barriers to B2B E-Commerce Adoption: Evidence from Prysmian Spain's Digital Platform

**Supervisor**
Ch. Prof. Francesco Rullani

**Graduand**
Chiara Golinelli
Matriculation Number 905996

**Academic Year**
2025 / 2026

# Abstract

This thesis contributes to B2B e-commerce adoption theory by showing, through the case of Prysmian Spain's digital platform, that adoption depends on multi-level alignment rather than platform access alone. While digital transformation has increased the strategic importance of online channels, many industrial firms continue to experience a significant gap between platform availability and actual customer usage.

The study addresses this issue by examining how customer characteristics, product catalog coverage, order-level conditions, and stock availability influence the migration from traditional ordering channels to a digital B2B environment. Drawing on internal SAP and Salesforce data covering all orders sold in Spain during 2025, the research adopts a multi-level analytical framework that combines descriptive statistics, K-means clustering, logistic regression models, ordinary least squares estimation, and propensity score matching.

The findings reveal that B2B adoption is not a binary outcome driven solely by platform access, but rather a process shaped by the alignment of customer readiness, product suitability, transaction feasibility, and operational reliability. A substantial gap emerges between customer enablement and actual platform usage, while customer segmentation identifies a large group of high-potential users characterized by strong catalog compatibility but limited digital engagement.

At the product level, standardized and recurrently purchased items are significantly more likely to be included in the digital catalog than customized or technically complex products. Furthermore, stock availability proves to be a critical operational factor, as higher inventory reliability is associated with a lower likelihood of offline ordering.

The study contributes to the literature by conceptualizing B2B e-commerce adoption as a multi-level alignment process involving customers, products, orders, and inventory conditions. From a managerial perspective, it proposes a data-driven framework for prioritizing customer activation, catalog expansion, and stock allocation, providing practical guidance for improving digital channel adoption and commercial performance in industrial B2B settings.

*"E per quanta strada
ancora c'è da fare,
amerai il finale."*

Cesare Cremonini, *Buon Viaggio*

## Table of Contents

# List of Tables

# List of Figures

# Introduction

## Digital adoption and barriers

Digital transformation has made e-commerce adoption a critical requirement for competitive advantage across both B2B and B2C markets (Plekhanov et al., 2023). Small and medium-sized enterprises (SMEs), as well as large industrial firms, are increasingly shifting their operations to digital platforms in order to expand market reach, improve knowledge management, and improve efficiency (Naunthong, 2022; Pérez Trujillo & Aleán-Romero, 2024; Shivakumar & Suresh, 2014). The COVID-19 pandemic further accelerated this transition, forcing organizations to adopt online sales models more rapidly to compensate for the reduction in physical interactions and maintain business continuity (Bravo, Gonzalez Segura, Temowo, & Samaddar, 2022; Amornkitvikai et al., 2022). However, managing this transition successfully requires strong strategic commitment and the ability to address a complex set of benefits, risks, and organizational obstacles (Ballerini et al., 2023; Wang & Ahmed, 2009).

Despite these potential benefits, the literature identifies numerous functional, psychological, and organizational barriers that hinder e-commerce adoption and continued usage. At the organizational level, a lack of IT expertise, inadequate digital infrastructure, financial constraints, and rigid strategic orientations often prevent firms from fully exploiting their digital potential (Styvén & Wallström, 2019; Sell et al., 2019). In addition, firms face significant operational challenges at the marketing-operations interface, particularly in assortment planning, limited inventory availability, and frequent stockouts, all of which can disrupt the customer journey and reduce the likelihood of online ordering (Bijmolt et al., 2021; Derhami, Montreuil, & Bau, 2021). In some contexts, these constraints are further reinforced by poor delivery infrastructure and limited knowledge of digital logistics (Boom-Cárcamo et al., 2024).

From the customer perspective, innovation resistance theory suggests that adoption is often constrained by tradition, perceived risk, and usage barriers (Talwar et al., 2021; Dimitrova, Öhman, & Yazdanfar, 2022). Among the most relevant obstacles are concerns over data privacy and security, especially when online platforms fail to provide transparent policies or credible trust signals (Beatty et al., 2007; Tchao, Diawuo, Aggor, & Kotey, 2017). In both consumer and B2B settings, customers may also perceive digital channels as less trustworthy, less flexible, or more impersonal than traditional face-to-face interactions (Sharma, Gupta, & Joshi, 2020; Yang et al., 2011). This resistance may be further reinforced by limited supplier flexibility, insufficient digital training, and broader forms of digital exclusion (Ghosh & Dash, 2023; Xiao, Xu, Xiao, Wang, & Skare, 2024).

To overcome these barriers and support migration toward digital channels, firms must develop strategies that effectively integrate physical and virtual touchpoints. The literature shows that phygital services and broader omnichannel ecosystems can enhance customer engagement, reduce friction across channels, and improve repurchase intentions (George, Panattil, & Sajid, 2025; Rahman et al., 2025). Developing an effective omnichannel ecosystem requires the harmonization of online and offline touchpoints, the reduction of price inconsistencies, and coordinated information sharing across customer-facing functions (Chou et al., 2025). In restricted B2B environments, targeted customer training, and digital service support are particularly important, as they can strengthen customers' confidence in the platform and improve the overall effectiveness of e-commerce initiatives (Costa & Rodrigues, 2024; Evans, 2023). More broadly, successful channel migration depends on the alignment of customer expectations, supplier capabilities, and platform functionality (Anand et al., 2025).

Within this evolving landscape, Artificial Intelligence (AI) is increasingly discussed as a promising enabler of digital adoption and e-commerce efficiency. AI-based tools can improve customer-facing services through chatbots, virtual assistants, and automated problem-solving functions, while also supporting back-end processes such as demand forecasting, inventory monitoring, and product availability management (Papastamoulou & Antonopoulos, 2025; Gao, Opute, Jawad, & Zhan, 2025). In addition, AI-generated summaries and recommendation systems may improve the perceived usefulness and trustworthiness of digital catalogs by reducing information overload and facilitating navigation (Bian & Che, 2025). Although AI is not the central focus of this thesis, it represents a relevant complementary perspective, particularly in contexts where operational frictions and information asymmetries limit platform adoption.

## Research context: Prysmian Spain

Against this theoretical background, this thesis examines the case of Prysmian's B2B e-commerce platform in the Spanish market. Prysmian is an Italian multinational company and a global leader in the cable systems industry, operating in the energy and telecommunications sectors. The company designs, manufactures, and supplies cables and integrated solutions for the transmission and distribution of electricity, as well as for digital connectivity and telecommunications infrastructure. Prysmian has a broad international presence, operating in more than 50 countries through over 100 production facilities and multiple research and development centers. Its business model is primarily business-to-business, serving industrial customers, utilities, installers, telecommunications operators, and other professional clients rather than final consumers.

Prysmian's offering spans a wide range of products and systems, including underground and submarine power cables, solutions for low-, medium-, and high-voltage transmission, and telecom cables and fiber-optic systems. Given its scale, industrial complexity, and strong international footprint, Prysmian represents a particularly relevant context for studying digital channel transformation in B2B markets. In sectors characterized by technical product specifications, customized solutions, and complex customer relationships, the shift from traditional sales and service channels to digital self-service platforms is often gradual and affected by multiple organizational and operational barriers.

Prysmian's digital strategy includes a dedicated B2B platform designed to facilitate online ordering, streamline customer interactions, and reduce reliance on traditional channels such as call centers and email. Despite this strategic intent, platform adoption in Spain remains partial. Only selected customer segments, especially Trade and Installers, are currently enabled; the product catalog is limited to more standardized items; and important operational constraints, including stock availability and customized production requirements, continue to shape both customer behavior and internal commercial practices.

For these reasons, Prysmian Spain represents a particularly useful empirical setting in which to investigate why digital migration remains incomplete and which barriers are associated with B2B platform adoption. The case is characterized by a combination of organizational, operational, and customer-related frictions, including limited B2B usage, catalog constraints, inventory shortages, price complexity, and internal resistance. This thesis treats B2B adoption as a multi-level process involving customer access, order-level channel choice, product catalog feasibility, and stock reliability. More specifically, it seeks to identify why some customers actively use the platform while others do not, why many enabled customers still rely on manual ordering, and how product availability and stock conditions influence channel choice.

Although the empirical analysis focuses on Prysmian Spain, the case is used to address a broader theoretical problem in B2B e-commerce research: why the existence of a digital platform and the formal enablement of customers do not automatically translate into actual online ordering. The thesis therefore treats Prysmian Spain as a setting through which to examine the conditions under which digital channel migration occurs in industrial B2B markets.

# Chapter 1. Theoretical Background and Research Gap

This chapter sets the theoretical basis for the thesis by explaining why B2B e-commerce adoption is not only a technological issue, but also a matter of customer behavior, product availability, and operational reliability. It then identifies the main research gap in the literature and presents the research question that guides the analysis.

Section 1.1 reviews the literature on B2B e-commerce adoption and explains why customers do not move online simply because a digital platform exists. It highlights the importance of customer fit, purchasing habits, product coverage, and stock availability in shaping adoption.

Section 1.2 presents the main factors that either hinder or encourage the shift from offline to online ordering. It discusses barriers such as resistance to change, lack of trust, limited assortment, and stock-outs, as well as drivers such as convenience, completeness, reliability, and customer enablement.

Section 1.3 explains what the existing literature does not fully capture, especially in restricted B2B industrial settings where customers, products, and stock conditions interact. It then introduces the central research question of the thesis, which examines how B2B actual usage, assortment completeness, and stock availability are associated with B2B e-commerce adoption.

## 1.1 Background and motivation based on literature review

This chapter frames B2B e-commerce adoption as more than a simple technological upgrade. The literature suggests that moving from offline to online channels does not depend only on the existence of a digital platform, but also on how well that platform fits the customer's purchasing habits and product needs (Amornkitvikai et al., 2022; Ghosh & Dash, 2023). A useful way to understand this process is to view digital channel adoption as gradual rather than immediate (Yang et al., 2011).

Customers do not necessarily move from offline to online all at once. Even when a digital channel is available, they may continue to rely on traditional ordering methods if the platform does not provide sufficient product coverage, reliable stock availability, or a simple purchasing experience (Bijmolt et al., 2021; Derhami et al., 2021). For this reason, adoption is not only about whether customers are enabled, but also about how much they actually use the channel and whether it can realistically replace offline ordering (Ghosh & Dash, 2023).

This issue is particularly relevant in industrial B2B settings, where the purchasing process is often more complex than in consumer markets. Customers frequently buy technical products with specific requirements, may need support from sales representatives, and expect a high level of reliability in product availability (Derhami et al., 2021; Ghosh & Dash, 2023). As a result, the shift to online channels is often slowed down by operational frictions and by customers' preference for established routines (Talwar et al., 2021).

The literature also shows that the success of digital transformation depends on the interaction between the digital channel and the operational structure behind it (Bijmolt et al., 2021). A platform can support adoption only if it is complete, reliable, and easy to use. If customers encounter missing products, stock-outs, or a mismatch between the online catalog and their actual needs, their experience with the digital channel may be disappointing, leading them to continue ordering through traditional manual channels (Bijmolt et al., 2021; Derhami et al., 2021).

This is why the present thesis focuses on the concrete factors that influence channel migration in a B2B environment. More specifically, it examines customer B2B usage, product compatibility with the B2B catalog, and stock availability as three interrelated conditions that shape the actual use of the online channel (Evans, 2023; Yang et al., 2011).

## 1.2 Barriers and drivers of customers' shift to online channels

The literature identifies several barriers that explain why customers may not shift from offline to online ordering. On the customer side, one of the main barriers is resistance to change (Yang et al., 2011). Customers are often accustomed to offline routines, may prefer direct human interaction, or may perceive online ordering as less flexible than relationship-based purchasing (Dimitrova, Öhman, & Yazdanfar, 2022). In B2B contexts, this can be reinforced by long-standing commercial ties with sales representatives, which create inertia and reduce the incentive to switch channels (Ghosh & Dash, 2023).

Trust is another important barrier (Tchao, Diawuo, Aggor, & Kotey, 2017). Customers may hesitate to use online channels if they do not trust the platform, the quality of product information, or the way private data are managed (Beatty et al., 2007). If the digital channel is perceived as incomplete or unreliable, customers may continue to rely on offline communication even when online ordering is technically possible.

Product and inventory availability are also central barriers. The literature shows that limited assortment planning and frequent out-of-stock events can strongly disrupt the customer journey (Bijmolt et al., 2021). If customers cannot find all the products they need in the digital catalog, or if they are uncertain about stock availability, they may return to offline ordering (Derhami, Montreuil, & Bau, 2021). If offline ordering appears more complete or more dependable, online adoption becomes less likely.

There are also organizational barriers on the supplier side (Amornkitvikai et al., 2022). Firms may lack digital expertise, may not have adequate infrastructure, or may fail to align internal incentives with channel migration (Styvén & Wallström, 2019; Naunthong, 2022). In B2B environments, sales teams may also resist digitalization if they fear losing control over customer relationships (Plekhanov et al., 2023). These internal frictions matter because a platform may formally exist, but still not be used actively if the organization does not support it (Wang & Ahmed, 2009).

At the same time, the literature points to several drivers that can encourage customers to move online. The first is convenience: when the digital channel makes ordering easier, reduces transaction costs, and simplifies repeat purchases, customers are more likely to adopt it (Costa & Rodrigues, 2024).

The second is completeness: customers are more willing to use B2B ordering when the catalog contains the products they need and when the platform fits their usual purchasing patterns (Bijmolt et al., 2021).

The third is reliability: online adoption becomes easier when customers trust that products will be available and that the ordering process will work without friction (Derhami, Montreuil, & Bau, 2021).

Another relevant driver is targeted enablement. Customers may need to be activated, trained, or supported before they can use the B2B platform effectively (Evans, 2023). This is especially important in restricted B2B environments, where not all customers are enabled in the same way. In this sense, enabling the right customers is not only a technical decision, but also a strategic one (Ballerini et al., 2023).

Finally, the literature suggests that digital adoption is easier when physical and digital channels are integrated into a coherent system (Chou et al., 2025). Even if this thesis does not focus broadly on omnichannel strategy, the underlying idea remains important: customers are more likely to shift when the digital channel complements the offline channel rather than replacing it in a disruptive way (Rahman et al., 2025). In practice, this means that good coordination among customer service, product availability, and ordering functionality can reduce resistance and improve usage (Bijmolt et al., 2021).

## 1.3 Research gap and research question

Existing research on B2B e-commerce adoption has often focused on customer-level willingness, organizational readiness, or perceived barriers to digital use. However, in restricted industrial platforms, adoption may fail even when customers are formally enabled and willing to use the channel, because the specific products, orders, or stock conditions required for digital transactions are not fully supported (Amornkitvikai et al., 2022; Yang et al., 2011; Bijmolt et al., 2021; Derhami, Montreuil, & Bau, 2021; Anand et al., 2025). This thesis addresses this gap by conceptualizing B2B adoption as a multi-level alignment problem involving customer access, product catalog feasibility, transaction characteristics, and operational reliability (Ghosh & Dash, 2023; Plekhanov et al., 2023; Beatty et al., 2007; Bijmolt et al., 2021; Derhami, Montreuil, & Bau, 2021; Talwar et al., 2021).

The empirical case of Prysmian Spain makes this question especially relevant. The platform is already in place, but adoption remains partial, the catalog is incomplete, and stock conditions continue to influence customer behavior. This makes Prysmian Spain a useful setting in which to study not only whether B2B works, but also where the main frictions are concentrated and how they may be addressed. This thesis contributes by showing that restricted B2B platform adoption depends on the interaction of customer access, catalog feasibility, and stock reliability at multiple empirical levels.

The central research question can therefore be formulated as follows: What customer-, order-, product-, and stock-level factors explain the gap between B2B platform enablement and actual B2B ordering in industrial B2B companies?

To address this research question, the study is structured around the following specific objectives:

O1: To measure the gap between formal B2B platform enablement and actual B2B usage in industrial B2B companies, considering customers, orders, turnover, and product coverage.

O2: To segment B2B-enabled customers according to their actual digital purchasing behavior and their compatibility with the current B2B catalog, in order to distinguish between mature digital users, latent adopters, and customers with limited structural fit.

O3: To identify the order-level factors associated with continued offline ordering among B2B-enabled customers, with particular attention to transaction value, catalog feasibility, stock availability, and temporal variation in channel choice.

O4: To examine which product characteristics are associated with B2B catalog inclusion and to identify high-demand products currently purchased offline that could be prioritized for future catalog expansion.

O5: To assess how stock availability and stock-related constraints are associated with B2B ordering behavior and to identify products for which inventory reliability may represent a barrier to digital channel migration.

O6: To develop a data-driven prioritization framework that supports managerial decisions on customer activation, catalog expansion, and stock allocation in industrial B2B e-commerce development.

In line with the literature on B2B e-commerce adoption, digital channel migration, product assortment, and operational reliability, the empirical analysis is guided by the following hypotheses.

H1: Among B2B-enabled customers, higher catalog compatibility, measured as the proportion of purchased products available in the digital catalog, is positively associated with actual B2B platform usage.

H2: Higher stock availability at order level is associated with a lower probability of offline ordering among B2B-enabled customers.

H3: Standardized and recurrently purchased products are more likely to be included in the B2B catalog than customized, project-based, or technically complex products.

H4: Customers with stronger prior exposure to electronic ordering channels and organizational contexts more aligned with digital purchasing are more likely to become active B2B platform users.

H5: Higher-value orders, used as a proxy for greater transaction complexity and need for commercial coordination, are more likely to be placed through offline channels than through the B2B platform.

# Chapter 2. Data and Methodology

This chapter describes the data sources, sample selection criteria, and empirical strategy adopted to investigate the barriers that hinder customers' shift to the B2B e-commerce channel in Prysmian Spain. The analysis is based on internal transactional and customer data extracted from SAP and Salesforce, two core systems that track Prysmian's sales, ordering, and customer management processes. The data are complemented by internal documents, reports, and qualitative evidence gathered through informal interviews with managers and sales personnel. These qualitative inputs provide context for interpreting the quantitative patterns and help clarify how organizational routines, sales incentives, and operational constraints shape customers' channel choices in practice (Ballerini et al., 2023).

This mixed-methods case study design is appropriate for the objectives of the thesis, as it combines detailed firm-level records with contextual qualitative evidence. By integrating quantitative analysis with qualitative insights, the study aims not only to measure the extent of B2B platform adoption, but also to identify the underlying mechanisms that drive customers to remain on offline channels. This approach is particularly suitable in a B2B industrial setting, where customer relationships are often long-term, orders may involve technical complexity, and stock availability plays a central role in service reliability (Ballerini et al., 2023; Ghosh & Dash, 2023).

The chapter is organized in two sections.

Section 2.1 presents the data sources and sample selection criteria, with a detailed description of the customer-level, order-level, and product-level datasets.

Section 2.2 outlines the empirical strategy, including the construction of key indicators, the descriptive analysis, and the regression-based models used to assess the associations among B2B usage, catalog completeness, and stock availability with the likelihood of placing orders through the B2B platform rather than through offline channels.

Together, these elements establish the methodological foundation for the empirical analysis and for the subsequent customer, product, and stock prioritization developed in Chapter 10, where customer-, product-, and stock-level configurations are evaluated based on their commercial potential, B2B compatibility, operational feasibility, and likelihood of migration to the digital B2B channel.

## 2.1 Data sources and sample selection

The empirical analysis relies on data extracted from Prysmian's SAP and Salesforce systems. More specifically, the study uses the Cockpit dataset and focuses exclusively on orders sold from Spain to customers located in Spain during the 2025 year. To ensure that the sample captures only market-based customer transactions, intercompany orders were excluded from the analysis.

The final dataset is structured around three main levels of analysis: customer, order, and product. This multi-level structure makes it possible to link customer characteristics, order composition, and product availability conditions to the observed use of the B2B channel. In this way, the dataset supports both descriptive and inferential analyses of the factors associated with online versus offline ordering behavior.

### 2.1.1 Customer-level data

At the customer level, the dataset includes customer ID, customer name, province, company, total turnover, and a binary indicator of whether the customer is enabled for the B2B platform. This enablement indicator is central to the thesis, as it captures one of the key operational levers that may influence customers' likelihood of adopting the online channel.

In addition, the customer-level database includes detailed information on orders placed and products purchased, which makes it possible to construct indicators related to purchasing intensity, order frequency, and product mix (Bijmolt et al., 2021). These variables are important because they help capture the operational conditions that may facilitate or hinder channel migration.

A relevant point concerns business segmentation. Most customers currently enabled for the B2B platform belong to the HFM business category "Trade and Installers" (T&I), mainly because they purchase standardized products that are generally available in the B2B catalog (Costa & Rodrigues, 2024). However, this classification should not be interpreted as a strict requirement for B2B enablement, since customers outside the T&I category may also be enabled.

Customers are assigned to a single business segment based on the types of products they purchase and the nature of their business activity. However, this classification is a simplification, because some customers operate across multiple business lines and purchase a diversified mix of products that does not fit neatly into a single category. As a result, the segmentation label may not fully capture the heterogeneity of a customer's actual business (Anand et al., 2025).

### 2.1.2 Order-level data

At the order level, each observation corresponds to a single order. The dataset includes order ID, purchase order type, order turnover, product details, and a binary variable distinguishing between online orders placed via the B2B platform and offline orders entered through manual channels.

This level of analysis is particularly important because channel choice is observed at the order level rather than only at the customer level. A single customer may place some orders online and others offline, depending on the composition of the order, the products requested, stock conditions, or the need for direct interaction with sales personnel (Sharma et al., 2020; Chou et al., 2025).

The order-level perspective therefore provides a more precise basis for identifying the determinants of channel choice. It makes it possible to analyze whether offline ordering is associated with catalog incompatibility, stock-related problems, order complexity, or other transactional characteristics that are not visible when the customer is treated as the only unit of analysis (Ghosh & Dash, 2023; Rahman et al., 2025).

### 2.1.3 Product-level data

At the product level, the dataset includes product ID, product name, product class, and a binary indicator of whether the product is available in the B2B catalog. In addition, the product-level data contain stock-related indicators, such as the percentage of time the product is out of stock and the frequency of stock-outs.

These variables are particularly relevant for the present research, since catalog completeness and product availability are among the main operational barriers that are associated with customers' migration to the online channel. If relevant products are not included in the B2B catalog, or if they are frequently unavailable, customers may prefer to rely on offline channels even when they are formally enabled to use the platform (Talwar et al., 2021; Derhami et al., 2021).

## 2.2 Empirical strategy

The empirical strategy adopted in this thesis aims to quantify how customer B2B usage, assortment completeness, and stock availability influence the likelihood that orders are placed through the B2B platform rather than through offline channels (Derhami et al., 2021). To do so, the study combines descriptive analysis with regression-based inference, using data extracted from SAP and Salesforce and integrating the quantitative results with qualitative insights from internal documents and informal interviews (Ballerini et al., 2023).

The first step consists of constructing a consolidated database by cleaning and merging customer-, order-, and product-level information through unique identifiers such as customer ID, order ID, and product ID. During this phase, duplicates are removed, codes are harmonized, and variables such as business segment and product class are standardized. Based on the merged dataset, several indicators are then created to capture the core dimensions of interest. These include measures of assortment completeness, such as the share of products purchased by a customer that are available in the B2B catalog, and stock-related measures, such as stock-out frequency and the share of order line items associated with products in stock (Bijmolt et al., 2021; Derhami et al., 2021).

The empirical analysis proceeds in two main stages. First, descriptive statistics and key performance indicators are used to examine the distribution of B2B enablement across customers, the share of turnover generated through B2B versus offline channels, the degree of catalog coverage relative to purchased products, and stock-out patterns across product classes. This first stage provides an initial overview of the relationship between B2B actual usage, assortment availability, stock conditions, and actual channel usage.

Second, inferential analyses are conducted at the customer, order, product and stock levels. At the customer level, a logistic regression model estimates the probability that a customer is actually active on the B2B platform, measured by whether the customer placed at least one B2B order. The final specification includes prior electronic ordering behavior, measured by *ceec_orders*, the customer's organizational group, measured by *first_level_reduced*, and geographic macro-area, measured by *province_macro*. Given the imbalance between B2B-active and non-active customers, the model is interpreted primarily as a ranking and prioritization tool rather than as a strict binary classifier.

At the order level, a second logistic regression estimates the probability that an order placed by a B2B-enabled customer is processed offline rather than through the B2B platform. The final model includes *log_order_turnover*, used as a proxy for the economic relevance and potential complexity of the transaction, *pct_in_stock_products*, which captures the share of products in the order that were available in stock, and monthly fixed effects to control for seasonality and time-specific shocks. Catalog-composition variables and line-item measures are not included in the final specification because they are mechanically linked to the channel outcome.

At the product level, a logistic regression estimates the probability that a product is included in the B2B catalog. The final specification includes product-class controls, *percentage_orders*, *product_turnover, total_orders, orders_from_b2b_customers,* and *average_orders_per_customer*. This model is used to identify which product characteristics are associated with catalog inclusion and to support catalog-expansion prioritization.

At the stock level, an ordinary least squares regression estimates product-level stock availability, measured by *pct_in_stock_product*. The model includes *log_product_turnover*, *log_total_orders, log_orders_from_b2b_customers, average_orders_per_customer*, and product-class controls. This specification is used to identify which product characteristics are associated with higher or lower in-stock availability. The results are interpreted as conditional associations rather than causal effects, since stock allocation and product demand may influence each other.

Overall, this combined descriptive and regression-based approach makes it possible to connect concrete operational levers, such as customer B2B usage, catalog expansion, and stock allocation, to observable channel outcomes. As a result, the empirical strategy provides an evidence-based framework for understanding the barriers to B2B e-commerce adoption in Prysmian Spain and for assessing which improvements may most effectively support customers' migration to the online channel.

A key methodological distinction in this thesis is the difference between prediction, association, and causal reasoning. Predictive models are useful for identifying customers, products, or orders that resemble current B2B users, but they do not necessarily indicate which managerial levers can change adoption outcomes. This distinction is important because the research question is not only descriptive, but also intervention-oriented: customer B2B usage, catalog inclusion, and stock allocation are decisions that the firm can potentially modify. For this reason, the empirical strategy combines predictive and explanatory models with a matching-based attempt to approximate causal reasoning. The matching analysis in Chapter 9 does not provide definitive causal identification, given the observational nature of the data, but it represents a necessary methodological step toward evaluating whether actionable levers are associated with different outcomes after improving comparability between treated and untreated units.

# Chapter 3. Empirical Context: Prysmian and the Spanish B2B Platform

This chapter presents the empirical context of the thesis by focusing on Prysmian's market in Spain and on the current development of its B2B e-commerce platform. The analysis refers to orders sold in Spain to customers located in Spain during 2025 and provides an overview of the current state of digital adoption in the Spanish business environment.

The chapter has two sections.

Section 3.1 describes the specific characteristics of the Spanish market and explains why it represents a relevant setting for studying B2B digital channel migration.

Section 3.2 presents a set of general key performance indicators that provide an initial picture of platform adoption in terms of turnover, customers, products, and orders.

Together, these elements help frame the empirical analyses developed in the following chapters by highlighting the gap between formal platform availability and actual usage, as well as the operational constraints that continue to limit migration toward the online channel.

## 3.1 Description of the Spanish market

From the perspective of this thesis, the Spanish market is particularly relevant because it combines the characteristics of a mature industrial B2B environment with the operational complexity of partial digitalization. The Spanish customer base includes a variety of professional clients, such as installers, distributors, and other business customers operating across sectors linked to construction, electrification, telecommunications, and industrial applications.

A central feature of the Spanish B2B platform is that customer enablement remains only partial. Not all customers have access to the platform, and most enabled customers belong to the Trade and Installers segment, which largely overlaps with the group of customers purchasing more standardized products that are easier to include in a digital catalog. This segmentation logic reflects the current phased approach to digitalization, in which the platform is primarily designed around product categories and customer profiles that are more compatible with self-service ordering.

The Spanish B2B platform therefore represents an appropriate setting for examining how customer B2B usage, product catalog availability, and operational constraints influence the adoption of the online channel.

## 3.2 General KPIs

The general KPIs provide an initial overview of the relative weight of the online channel within Prysmian Spain in 2025. Overall, the data show that the B2B platform still plays a marginal role compared with traditional ordering channels, both in terms of turnover and in terms of number of orders. At the same time, the platform is not irrelevant, as a meaningful group of customers has already been enabled, a non-negligible number of products is available online, and a subset of customers actively uses the B2B channel.

The KPI analysis is therefore useful not only to describe the current state of adoption, but also to contextualize the gap between enablement and actual online usage. This descriptive evidence provides the basis for the more detailed analyses developed in the following chapters.

### 3.2.1 Turnover KPIs

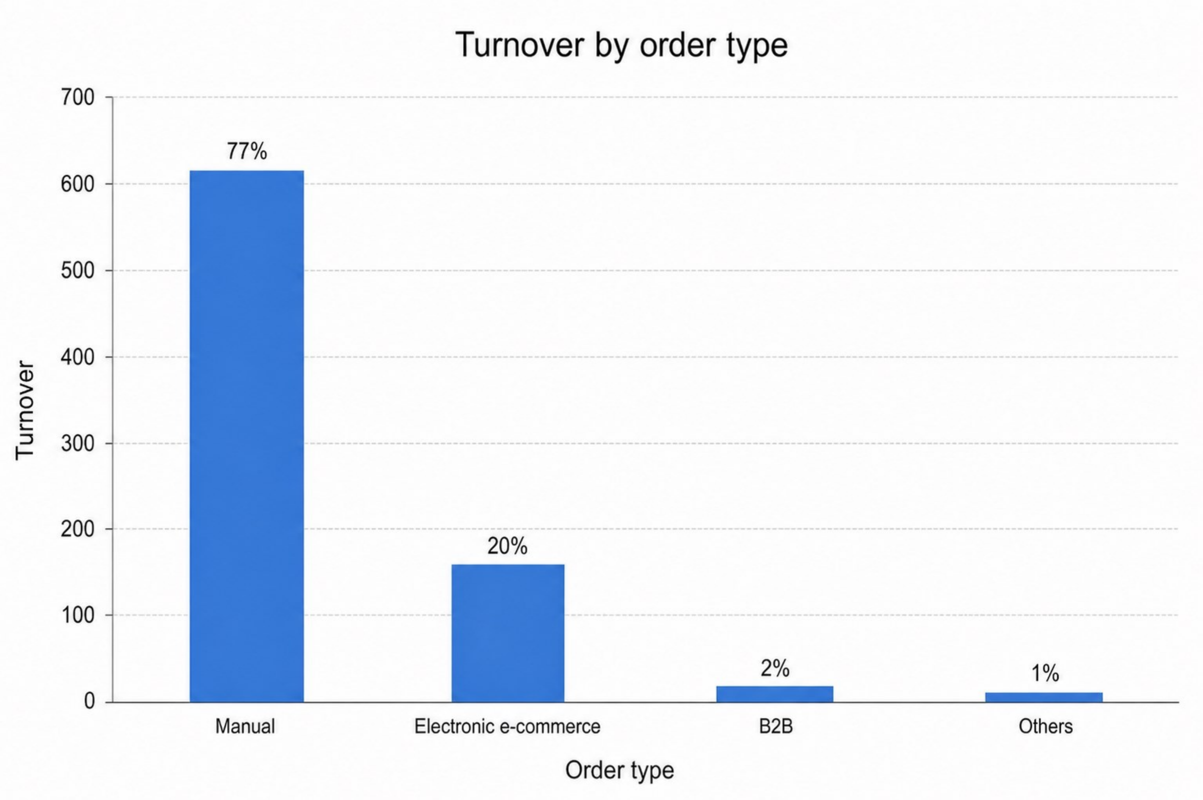


*Figure 3.1 - Turnover distribution by order type, highlighting the predominance of manual orders and the limited contribution of B2B orders.*

Total turnover generated in Spain in 2025 amounts to 808 million euros. Of this total, as shown in Figure 3.1, 14 million euros were generated through the B2B platform, corresponding to 1.7% of total turnover, while 623 million euros, corresponding to 77.2%, were generated through offline channels. The remaining turnover is distributed across other channels.

These figures show that the B2B platform still plays a limited role in Prysmian Spain's commercial activity. Although the company has invested in a dedicated digital channel, most of the revenue continues to be generated outside the platform. Digitalization is therefore not absent, but rather fragmented across different systems and ordering logics.

The B2B platform represents only one component of Prysmian's broader digital order ecosystem, and its relatively small weight reinforces the relevance of investigating why customers continue to rely predominantly on offline or alternative electronic channels.

### 3.2.2 Customer KPIs

The total number of customers in Spain in 2025 is 1,446. This figure is based on customer ID rather than customer name, since the ID provides a more precise and unique identifier. If customer name were used instead, the total number would fall to 1,302, reflecting the fact that names are less reliable for uniquely identifying customers across systems. In addition, the same customer may appear under different company codes while still referring to the same underlying entity. For this reason, customer ID is the most appropriate unit for the empirical analysis.

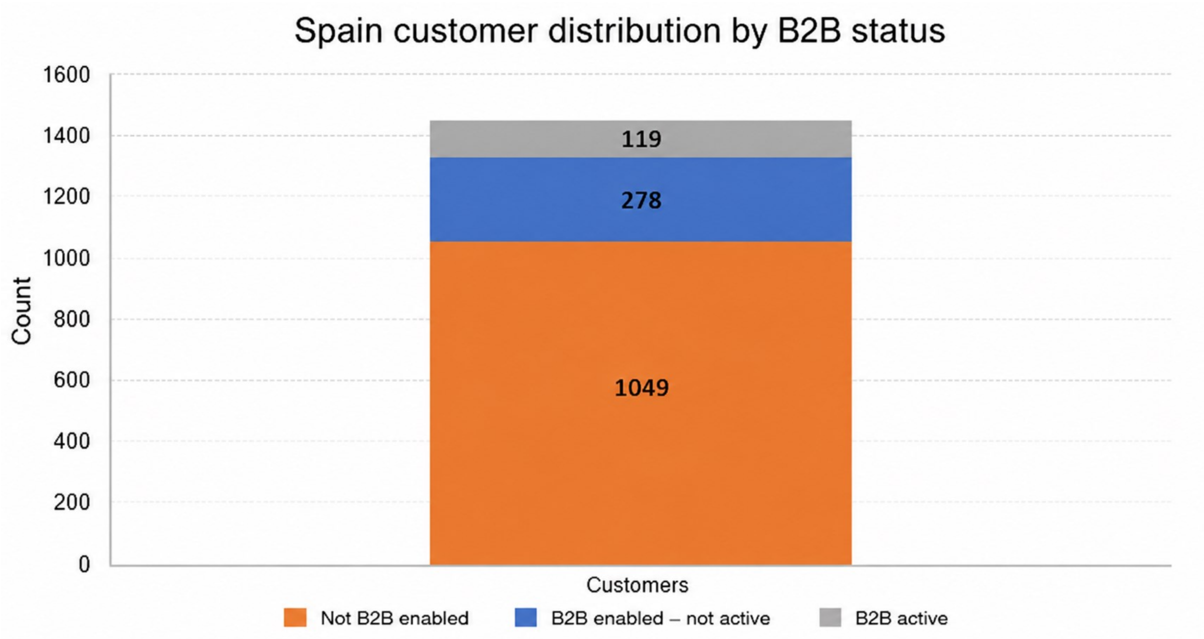


*Figure 3.2 - Customer distribution in Spain by B2B status, showing the prevalence of non-enabled customers and the small share of B2B-enabled active customers.*

Out of 1,446 Spanish customers, 1,049 were not B2B-enabled. The remaining B2B-enabled customers were reduced to 397 in the final analytical sample after excluding the seven customers using only the CEEC channel. Within this sample, 119 customers placed at least one B2B order, while 278 did not, as shown in the Figure 3.2. This is one of the clearest signs of the gap between potential and actual adoption. A relatively large set of customers has been formally enabled, but only a much smaller subset has actually used the platform.

This divergence suggests that enablement alone is not sufficient to drive migration to the online channel. Additional customer-related indicators further describe the Spanish commercial environment. In 2025, the business includes 24 companies at first level, 4 sales owners, and 148 provinces or billing cities. However, the data related to sales owners are not regularly updated and have not been properly maintained, which reduces their reliability and limits the extent to which this variable can be used in the empirical analysis.

A further issue concerns the mismatch between Salesforce and Cockpit data regarding active B2B customers. Salesforce reports 143 active B2B customers, defined as customers who placed at least one B2B order, whereas Cockpit reports only 119. The discrepancy appears to be driven by the initial filters applied to the Cockpit dataset.

Since the two systems agree on the same 119 customers, these common observations are treated as the most reliable measure of active B2B customers for the purpose of the analysis. This conservative choice favors consistency and reduces the risk of overestimating actual platform usage.

### 3.2.3 Product KPIs

In 2025, a total of 5,261 products were sold at least once in Spain. By contrast, the number of products currently available on the B2B platform is 1,103. This means that the digital catalog covers only a limited portion of the total assortment sold in the Spanish market. This is a crucial point for the thesis, because catalog completeness is one of the main operational dimensions that may affect customers' willingness to shift to the online channel.

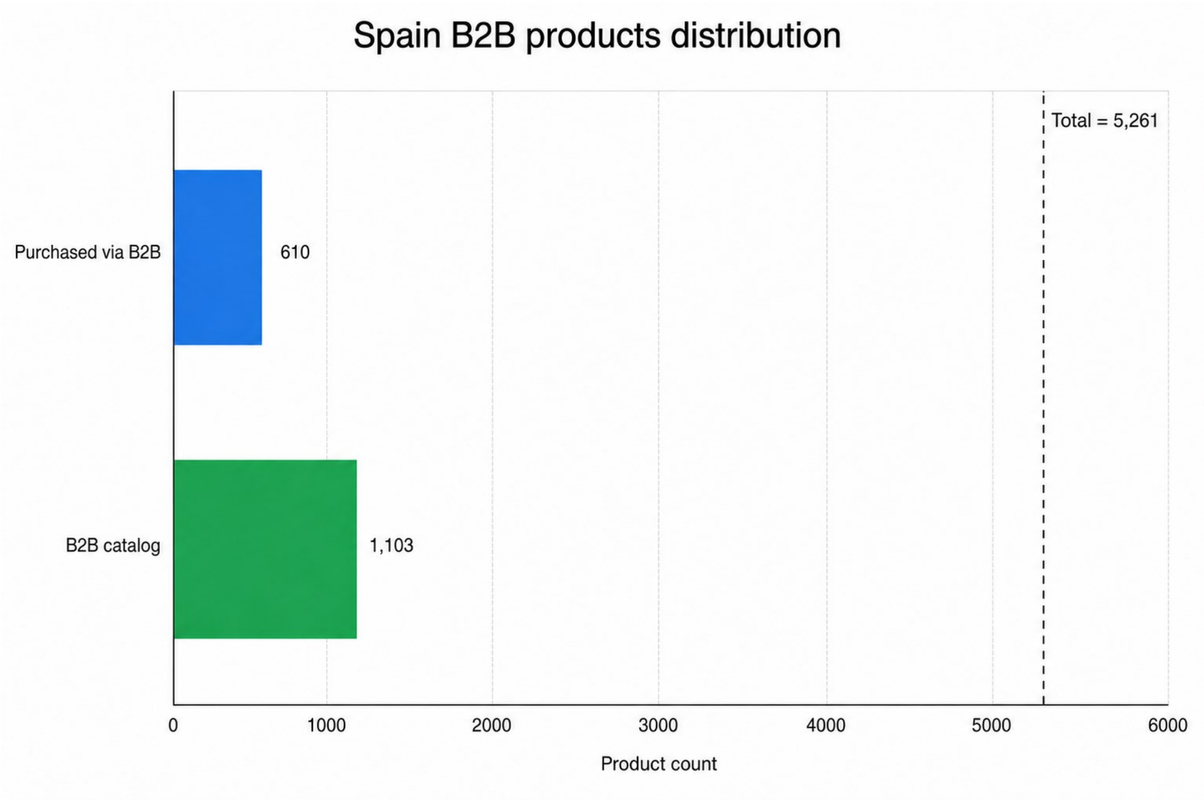


*Figure 3.3 - Spain product distribution by B2B status, showing the small share of B2B products and the smaller share of products available and purchased through the B2B channel.*

As shown in Figure 3.3, among the 1,103 products available on B2B, 610 were purchased at least once through the B2B channel during 2025. This indicates that slightly more than half of the products listed on the platform were transacted online. At the same time, it also implies that a substantial portion of the available online assortment was not purchased via B2B during the year, which may reflect either low demand for certain items or limited platform usage among enabled customers.

### 3.2.4 Order KPIs

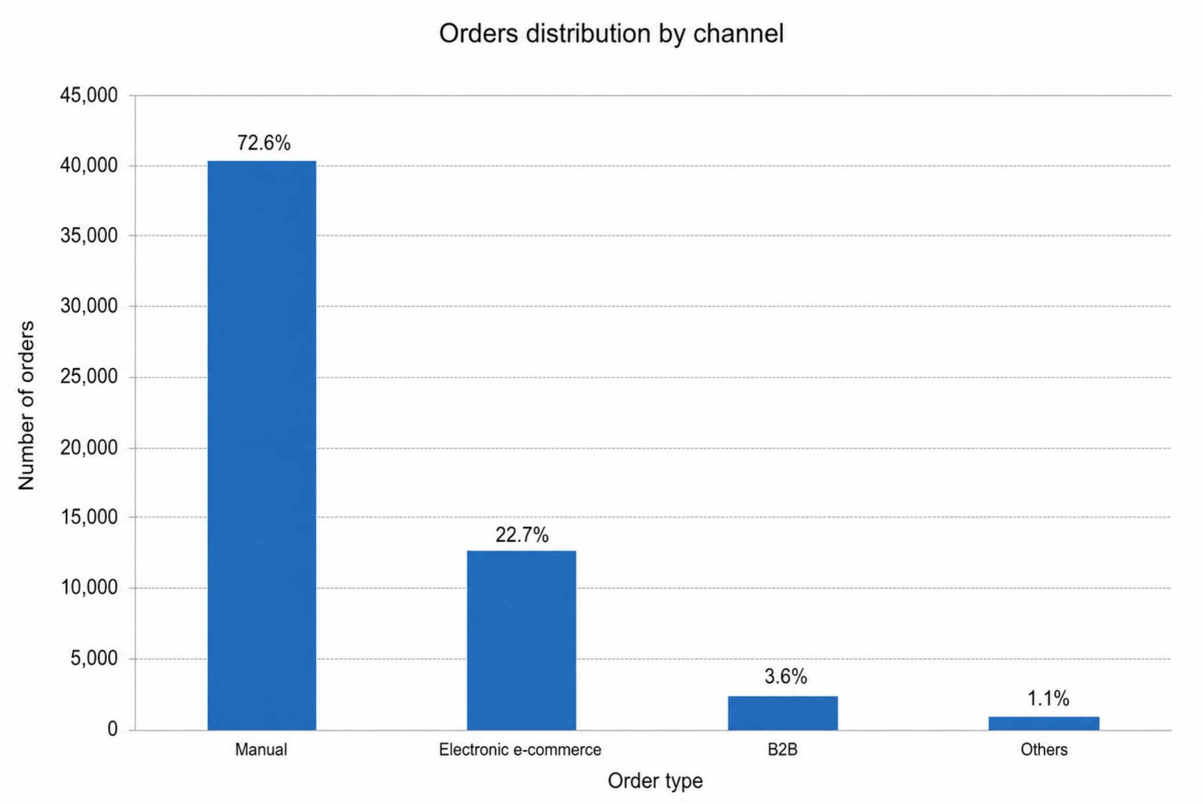


*Figure 3.4 - Orders distribution by channel, highlighting the predominance of manual orders and the small share of B2B orders.*

As shown in Figure 3.4, the total number of orders placed in Spain in 2025 is 55,905. Of these, 2,037 were placed online through the B2B platform, corresponding to 3.6% of all orders, while 40,571 orders, or 72.6%, were placed offline. The remaining share is attributable to other purchase order types.

These figures are consistent with the turnover evidence and reinforce the conclusion that the B2B platform still accounts for only a small portion of commercial activity in Spain. Interestingly, the share of orders placed via B2B is higher than the share of turnover generated via B2B. This suggests that B2B orders are, on average, smaller in value than offline orders. In turn, this may indicate that the online channel is more frequently used for lower-value purchases, whereas larger orders continue to be managed through offline interaction.

The breakdown by purchase order type confirms the predominance of offline ordering. Specifically, 72.6% of all orders are offline, 22.7% are classified as Electronic E-Commerce, and the remaining share falls under other channels.

Overall, the KPIs presented in this chapter show that Prysmian Spain has already taken important steps toward digitalization, but actual B2B platform adoption remains limited relative to the size of the business. The Spanish market therefore provides a highly relevant empirical setting in which to analyze the gap between enablement and usage, the role of catalog completeness, and the operational barriers that continue to favor offline ordering.

# Chapter 4. K-Means Customer Segmentation

This chapter introduces a customer segmentation analysis based on K-means clustering, designed to identify homogeneous groups of B2B-enabled customers in Prysmian Spain according to their purchasing behavior and their compatibility with the B2B catalog. The objective is to distinguish between different types of customers who may face different barriers and therefore require different managerial interventions. By clustering customers based on actual behavior rather than predefined categories, this chapter aims to uncover distinct patterns of B2B adoption and provide a deeper understanding of the relationship between product compatibility, digital usage, and commercial potential.

The clustering results are theoretically relevant because they show that formal enablement does not define a homogeneous adoption state. Enabled customers differ substantially in the relationship between catalog compatibility and actual digital usage, suggesting that adoption should be treated as a behavioral continuum rather than a binary condition.

The chapter is organized into two main sections.

Section 4.1 presents the methodology used to construct the customer segments, including the choice of the K-means algorithm, the variables selected for clustering, and the criteria used to determine the optimal number of clusters.

Section 4.2 presents the clustering results and provides an interpretation of each segment, linking the observed patterns to different forms of B2B adoption and to different strategic priorities.

Rather than treating all enabled customers as a single population, the analysis identifies three distinct groups: B2B Champions, Hidden Giants, and Non-B2B Fit customers. These segments represent different combinations of purchasing behavior and product compatibility and provide a practical basis for targeted commercial action.

## 4.1 Methodology: clustering B2B-enabled customers by behavior

The objective of this analysis is to segment B2B-enabled customers into homogeneous groups based on their purchasing behavior, with the aim of identifying targeted actions for each segment and increasing B2B sales. Rather than treating all enabled customers as a single group, this approach recognizes that customers may differ substantially in how they actually use the B2B platform and in the compatibility of their product mix with the current B2B assortment. Segmenting customers based on these characteristics makes it possible to distinguish between customers who are already mature digital users, customers with high unrealized potential, and customers whose limited adoption is mainly explained by structural constraints.

The analysis includes only customers who are enabled for the B2B platform. After excluding customers who place orders exclusively through the Electronic E-Commerce channel made specifically for Spain, the final sample consists of 397 B2B-enabled customers. The analysis considers only customers with complete information on these variables; any row containing a missing value was removed from the dataset to be able to apply the K-means customer segmentation; for this reason, the final sample consists of 378 customers.

The clustering is performed on a customer-level dataset in which each observation corresponds to one customer and includes customer code, customer name, turnover from B2B orders, total turnover, number of purchased products available in the B2B catalog, and total number of purchased products.

From these variables, two normalized indicators are constructed and used for clustering: the percentage of B2B turnover over total turnover and the percentage of purchased products available in the B2B catalog. The use of normalized variables is particularly important in this context. Variables expressed in absolute values, such as total turnover or total B2B turnover, tend to dominate the distance measure used by the K-means algorithm because of their larger numerical scale. As a result, customers with high turnover would disproportionately influence the clustering outcome, regardless of their actual behavioral similarity to other customers. To avoid this issue, the analysis relies on relative indicators, which allow customers to be compared on a more consistent basis and ensure that the clustering reflects behavioral patterns rather than scale effects (Jain, 2010).

K-means was selected as the clustering technique because it partitions observations into groups by assigning each observation to the nearest centroid and iteratively minimizing the within-cluster sum of squared distances. In this way, the algorithm treats customers as points in a multidimensional space and groups together those that are closest in terms of the selected variables. Since K-means requires the number of clusters to be defined in advance, different values of K, from 2 to 5, were tested. The optimal number of clusters was determined by combining two criteria: the Elbow Method, which evaluates how within-cluster variance changes as the number of clusters increases, and the Silhouette Score, which assesses the balance between cluster cohesion and cluster separation (Rousseeuw, 1987).

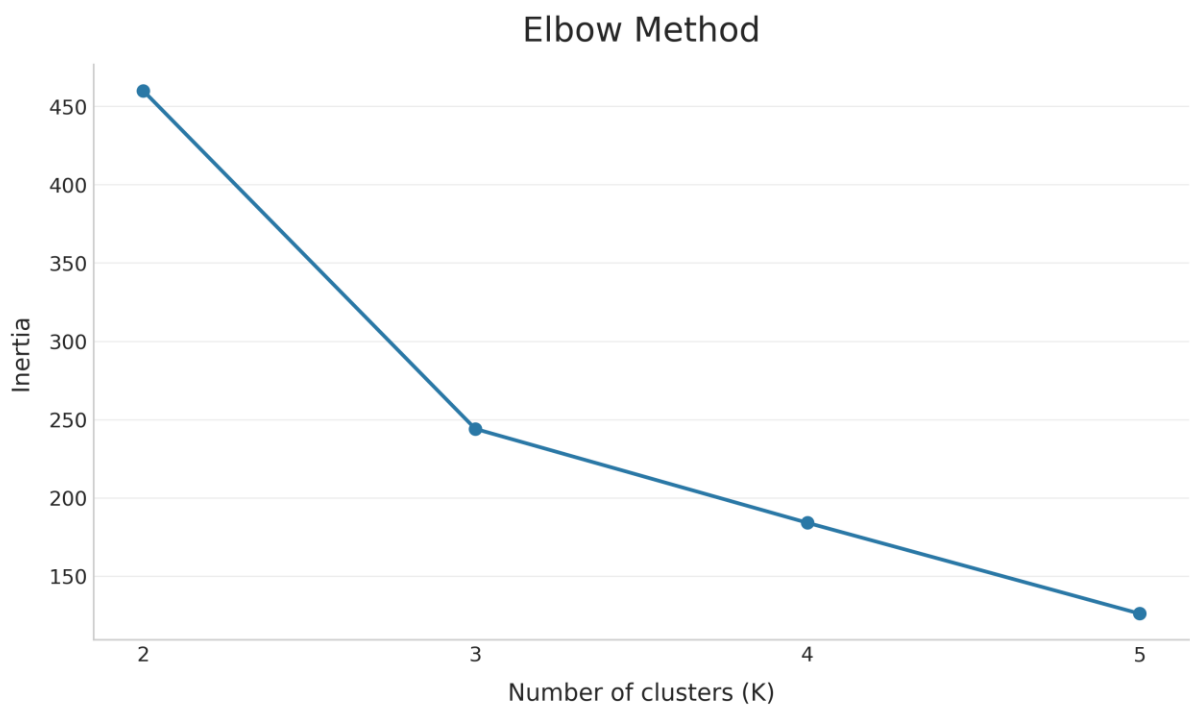


*Figure 4.1 - Elbow method for determining the optimal number of clusters in the k-means analysis.*

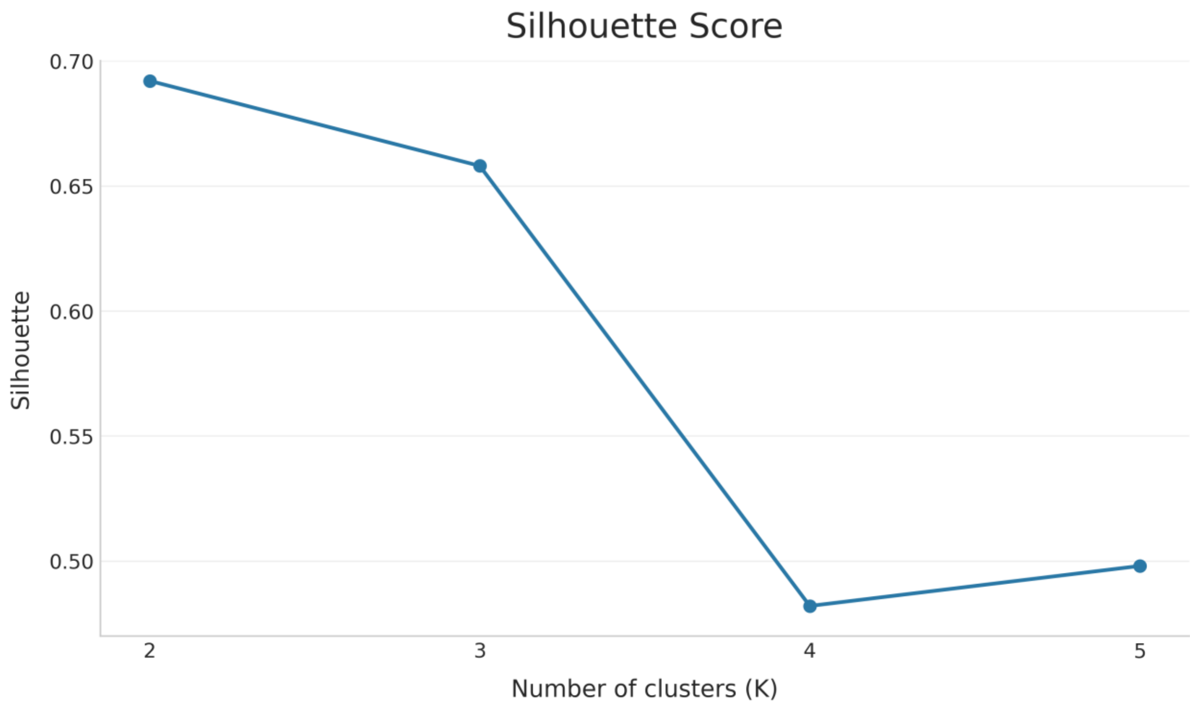


*Figure 4.2 - Silhouette score for determining the optimal number of clusters in the k-means analysis.*

Based on the combined results of these two methods, as shown in Figure 4.1 and Figure 4.2, K = 3 was selected as the optimal number of clusters. The Elbow Method (Figure 4.1) shows a substantial reduction in inertia as the number of clusters increases, followed by a visible flattening of the curve after three clusters, indicating diminishing returns from adding further groups. Although the Silhouette Score (Figure 4.2) is highest at K = 2, the value at K = 3 remains relatively close and is still clearly better than at K = 4, suggesting that K = 3 offers a reasonable balance between cluster cohesion and separation.

Overall, this choice offers a segmentation structure that is sufficiently detailed to capture meaningful customer differences, while remaining interpretable from a managerial perspective.

## 4.2 Results and cluster interpretation

The clustering analysis in Figure 4.3 identifies three distinct customer segments, each characterized by a different combination of B2B adoption and compatibility with the B2B product assortment.

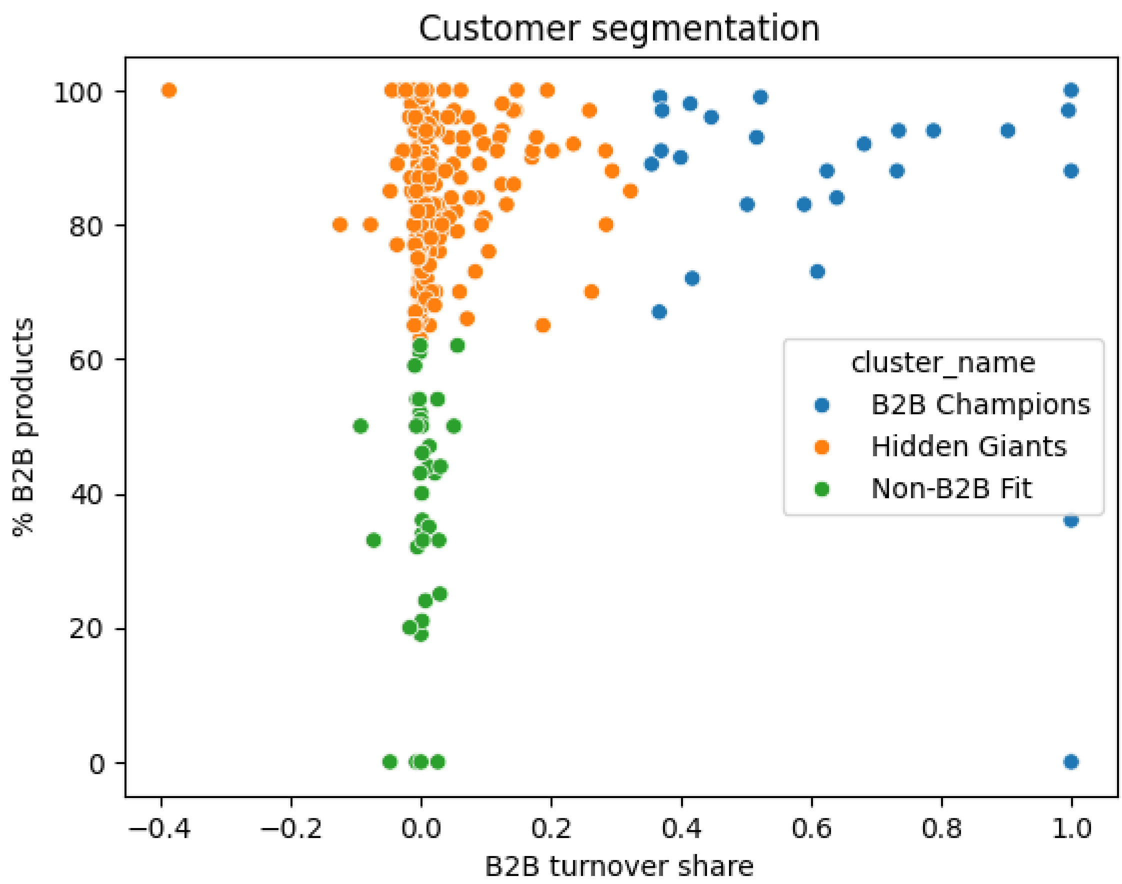


*Figure 4.3 - Customer segmentation by B2B turnover share and percentage of B2B products bought, distinguishing B2B Champions, Hidden Giants, and Non-B2B Fit customers.*

These differences are captured by the centroids of the two clustering variables: the average B2B turnover share and the average share of purchased products available in the B2B catalog. It should be noted that, in certain cases, the B2B turnover share may take negative values; this is due to adjustments such as product returns, invoice disputes, or refunds issued to customers.

Taken together, the three segments provide a structured view of the gap between potential and actual B2B usage among enabled customers (Plekhanov et al., 2023).

| Cluster name | Number of customers | Average B2B turnover share (%) | Average share of purchased products available in the B2B catalog (%) |
|---|---|---|---|
| Hidden giants | 313 | 2% | 87% |
| Non-B2B Fit | 38 | 0.4% | 38% |
| B2B Champions | 27 | 64.3% | 81% |

*Table 4.1 - Average customer characteristics by cluster, showing the number of customers, average B2B turnover share, and average share of purchased products available in the B2B catalog for each cluster.*

As shown in Table 4.1, the largest segment, labelled Hidden Giants, includes 313 customers. These customers show a very high alignment with the B2B catalog, with an average of 86.7% of purchased products already available in B2B, but a very low

B2B turnover share, around 2.0%. This means that most of the products they buy could already be ordered online, yet their actual use of the platform remains extremely limited. This segment therefore represents the main target for commercial and operational initiatives aimed at increasing B2B adoption, as the main issue does not appear to be product availability but rather the failure to convert compatibility into effective usage (Talwar et al., 2021; Yang et al., 2011).

The second segment, labelled Non-B2B Fit, includes 38 customers and is characterized by both low B2B turnover share and low compatibility with the B2B catalog. On average, only 38.5% of the products purchased by these customers are available in B2B, while their B2B turnover share remains close to zero, around 0.4%. This suggests that these customers are structurally less suited to the current configuration of the platform, since a large part of their demand concerns products that are not included in the digital catalog. For this group, the main barrier is not behavioral resistance to the platform, but limited assortment compatibility (Ghosh & Dash, 2023).

The third segment, labelled B2B Champions, includes 27 customers and represents the group with the highest degree of digital maturity. These customers combine a high B2B turnover share, around 64.3%, with a very high share of purchased products available in B2B, equal to 80.9%. Their profile indicates strong alignment with the digital channel both in terms of product compatibility and actual purchasing behavior (Ballerini et al., 2023). As such, they can be considered benchmark customers, since they represent the closest approximation to the desired adoption pattern within the current business model (Costa & Rodrigues, 2024).

Overall, the clustering provides an empirical basis for differentiating managerial actions according to the actual profile of each customer segment, rather than applying a uniform strategy to all enabled customers (Evans, 2023). For B2B Champions, the priority is to consolidate advanced usage and potentially test more sophisticated platform functionalities. For Hidden Giants, the priority is activation, training, and targeted support aimed at converting high compatibility into concrete B2B usage. For Non-B2B Fit customers, the evidence suggests that large investments in activation and onboarding are less justified under the current platform configuration, and that offline or assisted channels should continue to play a central role.

# Chapter 5. Customer-Level Analysis

This chapter presents a customer-level regression analysis aimed at identifying which customer characteristics are associated with a higher likelihood of being active on Prysmian Spain's B2B e-commerce platform, meaning that they have actually placed orders through the channel. The analysis addresses the following question: which observable customer attributes are most strongly associated with B2B activity?

This customer-level perspective is important because actual use of the platform is one of the key outcomes of digital channel adoption. Before customers can order through B2B, they must first be selected, onboarded, and enabled by the company; however, the focus of this chapter is not enablement itself, but actual channel usage. Understanding which customer profiles are more likely to be active on the B2B platform therefore helps clarify the logic behind Prysmian Spain's digital adoption process and provides an empirical basis for evaluating whether that process is sufficiently targeted, scalable, and coherent with the firm's broader digital strategy.

The chapter is structured in three main sections.

Section 5.1 presents the model specification and the variables included in the analysis.

Section 5.2 reports the estimation results and evaluates model performance.

Section 5.3 translates the findings into managerial implications, with particular attention to how Prysmian Spain can prioritize customers more selectively and reduce the gap between formal activation and effective digital adoption.

## 5.1 Model specification and variables

This section presents the customer-level regression model used to identify which customer characteristics are associated with a higher likelihood of placing orders through Prysmian Spain's B2B e-commerce platform. The purpose of the analysis is not simply descriptive. Rather, it is to estimate, in a structured and statistically interpretable way, how observable customer attributes relate to the probability of B2B activity and, therefore, to Prysmian's current digital channel expansion logic.

### 5.1.1 Variables

The target variable is *b2b_active*, a binary variable indicating whether the customer placed at least one order through the B2B platform (1 = yes, 0 = no).

The explanatory variables were selected to capture dimensions that are likely to influence B2B activity: purchasing behavior, order complexity, prior channel behavior, and organizational or geographic context. Together, these variables provide a picture of whether a customer is behaviorally compatible with the B2B model.

The first variable is *ceec_orders*, which counts the number of Electronic E-Commerce orders. This variable captures prior exposure to alternative digital ordering channels. Its inclusion is important because customers already familiar with electronic transactions may be more likely to use Prysmian's B2B platform. A positive relationship is expected: customers with prior digital channel experience should have higher odds of being active on the platform, as they demonstrate digital familiarity and lower adoption barriers. However, the effect may also be neutral or even negative if these customers are already well served through another electronic channel and therefore have less need to migrate to Prysmian's dedicated B2B platform.

A second categorical variable is *first_level_reduced*, which captures the customer's first-level organizational group after category reduction. This is one of the most relevant variables in the model because it reflects broader organizational and network-level structures. Customers belonging to the same group may share procurement practices, commercial governance, digital readiness, or strategic relationships with Prysmian. Including this variable allows the model to test whether B2B activity is shaped not only by the characteristics of individual customers, but also by the broader organizational environment in which they operate.

The model also includes *province_macro*, a grouped geographical variable that classifies customers into broader macro-areas such as North, Center, South, Islands, and Other. Geography may matter because regional markets can differ in customer composition, sales organization, logistics, and digital maturity. At the same time, using broad macro-areas instead of individual provinces avoids excessive fragmentation and helps preserve statistical stability.

Finally, the original specification included *sales_owner*, which identifies the assigned sales representative. This variable was intended to control for possible sales-driven differences, such as heterogeneous onboarding practices, commercial priorities, or different trainings for digital migration.

### 5.1.2 Variables excluded from the final specification

Several variables that were initially considered were excluded from the final model specification. This exclusion was based on methodological concerns related to multicollinearity, endogeneity, quasi-separation, and data quality issues.

The original specification included *avg_products_per_order* (average number of products per order), *log_total_orders* (logarithm of the total numbers of orders) and *log_total_products* (logarithm of the total number of products bought). However, these variables were not retained in the final specification because they did not improve the final model and could introduce redundancy with other customer-size measures.

During the model estimation procedure, several categorical groups within *first_level_reduced* exhibited quasi-separation (also called complete separation). This statistical problem occurs when a category perfectly predicts the outcome: all observations in that category have *b2b_active* = 0 or all have *b2b_active* = 1. When quasi-separation occurs, the logistic regression coefficients become extremely large, the standard errors become infinite or undefined, and the model fails to converge properly. The categories IDEE, SALTOKI, and UNASE exhibited this problem. To address it, these categories were collapsed into a residual category called "*Other_sep*". However, this collapsed category still shows unstable estimation (extremely large coefficient with infinite standard error), indicating that it remains statistically unidentifiable. For the purposes of the thesis analysis, this category is excluded from interpretation.

Additionally, a minimum-frequency threshold of 30 observations was applied to all categories in *first_level_reduced* and *province_macro*. Categories with fewer than 30 observations were collapsed into "*Other*". This threshold-based approach prevents estimation instability that arises when dummy variables are based on very small numbers of observations, which would otherwise generate unreliable coefficients with large standard errors.

The variable *sales_owner* was excluded from the final model because the underlying data is not regularly updated and appears to be unevenly distributed across customers. Many customers have missing or inconsistent *sales_owner* assignments, which limits the variable's explanatory value and introduces measurement error. While sales teams may indeed influence the target variable, the available variable is not sufficiently informative to capture those differences reliably within the statistical model.

The variable *business_desc* (customer business description) was also excluded because it has very high cardinality (many unique categories) and most categories have very few observations. Collapsing this variable would result in overly broad categories that lose meaningful heterogeneity, while keeping it would create the same sparsity and instability problems addressed by collapsing *first_level_reduced*.

The final specification therefore includes *ceec_orders, first_level_reduced,* and *province_macro.*

### 5.1.3 Grouping of categorical variables

A key methodological issue in the specification concerns the treatment of high-cardinality categorical variables. Some original variables, particularly province and first-level customer group, include many categories. If entered directly into the model without transformation, these variables would generate a high number of dummy coefficients, many of them based on a very small number of observations. This would reduce estimation stability, increase the risk of overfitting, and make interpretation unnecessarily complex.

To address this issue, the province variable was aggregated into broader macro-geographical areas: North, Center, South, Islands, and Other. This grouping preserves the main regional variation in customer location while avoiding the sparsity problems associated with a highly fragmented location variable. The purpose of this transformation is not to eliminate geographic heterogeneity, but to represent it at a level that is both statistically manageable and commercially meaningful.

A similar logic was applied to the first-level customer group. The original variable contains many categories, several of which are represented by only a few customers. To ensure stable estimation and address quasi-separation issues, a minimum-frequency threshold of 30 observations was used: only first-level groups with at least 30 observations were kept as separate categories, while all smaller groups were merged into a residual *Other* category. Additionally, categories with perfect prediction of the outcome were collapsed into *Other_sep*. This threshold-based approach balances two competing objectives: on the one hand, it preserves the most commercially relevant organizational groups, which may differ substantially in B2B usage patterns; on the other hand, it avoids generating unstable coefficients for rare groups that would otherwise be estimated on an insufficient empirical basis.

This grouping procedure is particularly important because *first_level_reduced* is expected to be one of the strongest predictors in the model. Reducing the variable in a disciplined way therefore improves both robustness and interpretability. It allows the analysis to isolate meaningful variation across major customer organizations without overfitting the model to small and potentially idiosyncratic subgroups.

### 5.1.4 Model estimation procedure

The customer-level dataset was built by grouping all transactional information so that each row represents one customer. For every customer, we calculated the variables described above (such as the number of electronic orders, the organizational group, and the geographic area) and we also recorded whether the customer is purchasing with the B2B platform or not. This gives us a clean table where each customer is one observation, and we can analyze which characteristics are linked to B2B usage. After removing observations with missing or unusable values in the variables required by the final specification, the modeling sample consists of 1,384 customers.

The sample was then split into two parts: a training set and a test set. Roughly 70% of the customers (968 customers) were used to build the model, and the remaining 30% (416 customers) were kept aside to check how well the model works on new data that it has never seen. This split was done in a way that keeps the same proportion of active and non-active customers in both parts, so that the model is trained and tested on balanced representations of the real situation.

We used a logistic regression model, which is a standard statistical method for predicting a binary outcome (in this case, active = 1 or not active = 0). The model estimates how much each variable increases or decreases the probability of being B2B active. The results are expressed in terms of coefficients and odds ratios, which are easy to interpret: a positive coefficient means the variable increases the chance of using B2B for ordering, while a negative coefficient means it decreases it.

Because the number of actual B2B users is much smaller than the number of non-users, the model should not be interpreted as a simple yes-or-no decision rule. Instead, it is more appropriately used as a ranking tool. The model assigns each customer a probability of using the B2B platform, allowing customers to be ordered from the most likely to the least likely candidates for B2B usage. This is more useful from a business perspective because it enables the commercial team to build a shortlist of promising customers and concentrate efforts where they are most likely to be effective, rather than attempting to activate all customers simultaneously.

Before finalizing the model, we also checked for statistical problems that can occur when some categories have very few customers or when a category perfectly predicts the outcome (for example, if all customers in a certain group are non-active). When this happens, the model cannot estimate reliable coefficients for those categories. To fix this, we merged small categories into larger groups and, in some cases, excluded categories that remained unstable even after merging. This ensures that the final model produces stable and interpretable results.

In summary, the estimation procedure is designed to be straightforward and reliable: we build a logistic regression model on a training sample, test it on a separate sample, and interpret it as a ranking tool that helps identify which customer profiles are most associated with B2B actual usage. The goal is not to predict B2B usage with perfect accuracy, but to provide a statistically grounded way to prioritize customers for future activation efforts.

## 5.2 Results and model performance

This section presents the results of the customer-level logistic regression and evaluates how effectively the model distinguishes between B2B users and non-users. The analysis is useful for two reasons. First, it identifies which observable customer characteristics are most strongly associated with B2B usage. Second, it clarifies whether the model can support managerial prioritization by ranking customers according to their predicted probability of using the B2B platform.

### 5.2.1 Main drivers

Table 5.1 reports the estimated coefficients and odds ratios for the customer-level logistic regression model.

| Variable | Coefficient | Standard Error | Odds Ratio | P-value |
|---|---|---|---|---|
| Intercept | -1.351 | 0.474 | 0.259 | 0.004** |
| ceec_orders | 0.001 | 0.000 | 1.001 | 0.001*** |
| province_macro_North | 0.059 | 0.445 | 1.060 | 0.895 |
| province_macro_South | -0.099 | 0.435 | 0.906 | 0.821 |
| province_macro_Other | -0.348 | 0.386 | 0.706 | 0.367 |
| province_macro_Islands | -0.701 | 0.707 | 0.496 | 0.322 |
| first_level_reduced_NOU GRUP | 0.486 | 0.555 | 1.626 | 0.381 |
| first_level_reduced_IMELCO | 0.376 | 0.422 | 1.456 | 0.374 |
| first_level_reduced_NORIA | 0.001 | 0.488 | 1.001 | 0.998 |
| first_level_reduced_GRUPO ELECTRO STOCKS | -0.943 | 0.608 | 0.389 | 0.121 |
| first_level_reduced_GRUPO SONEPAR | -2.764 | 0.838 | 0.063 | 0.001*** |
| first_level_reduced_Other | -2.221 | 0.490 | 0.109 | 0.000*** |
| first_level_reduced_Other_sep | -18.506 | 1835.891 | 0.000 | 0.992 |

**Notes**: Significance levels: $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

*Table 5.1 - Customer-level logistic regression coefficients and odds ratios, highlighting the main determinants of B2B usage.*

The results show that *ceec_orders* is the only continuous variable that remains clearly significant in the final specification, while the main source of variation in B2B usage comes from *first_level_reduced*, namely the customer's organizational group. Relative to the reference category FEGIME INTERNACIONAL, several groups display markedly different odds of being B2B active users, which confirms that B2B usage is not distributed uniformly across the customer base but is strongly shaped by organizational affiliation.

The regression does not provide statistically significant evidence that NOU GRUP, IMELCO, or NORIA differ from the reference category in their probability of being B2B active. Therefore, these groups should not be prioritized on the basis of coefficient direction alone. The only robust group-level evidence concerns the significantly lower odds of B2B usage for GRUPO SONEPAR and for the residual Other category. GRUPO SONEPAR has a coefficient of $\beta = -2.764$ with $p = 0.001$, which means that customers in this group have substantially lower odds of being B2B active than the reference category. The odds ratio is approximately 0.063, implying a very large reduction in the likelihood of being B2B active. The category Other is even more clearly significant, with $\beta = -2.221$, $p < 0.001$, and an odds ratio of about 0.109. This means that smaller or miscellaneous groups collapsed into this category are much less likely to be B2B active than customers in FEGIME INTERNACIONAL. These are the strongest and most interpretable group-level findings in the model.

The category *Other_sep* must be treated differently. It has an extremely large negative coefficient ($\beta = -18.506$), with a p-value of 0.992. This is not a meaningful effect. Instead, it is a statistical sign of remaining quasi-separation, meaning that this group is still not reliably identifiable in the model. For this reason, *Other_sep* should be excluded from interpretation in the thesis, and it should be presented only as a technical limitation of the estimation.

The variable *ceec_orders* is positive and statistically significant, with $\beta = 0.001$, $p = 0.001$, and an odds ratio of 1.001. Although the coefficient is numerically small because the variable is measured in units of orders, its sign is clear and its significance is strong. The interpretation is that customers with more electronic orders are slightly more likely to be B2B active.

The geographic variable *province_macro* does not emerge as a meaningful determinant in the final model. The coefficients for Islands, North, Other, and South are all statistically insignificant, with p-values well above 0.05. This means that, once organizational group and prior electronic ordering are considered, regional location does not appear to explain the target variable in a robust way. Geography may still matter indirectly through customer composition or commercial structure, but it is not a key driver in the model as estimated here.

Overall, the corrected results indicate that B2B usage is mainly associated with organizational affiliation and, to a lesser extent, prior electronic channel usage. The strongest and most reliable effects are the negative ones for GRUPO SONEPAR and Other, while *ceec_orders* provides a modest but statistically meaningful positive association. By contrast, geography is not significant, and the *Other_sep* category remains unstable and should not be interpreted.

### 5.2.2 Model evaluation and performance metrics

The predictive performance of the model was evaluated on the test set using standard classification metrics.

| Metric | Value |
|---|---|
| Accuracy | 0.745 |
| Precision | 0.243 |
| Recall | 0.917 |
| F1-score | 0.384 |
| ROC-AUC | 0.826 |

*Table 5.2 - Test-set classification performance for the customer-level logistic regression model, including accuracy, precision, recall, F1-score, and ROC-AUC.*

| | Predicted 0 | Predicted 1 |
|---|---|---|
| **Actual 0** | 277 | 103 |
| **Actual 1** | 3 | 33 |

*Table 5.3 - Confusion matrix for the customer-level logistic regression model.*

Table 5.2 reports accuracy, precision, recall, F1-score, and ROC-AUC, while Table 5.3 presents the corresponding confusion matrix. Taken together, these indicators show that the model performs well as a discriminative tool, but much less well as a strict binary classifier at the standard 0.5 threshold (Sujon et al., 2025).

The ROC-AUC value of 0.826 indicates that the model has good discriminative ability. This means that the model can distinguish reasonably well between active and non-active customers across different possible thresholds. Therefore, even though the model does not classify every customer correctly at the selected cutoff, it can still be considered useful for ranking customers according to their estimated probability of B2B usage.

The confusion matrix shows that the model correctly classified 277 non-active customers as non-active, while 103 non-active customers were incorrectly classified as active. For the B2B users class, the model correctly identified 33 active customers, while only 3 active customers were missed. This result is reflected in the high recall of 0.917, which shows that the model is particularly effective at capturing customers who are actually using B2B.

However, the precision of 0.243 indicates that many customers predicted as active are actually non-active. This produces a higher number of false positives, but it also allows the model to miss very few B2B active customers.

The F1-score of 0.384 reflects the trade-off between the high recall and the lower precision. Overall, these results suggest that the model is more suitable as a screening or prioritization tool than as a strict binary classifier.

These results imply that the model prioritizes capturing the majority of B2B active customers rather than minimizing false positives. The high recall of 92% for the B2B active class means that the model is effective at identifying customers who are likely to be suitable for B2B actual usage.

From a managerial point of view, this is valuable. The model's ability to capture 92% of B2B users suggests that it can serve as an effective screening tool to help the commercial team focus on the most promising candidates.

Overall, the performance metrics confirm that the model is empirically informative but should be used with the right interpretation. Its main contribution is not to predict B2B usage perfectly, but to provide a reliable statistical ranking of customer profiles. This makes the model a useful input for strategic prioritization, allowing the firm to identify a shortlist of high-potential customers while accepting a high level of false positives. These findings provide the empirical foundation for the strategic implications discussed in the following Section 5.3.

## 5.3 Implications

The customer-level regression results provide a practical basis for thinking about B2B actual usage more selectively. Rather than interpreting the model as a decision rule, it is more appropriate to treat it as a screening tool that helps identify customer profiles that are more or less aligned with the current platform configuration.

A first implication concerns the strong role of organizational group. Since first-level customer group emerges as the most important predictor of B2B actual usage, organizational affiliation appears to be closely linked to the current purchasing logic on the platform. There is no statistically significant evidence that NOU GRUP, IMELCO, or NORIA differ from the reference category, FEGIME INTERNACIONAL, in their probability of being B2B active. Therefore, these groups should not be interpreted as more compatible on the basis of coefficient direction alone.

By contrast, GRUPO SONEPAR and the residual *Other* category show significantly lower odds of B2B usage. SALTOKI and IDEE are included in *Other_sep*, which remains affected by quasi-separation and should not be interpreted separately.

Similarly, *ceec_orders* is statistically significant but has a small marginal effect per order. This means that prior use of another electronic channel is associated with B2B actual usage, but only weakly, so it should not be interpreted as a strong stand-alone indicator of purchasing activity on the platform.

The broader implication is that B2B actual usage should be interpreted in a more segmented way. Rather than assuming that all customers follow the same purchasing pattern, the results suggest that usage on the platform is more concentrated among customers whose organizational context and prior digital behavior are more compatible with the current B2B setup. At the same time, some customers with lower predicted usage may still have commercial potential, so they should not be treated as excluded categories. Instead, their lower observed activity may reflect factors that are not captured by the model, such as ordering routines, product fit, or the way purchasing responsibilities are organized internally.

This segmentation logic is also consistent with the broader evidence developed in the thesis. The descriptive analysis shows a large gap between formal enablement and actual B2B usage, while the clustering analysis identifies a large segment of “Hidden

Giants," namely enabled customers with high catalog compatibility but very limited platform use.

For this reason, the most effective strategy is likely to combine predictive prioritization with targeted onboarding actions. Customers identified by the model as high-potential candidates should not only be enabled but also supported through structured activation measures such as commercial follow-up, onboarding guidance, and clear communication of the operational advantages of the platform.

Overall, the results suggest that Prysmian Spain's current B2B purchasing pattern is shaped primarily by organizational structure and, to a lesser extent, by prior electronic channel experience. Geography does not appear to be a robust determinant. The model therefore helps describe where B2B actual usage is concentrated, but it should not be interpreted as a complete explanation of purchasing behavior.

# Chapter 6. Order-Level Analysis

This chapter presents an order-level regression analysis aimed at identifying which order characteristics are associated with a higher likelihood of being placed through the offline channel rather than via Prysmian Spain's B2B platform. While Chapter 5 focused on actual B2B usage, the present analysis addresses a different but equally important question: once a customer has access to the platform, which order characteristics make offline ordering more likely?

The order-level perspective is particularly relevant because digital adoption is ultimately expressed through individual purchasing decisions rather than through formal enablement alone. A customer may be enabled for B2B and still rely on offline ordering whenever the order involves high economic value, is affected by stock-related constraints, or includes characteristics that make offline interaction more convenient or necessary. The objective of this chapter is therefore to identify the operational barriers that continue to divert demand away from the digital channel even after access to the platform has already been granted.

The chapter is structured in three main sections.

Section 6.1 presents the model specification and the variables included in the analysis.

Section 6.2 reports the estimation results and evaluates model performance.

Section 6.3 translates the findings into managerial implications for stock reliability and high-order value support.

## 6.1 Model specification and variables

This section presents the specification of the order-level regression model used to explain the choice between offline and B2B ordering channels among customers who are already enabled for Prysmian Spain's B2B platform. The purpose of the analysis is not to study access to the digital channel, which has already been addressed at the customer level, but rather to understand why enabled customers still place some orders offline instead of through the platform. In this sense, the order-level model focuses on realized channel choice conditional on enablement and examines how order value, stock availability, and time effects are associated with that decision.

The order-level perspective is particularly important because digital adoption is ultimately reflected in individual transactions rather than in formal enablement alone. A customer may be fully enabled for B2B and still rely on offline ordering when stock conditions are unfavorable or when the order has characteristics that make offline interaction more convenient. The model therefore aims to isolate the operational and transactional factors that are associated with demand continuing to move away from the B2B platform even after digital access has already been granted.

### 6.1.1 Variables

The dependent variable is a binary indicator, *offline_order_dummy*, which takes value 1 if the order is placed offline and 0 if the order is placed through the B2B platform.

This modeling choice makes the interpretation of the coefficients more intuitive, since positive effects indicate factors associated with a higher likelihood of offline ordering, while negative effects indicate conditions associated with a greater propensity to use the B2B platform. The dependent variable is defined only for customers who are already enabled for the platform, so the analysis captures channel choice conditional on access rather than the broader decision of whether to adopt B2B at all.

Two variables related to order composition and catalog compatibility were initially considered for the analysis: *order_composition*, a categorical indicator classifying each order as fully non-B2B, mixed, or fully B2B-compatible, and *pct_b2b_line_items*, measuring the share of products in the order that were available in the B2B catalog. However, both variables were ultimately excluded from the final specification because they were too mechanically linked to the outcome. Orders composed entirely of non-B2B products were almost always executed offline by construction, and a very high or very low B2B-compatible share would have introduced a similarly tautological relationship with the dependent variable. For this reason, these variables were not retained in the final model, as their inclusion would have reduced the analytical value of the regression by embedding catalog feasibility too directly into the prediction of offline ordering.

The model includes the logarithm of order turnover (*log_order_turnover*), which captures the economic relevance of the transaction. Higher-value orders may be more likely to remain offline because they often involve additional negotiation, verification, or coordination with sales personnel. In this context, the logarithmic transformation is useful because it reduces the influence of extreme values and allows the effect of order value to be interpreted in relative rather than purely absolute terms. This variable therefore captures whether larger or more commercially important transactions are less compatible with the current standardized structure of the B2B platform.

Another important explanatory variable is *pct_in_stock_products*, defined as the percentage of products in the order that are available in stock. This variable captures the operational feasibility of completing the order smoothly through the digital channel. Orders with higher stock availability are expected to be less likely to be placed offline, because reliable availability reduces uncertainty and increases the practical usability of the platform. In this sense, stock visibility and stock reliability are treated as factors associated with channel choice rather than as downstream operational outcomes.

In addition to these order-level covariates, the model includes monthly fixed effects to control for seasonality and time-specific shocks affecting all customers. These temporal controls account for common factors such as demand fluctuations, commercial cycles, or operational disruptions that may influence the ordering channel independently of the characteristics of individual orders. Since the dataset covers only the year 2025, monthly dummies provide a sufficiently granular and stable representation of temporal heterogeneity. This choice avoids the excessive noise that could result from more detailed time controls while still capturing systematic variation across the year.

### 6.1.2 Model estimation procedure

The order-level analysis follows the same general procedure used in Chapter 5, but applies it at the order level. First, an order-level dataset was constructed using transactional data from customers already enabled for the B2B platform, so that each observation corresponds to a single order and includes a binary target variable indicating whether the order was placed offline or via B2B. The explanatory variables were then linked to each order to describe its turnover and stock conditions.

The final dataset was divided into training and test sets while preserving the original distribution of offline and B2B orders. A logistic regression model was then estimated on the training set and evaluated on the test set using standard classification metrics. However, given the likely presence of unobserved behavioral factors and the operational complexity of channel choice, the model should be interpreted primarily as an explanatory framework rather than as a purely predictive tool. Its main value lies in identifying which order characteristics are systematically associated with offline ordering among customers who already have access to the platform (Sivakumar, Parthasarathy, & Padmapriya, 2024; Vrigazova, 2021).

This regression analysis is particularly relevant because it addresses a core operational issue in the adoption of digital sales channels. If customers who are already enabled for B2B still place a substantial share of their orders offline, the issue cannot be explained by access alone. Instead, it suggests that channel migration is constrained by the interaction between transactional characteristics and operational reliability. The model therefore provides an empirical basis for assessing whether offline ordering is mainly driven by persistent customer habits or by limitations in the current configuration of the B2B offer (Evans, 2023; Ghosh & Dash, 2023).

## 6.2 Results and model performance

This section presents the estimation results of the order-level logistic regression model and evaluates its predictive performance. The objective is to identify which order characteristics are most strongly associated with offline ordering among customers who are already enabled for the B2B platform, and to assess how effectively the model distinguishes between orders placed through the digital channel and those that continue to be processed offline.

### 6.2.1 Odds ratios and main drivers

The results in the Table 6.1 provide empirical evidence on the main determinants of channel choice at the order level. Since the dependent variable is *offline_order_dummy*, the estimated odds ratios indicate how the odds of placing an order offline change with a one-unit increase in each explanatory variable, holding all other variables constant. As in the previous chapter, an odds ratio greater than 1 indicates a positive association with offline ordering, whereas an odds ratio below 1 indicates a lower likelihood of offline ordering and, therefore, a greater propensity to use the B2B platform (Sperandei, 2014; Szumilas, 2010).

| Variable | Coefficient | Standard Error | Odds Ratio | P-value |
|---|---|---|---|---|
| Intercept | 1.662 | 0.174 | 5.272 | 0.000*** |
| log_order_turnover | 0.231 | 0.026 | 1.260 | 0.000*** |
| C(month)[T.6] | 0.182 | 0.191 | 1.200 | 0.340 |
| C(month)[T.3] | -0.195 | 0.190 | 0.822 | 0.303 |
| C(month)[T.2] | -0.270 | 0.209 | 0.764 | 0.197 |
| C(month)[T.12] | -0.428 | 0.177 | 0.652 | 0.016* |
| pct_in_stock_products | -0.479 | 0.122 | 0.620 | 0.000*** |
| C(month)[T.5] | -0.557 | 0.185 | 0.573 | 0.003** |
| C(month)[T.10] | -0.564 | 0.180 | 0.569 | 0.002** |
| C(month)[T.7] | -0.609 | 0.184 | 0.544 | 0.001*** |
| C(month)[T.4] | -0.669 | 0.183 | 0.512 | 0.000*** |
| C(month)[T.11] | -0.733 | 0.182 | 0.480 | 0.000*** |
| C(month)[T.9] | -0.750 | 0.176 | 0.472 | 0.000*** |
| C(month)[T.8] | -0.791 | 0.189 | 0.454 | 0.000*** |

**Notes**: Significance levels: $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

*Table 6.1 - Order-level logistic regression coefficients and odds ratios, highlighting the main determinants of offline ordering among B2B-enabled customers.*

The results in Table 6.1 show that the economic value of the transaction plays a relevant role in shaping channel choice. The coefficient associated with *log_order_turnover* is positive and highly significant (β = 0.231, p < 0.001), with an odds ratio of 1.260. This means that a one-unit increase in the logarithm of order turnover is associated with a 26% increase in the odds of offline ordering. This finding is consistent with the idea that commercially more important transactions are often handled through more interactive and flexible channels, especially in B2B settings where negotiation, customer-specific pricing, or additional coordination may be required. Higher-value orders are therefore more likely to remain offline (Sharma, Gupta, & Joshi, 2020; Ghosh & Dash, 2023).

A second key result concerns stock availability. The coefficient associated with *pct_in_stock_products* is negative and highly significant (β = -0.479, p < 0.001), with an odds ratio of 0.620. This implies that better stock availability reduces the likelihood that the order is placed offline. In odds-ratio terms, a one-unit increase in the proportion of in-stock products (which corresponds to moving from 0% to 100% availability) is associated with a 38% reduction in the odds of offline ordering. Since *pct_in_stock_products* is expressed as a proportion (from 0 to 1), a one-unit increase corresponds to a very large change in availability; therefore, smaller and more realistic increases, such as 10 percentage points, would imply a more moderate but still meaningful reduction in the odds. This is one of the most important findings of the model, because it shows that the reliability of product availability is not only an operational condition, but also an important correlate of channel choice itself. When customers can rely on the products being available, the B2B platform becomes more suitable as a purchasing channel (Bijmolt et al., 2021; Derhami, Montreuil, & Bau, 2021).

Taken together, these results suggest that offline ordering among enabled customers is primarily associated with structural transaction characteristics rather than with enablement status alone. Higher turnover increases the tendency to remain offline, while greater stock availability supports channel migration toward B2B. This confirms that access to the platform is a necessary but not sufficient condition for actual digital adoption (Yang et al., 2011; Derhami, Montreuil, & Bau, 2021).

The monthly fixed effects also reveal a temporal pattern in ordering behavior; several month effects are statistically significant at conventional levels:

- Month 12 (December): $\beta = -0.428$, $p = 0.016$, odds ratio = 0.652
- Month 5 (May): $\beta = -0.557$, $p = 0.003$, odds ratio = 0.573
- Month 10 (October): $\beta = -0.564$, $p = 0.002$, odds ratio = 0.569
- Month 7 (July): $\beta = -0.609$, $p = 0.001$, odds ratio = 0.544
- Month 4 (April): $\beta = -0.669$, $p < 0.001$, odds ratio = 0.512
- Month 11 (November): $\beta = -0.733$, $p < 0.001$, odds ratio = 0.480
- Month 9 (September): $\beta = -0.750$, $p < 0.001$, odds ratio = 0.472
- Month 8 (August): $\beta = -0.791$, $p < 0.001$, odds ratio = 0.454

Several monthly fixed effects are statistically significant relative to January, indicating time variation in offline ordering. However, the model does not identify whether this pattern reflects seasonality, customer learning, rollout dynamics, commercial campaigns, or operational shocks. Therefore, the month coefficients should be interpreted as controls for temporal heterogeneity rather than as direct evidence of gradual adoption.

Overall, the estimated coefficients provide a coherent picture of channel choice in Prysmian Spain. Orders are more likely to remain offline when they are economically more important and when stock conditions are less favorable. Conversely, the B2B channel becomes more likely when the order is operationally feasible and supported by available stock. This interpretation is consistent with the broader theoretical framework of the thesis, which emphasizes that digital migration in B2B contexts depends on the interaction between customer behavior and operational reliability.

### 6.2.2 Model evaluation and performance metrics

The predictive performance of the order-level model is more limited than that of the customer-level model.

| Metric | Value |
|---|---|
| Accuracy | 0.783 |
| Precision | 0.783 |
| Recall | 0.999 |
| F1-score | 0.878 |
| ROC-AUC | 0.608 |

*Table 6.2 - Test-set classification performance for the order-level model, including accuracy, precision, recall, F1-score, ROC-AUC.*

| | Predicted 0 | Predicted 1 |
|---|---|---|
| **Actual 0** | 0 | 607 |
| **Actual 1** | 2 | 2196 |

*Table 6.3 - Confusion matrix for the order-level model.*

As reported in Table 6.2 , the model achieves a ROC-AUC of 0.608, which indicates only a modest improvement over random classification. This means that the model captures some underlying signal, but that the explanatory variables included in the specification are not sufficient to fully distinguish between offline and B2B orders. In practical terms, the model is therefore more useful as an explanatory framework than as a predictive instrument for classifying individual transactions.

This limited discriminative power is substantively meaningful rather than merely technical. Channel choice at the order level is likely influenced not only by measurable transactional variables, but also by a range of unobserved behavioral and organizational factors that are not captured in the dataset. These may include customer habits, urgency, relationship-specific negotiation practices, preferences for direct human interaction, internal routines, or temporary commercial circumstances. The relatively modest ROC-AUC therefore suggests that even though stock and order-value variables matter, they do not exhaust the explanation of why enabled customers continue to use offline ordering channels.

At the same time, the limited classification accuracy does not invalidate the substantive interpretation of the coefficients. Logistic regression remains valuable here because its purpose is not only prediction, but also the estimation of structured associations between order characteristics and channel outcomes. Even a model with moderate predictive power can provide useful managerial evidence when the coefficients point consistently to specific operational frictions. In this case, the model clearly indicates that stock reliability is a central factor associated with channel usage, while order value continues to favor offline processing. These findings are actionable even if the model is not sufficiently strong to forecast the channel of every individual order with high precision.

The confusion matrix reported in Table 6.3 should therefore be interpreted with this limitation in mind. Since *offline_order_dummy* = 1 represents offline orders and 0 represents B2B orders, the confusion matrix shows that the model predicts almost all observations as offline. It correctly classifies 2,196 out of 2,198 offline orders, but fails to identify B2B orders, correctly classifying 0 out of 607 B2B orders. Therefore, the high accuracy, precision, recall, and F1-score are misleading because they are driven by the dominance of the offline class. The ROC-AUC of 0.608 provides a more realistic assessment and indicates weak discriminative power (Sujon et al., 2025). The model should therefore be used for interpreting associations, not for classifying individual orders.

The main contribution of the model lies in revealing which order-level factors systematically shift the probability of offline ordering, rather than in producing highly accurate transaction-level predictions. This is especially appropriate in the context of the present thesis, whose objective is to identify operational barriers to B2B adoption and to derive strategic implications for catalog and stock management rather than to build a production-level prediction engine (Bijmolt et al., 2021; Vrigazova, 2021).

Overall, the evaluation results confirm that the order-level specification should be interpreted primarily as an explanatory model. Its discriminative power remains limited, but the estimated relationships are still informative and consistent with the descriptive evidence presented earlier in the thesis. The model therefore provides a useful empirical bridge between the observed persistence of offline ordering and the operational characteristics of the current B2B offer, highlighting the importance of stock availability and transaction value in shaping customers' actual channel behavior (Evans, 2023; Ghosh & Dash, 2023).

## 6.3 Implications for stock reliability and high-order value support

The order-level analysis provides a clear managerial implication: improving B2B adoption among already enabled customers does not depend only on activation efforts, but also on the ability of the platform to support transactions under reliable operational conditions. The results show that offline ordering is more likely when stock availability is weaker and when the economic value of the order is higher, while the probability of offline ordering decreases when the transaction can be fulfilled more smoothly through the platform (Bijmolt et al., 2021; Derhami, Montreuil, & Bau, 2021).

This finding is important because it shifts attention from customer enablement alone to the practical usability of the digital offer. A customer may already be enabled and may even be willing to use the B2B platform, but if stock conditions are uncertain or if the order requires greater coordination, the transaction is less likely to be completed entirely through the digital channel. In these cases, offline ordering remains the more convenient solution, not necessarily because of customer resistance, but because the current platform does not yet fully support the operational conditions required by the transaction (Yang et al., 2011; Ghosh & Dash, 2023).

The main implication is therefore that Prysmian Spain should treat stock reliability as a core lever for digital channel migration. The results indicate that higher in-stock availability is strongly associated with a lower likelihood of offline ordering ($\beta = -0.479$, $p < 0.001$, odds ratio = 0.620). This suggests that improving product availability, particularly for items that are frequently involved in B2B-relevant transactions, may help reduce friction and increase the practical attractiveness of the platform.

A second implication concerns order value. The positive and significant effect of *log_order_turnover* (β = 0.231, $p < 0.001$, odds ratio = 1.260) indicates that higher-value orders are more likely to remain offline. This suggests that Prysmian should consider whether the current B2B platform adequately supports large transactions. For high-value orders, the company may need to offer additional features such as customer-specific pricing, dedicated sales support, or enhanced verification processes to make the digital channel more competitive with offline ordering.

The temporal pattern revealed by the month effects indicates time variation in offline ordering. However, these coefficients should be interpreted only as controls for temporal heterogeneity; the model does not identify whether the pattern reflects learning, seasonality, rollout dynamics, commercial campaigns, or operational shocks.

Overall, the order-level results suggest that Prysmian Spain should focus on improving stock availability and assessing whether the B2B platform adequately supports high-value transactions. By addressing these operational barriers, the company can increase the likelihood that enabled customers will choose the digital channel for a broader range of orders, rather than continuing to rely on offline ordering for transactions that are operationally feasible or commercially important.

# Chapter 7. Product-Level Analysis

This chapter examines the product-level drivers of B2B catalog inclusion in Prysmian Spain. While the previous chapters focused on customers and orders, this section shifts the unit of analysis to the product itself to understand which items are more likely to be enabled for the B2B platform and which product characteristics make digital self-service more or less feasible.

The product-level perspective is important because channel migration in B2B markets does not depend only on customer readiness or order composition. It also depends on whether the products being sold are sufficiently standardized, operationally manageable, and commercially suitable for online ordering. In this sense, product enablement is not merely a technical catalog decision, but a strategic choice that reflects how closely a product fits the logic of digital commerce and the needs of digitally active customers.

This chapter therefore investigates the relationship between product characteristics and the probability of B2B inclusion, using a logistic regression framework to identify the main determinants of catalog access. The analysis focuses on how product standardization, market relevance, and operational compatibility shape the likelihood that a product will be made available through the B2B platform. This makes it possible to distinguish between products that are naturally aligned with digital channel migration and products that remain better suited to offline commercial processes.

The chapter is structured in three main sections.

Section 7.1 presents the model specification and the variables used in the product-level regression.

Section 7.2 discusses the estimation results and model performance, highlighting the main drivers of B2B catalog inclusion and the predictive quality of the specification.

Section 7.3 translates these findings into managerial implications for product enablement strategy, with particular attention to how Prysmian Spain can prioritize the right products for digital expansion

## 7.1 Model specification and variables

This section presents the model specification adopted for the product-level analysis and describes the variables used to explain whether a product is included in Prysmian Spain's B2B catalog. The objective is to move from a descriptive view of catalog composition to a more structured understanding of the characteristics that make some products more likely than others to be enabled for digital self-service. In this context, B2B inclusion is interpreted not simply as a technical listing decision, but as the outcome of a broader commercial and operational logic that reflects product suitability for online ordering.

The explanatory variables capture three main dimensions. First, commercial relevance is measured through indicators of market demand and transactional frequency, since products with broader and more recurring demand are more likely to justify digital inclusion. Second, the model includes variables linked to digitally active customers, because the B2B platform is designed to support products that fit the purchasing patterns of customers already close to online ordering. This approach is consistent with the idea that platform development should follow actual customer behavior rather than internal catalog logic. Third, product class is included to account for differences in structure and suitability: standardized products are easier to display and order online, while customized or technically complex products are typically less compatible with self-service ordering.

Product-class categories are converted into dummy variables so that each major family can be compared with a reference group (*HV_Energy_Infrastructure*). This makes it possible to test whether belonging to a specific class significantly changes the probability of B2B inclusion after controlling for the other variables. Overall, the specification is designed not only to explain past enablement decisions, but also to identify which products are most suitable for future B2B expansion in a way that is both statistically rigorous and managerially useful.

### 7.1.1 Variables

The dependent variable is a binary indicator of B2B inclusion, *b2b_enabled_product*. It takes value 1 if the product is currently available in the B2B catalog and 0 if it is not. This variable captures the product-level enablement decision that the chapter seeks to explain.

Among the independent variables, *percentage_orders* measures the share of a product's orders placed by B2B-enabled customers, calculated as *orders_from_b2b_customers / total_orders*. This variable captures the extent to which a product is commercially relevant for customers who are already connected, directly or indirectly, to the digital channel. Products with a higher exposure to B2B-enabled customers are expected to be more likely to enter the B2B catalog, because enabling them can support adoption among customers who are already structurally closer to online ordering.

The model also includes *product_turnover*, *total_orders*, *orders_from_b2b_customers*, and *average_orders_per_customer*. These variables capture different dimensions of product demand, including economic value, overall order intensity, demand coming specifically from B2B customers, and the average recurrence of orders per customer. Their inclusion helps assess whether B2B inclusion is driven only by relevance to digital customers or also by the volume, concentration, and monetization of demand at the product level.

Another key explanatory component is *product_class_grouped*. This categorical structure acts as a proxy for standardization, complexity, and digital suitability. It is particularly relevant because the current B2B catalog is not neutral across product families: it is already skewed toward more standardized products that can be ordered autonomously without heavy commercial intermediation. Including these categories therefore makes it possible to test whether the existing assortment logic is statistically reflected in the product-level data.

Taken together, these variables represent three complementary dimensions of product suitability for B2B inclusion: demand relevance, digital-customer relevance, and structural product compatibility.

### 7.1.2 Model estimation procedure

The estimation procedure follows the same general logic adopted in the previous regression chapters (Chapter 5 and Chapter 6) but applies it at the product level. First, a product-level dataset is constructed in which each observation corresponds to a single product sold in Spain during the period of analysis. For each product, the relevant explanatory variables are calculated by aggregating transactional and customer information, making it possible to connect product characteristics with B2B catalog status.

The final dataset is then divided into training and test sets in order to evaluate the model on observations not used for estimation. A logistic regression model is estimated on the training sample and assessed on the test sample using standard classification metrics. This procedure allows the analysis to move beyond coefficient interpretation alone and to examine whether the model is also effective at distinguishing between products that are included in the B2B catalog and products that are not.

From a methodological perspective, logistic regression is appropriate because the dependent variable is binary and because the model produces interpretable odds ratios that are useful for managerial discussion. At the same time, the model should not be understood as a fully automatic catalog decision system. Rather, it should be interpreted as an evidence-based screening tool that identifies which products have the strongest statistical profile for B2B enablement and which products appear structurally less aligned with digital self-service.

This distinction is important because B2B inclusion ultimately depends not only on measurable statistical patterns, but also on practical considerations such as technical feasibility, assortment coherence, stock availability, and strategic commercial priorities. Even so, the regression provides a rigorous empirical framework for narrowing the set of candidate products and for making product enablement decisions more transparent and data driven. In this sense, the model creates the analytical bridge between product characteristics and the managerial logic of catalog expansion that will be discussed in the following sections.

## 7.2 Estimation results and model performance

This section presents the results of the product-level logistic regression and evaluates the model’s predictive performance (Sperandei, 2014; Szumilas, 2010). The objective is twofold. First, the analysis identifies which product characteristics are most strongly associated with B2B catalog inclusion. Second, it assesses how effectively the model distinguishes between products that are enabled in the B2B catalog and those that are not.

### 7.2.1 Odds ratios and main drivers

The regression results in Table 7.1 confirm that product class is the strongest determinant of whether a product is included in the B2B catalog.

| Variable | Coefficient | Standard Error | Odds Ratio | P-value |
|---|---|---|---|---|
| product_class_grouped_Standard_Installer | 2.907 | 0.262 | 18.308 | 0.000*** |
| percentage_orders | 0.766 | 0.142 | 2.151 | 0.000*** |
| product_turnover | 0.003 | 0.000 | 1.003 | 0.000*** |
| total_orders | 0.003 | 0.002 | 1.003 | 0.236 |
| orders_from_b2b_customers | 0.001 | 0.005 | 1.001 | 0.868 |
| average_orders_per_customer | -0.030 | 0.050 | 0.970 | 0.547 |
| product_class_grouped_Project_Customized | -1.954 | 0.459 | 0.142 | 0.000*** |
| const | -3.590 | 0.265 | 0.028 | 0.000*** |
| product_class_grouped_Other | -32.759 | 48546.365 | 0.000 | 0.999 |
| product_class_grouped_LV_MV_Cables | -44.846 | 1304.563 | 0.000 | 0.973 |

**Notes**: Significance levels: $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

*Table 7.1 - Product-level logistic regression coefficients and odds ratios, highlighting the main determinants of B2B catalog inclusion.*

Products belonging to the *Standard_Installer* group are by far the most likely to be enabled in B2B, with odds about 18.3 times higher than those of the reference category *HV_Energy_Infrastructure*. This very large effect indicates that these standardized products are structurally the most compatible with digital self-service channels such as B2B, where customers can purchase autonomously without the need for negotiation or customization.

By contrast, *Project_Customized* products show a robust negative association with B2B enablement, whereas the coefficients for *Other* products and *LV_MV_Cables* are affected by estimation or separation issues, so they do not support reliable conclusions. Overall, the only clearly negative and robust effect among the product classes is observed for *Project_Customized* products.

Beyond product class, demand from digitally enabled customers also emerges as a major driver of product enablement. *Percentage_orders* has a strong positive effect, with an odds ratio of about 2.15. This means a 10 percentage-point increase in *percentage_orders* corresponds to an odds ratio of 1.08, so the odds of B2B inclusion increase by about 8 percent. This is an important result because it shows that products that are already demanded by B2B-active customers are more likely to be perceived as candidates for digital onboarding.

The number of orders from B2B customers (*orders_from_b2b_customers*) has a negligible effect, with an odds ratio of about 1.00, and it is not statistically significant. Since the value is essentially equal to 1, the relationship can be considered practically null, suggesting that the sheer number of B2B customer orders does not affect the probability of product enablement.

Similarly, the remaining demand-side variables show mixed effects. *Total_orders* has a positive coefficient, but its magnitude is limited and not statistically significant. Although *product_turnover* is statistically significant, its substantive magnitude is small at the unit level and depends on the turnover scale. Therefore, it should be interpreted as a secondary demand-related association, not as the main driver of product enablement.

The variable *average_orders_per_customer* also shows a non-significant effect, with an odds ratio of about 0.97, as well as *orders_from_b2b_customers*.

Overall, these findings show that product enablement is primarily associated with standardization and digital relevance, while absolute sales volume, turnover, and concentrated repeat demand play a clearly secondary role.

### 7.2.2 Model evaluation and performance metrics

From a predictive perspective, the product-level model performs strongly and confirms its usefulness as a screening and prioritization tool.

| Metric | Value |
|---|---|
| Accuracy | 0.897 |
| Precision | 0.717 |
| Recall | 0.702 |
| F1-score | 0.710 |
| ROC-AUC | 0.956 |

*Table 7.2 - Test-set classification performance for the product-level logistic regression model, including accuracy, precision, recall, F1-score, and ROC-AUC.*

| | Predicted 0 | Predicted 1 |
|---|---|---|
| **Actual 0** | 1218 | 78 |
| **Actual 1** | 84 | 198 |

*Table 7.3 - Confusion matrix for the product-level logistic regression model.*

The overall accuracy is 89.7%, shown in Table 7.2 and Table 7.3, which indicates that the model correctly classifies a large majority of products. This is a solid result, although it also reflects some non-negligible misclassification, especially among products that are predicted as B2B-enabled but are not actually included in the catalog (Sujon et al., 2025).

The most important metric is the ROC-AUC, which reaches 0.956. A value at this level indicates excellent discriminatory power, meaning that the model is highly effective at distinguishing between products that are more likely and less likely to be enabled for B2B, confirming that the model is particularly suitable as a ranking tool.

A defining feature of this specification is its high recall for the B2B-enabled class, which is 0.702. The model correctly identifies 70.2% of products that are enabled in the B2B catalog, corresponding to 198 out of 282 B2B-enabled products. This means that the model is effective at capturing a substantial share of relevant products, although it does miss a meaningful minority of them.

This recall level comes with a precision of 0.717 for the B2B class. This means that around 71.7% of the products predicted as B2B-enabled are actually enabled, while the remaining share are false positives. This pattern indicates that the model maintains a more balanced trade-off between identifying candidate products and limiting erroneous positive predictions than in the earlier specification.

The classification results show that the model is useful as a first-stage screening tool, but final product-enablement decisions should still be validated by the business team. For the B2B-enabled class, the model combines good recall with moderate precision. For the non-B2B class, the pattern is similar but slightly stronger: precision is 0.94 and recall is 0.94, showing that the model is very reliable when identifying products that are not enabled and also captures most of them correctly. The F1-score for the B2B class is 0.71, reflecting the balance between precision and recall and confirming that the model performs well across both classes.

Taken together, these performance metrics suggest that the model is especially useful for identifying and prioritizing candidate products for B2B enablement, rather than for making definitive operational decisions on its own. The very high ROC-AUC and solid class-level performance indicate that the model successfully captures the broad profile of products that belong in the B2B catalog. At the same time, the moderate precision and recall mean that some of the products identified as suitable by the model will still require additional business validation before inclusion.

This interpretation is important from a managerial perspective. The model should be seen as a first-stage screening mechanism that helps narrow down the universe of potential products by highlighting those with the strongest profile for B2B enablement. Final enablement decisions should then be complemented with operational checks by the Spain Business Team, such as technical feasibility, assortment coherence, stock availability, and commercial compatibility. Used in this way, the model provides a powerful and efficient support tool for catalog expansion, while avoiding the risk of treating statistical predictions as fully automatic decisions.

To summarize, the results of the product-level regression point to a clear strategic conclusion: products are far more likely to be enabled for B2B when they are standardized, broadly relevant, and already demanded by digitally active customers. Conversely, products that are technically complex, customized, or concentrated in specific customer relationships are structurally less compatible with digital self-service channels. The strong discriminatory power of the model confirms that these patterns are robust and operationally meaningful, even if the presence of false positives and missed positives means that business validation remains necessary.

These findings provide the empirical basis for the next Section 7.3, which translates the statistical evidence into concrete implications for product enablement strategy.

## 7.3 Implications for product enablement strategy

The product-level regression analysis shows that B2B product inclusion is primarily driven by product standardization and relevance for digitally active customers, while turnover is statistically significant but has a small per-unit effect. Total orders and other demand-intensity variables are not significant. Overall, this suggests that the B2B catalog is shaped more by compatibility with autonomous digital ordering than by purely commercial performance, providing a clear strategic logic for future enablement decisions.

Prioritize standardized products, avoid complex ones. *Standard_Installer* products have odds 18.3 times higher than the reference category *HV_Energy_Infrastructure* for B2B inclusion, indicating they are structurally most suitable for digital self-service. The Spanish team should analyze the current catalog to identify *Standard_Installer* products not yet enabled and use *Standard_Installer* as the benchmark for evaluating new products.

*Project_Customized* products show a statistically significant negative association with B2B catalog inclusion, suggesting that customized or project-based products are less compatible with the current self-service logic of the platform. By contrast, the coefficients for Other and *LV_MV_Cables* should not be interpreted: their very large negative estimates, extremely large standard errors, and non-significant p-values indicate estimation instability, likely due to separation or sparse categories. Therefore, the model provides robust negative evidence only for *Project_Customized* products. For this product category, the company could explore introducing an AI agent to support customization requests, enabling customers to configure tailored solutions, even though these products are not the primary focus for catalog expansion (Gao, Opute, Jawad, & Zhan, 2025). At the same time, for products where customization remains complex or highly specific, the company should continue to rely on alternative channels such as project-specific sales processes or specialized offline ordering (Bijmolt et al., 2021; Rosales et al., 2018). Resource allocation should therefore remain focused on products with high B2B customer relevance, where digital adoption is strongest and the return on enablement investments is highest (Ballerini et al., 2023).

The percentage of orders from B2B-enabled customers has an odds ratio of 2.15, meaning higher digital relevance roughly doubles enablement odds. Products should be prioritized by calculating B2B customer share for each product and focusing on the ones where B2B customers represent substantial demand, prioritizing products purchased by many distinct customers, and avoiding products tied to single customers or small sets (Ballerini et al., 2023).

The number of B2B customer orders and the average orders per customer variable have a negligible effect, meaning that these variables are not related to the probability of B2B enablement.

The model has excellent discriminatory power (ROC-AUC = 0.956) and good recall (0.702), making it useful for screening candidate products. However, moderate precision (0.717) means that some predicted B2B products will still be false positives. Therefore, a second-stage validation check should be implemented to assess technical feasibility, stock availability, and commercial compatibility (Derhami, Montreuil, & Bau, 2021).

By implementing these strategies, Prysmian Spain can align the B2B catalog with customer demand, close offline ordering gaps, and focus resources on products with the strongest statistical profile for B2B suitability.

# Chapter 8. Stock-Level Analysis

This chapter examines which product characteristics are associated with stock availability in Prysmian Spain and how these patterns help interpret stock-related barriers to B2B adoption. While the previous chapters focused on customers, orders, and product enablement, this section shifts the attention to whether products with higher turnover and stronger B2B relevance tend to be more consistently available in stock.

The analysis is important because stock availability is not only a logistical issue, but also a commercial one. In B2B environments, customers are less likely to place orders through the digital channel if products are frequently unavailable, since stockouts interrupt the purchasing process and reduce the reliability of the platform. For this reason, stock status can be seen as one of the key operational conditions behind digital adoption and sales performance.

To assess this relationship, the chapter uses a regression model where the dependent variable is the percentage in stock, and the main objective is to identify which product characteristics are associated with better availability. The results help clarify whether the products that matter most commercially are also the ones that are more reliably available, and whether stock allocation is aligned with turnover potential.

The chapter is structured in five main sections.

Section 8.1 presents the model specification and the variables used in the analysis.

Section 8.2 discusses the estimation results and model performance.

Section 8.3 examines the relationship between stock issues and B2B ordering behavior.

Section 8.4 translates the findings into implications for stock allocation.

Section 8.5 discusses the limitations of the analysis.

A key idea underlying this chapter is that stock availability should not be interpreted only as an operational outcome. It is also a strategic lever that can support both turnover and digital channel usage, especially when high-value or high-relevance products are consistently available to customers.

## 8.1 Model specification and variables

This section presents the empirical model used to examine the relationship between stock availability and product turnover. The objective is to understand whether some product characteristics are systematically associated with a higher percentage in stock, and therefore with better operational availability. In this context, stock availability is treated not only as a logistical outcome, but also as a relevant commercial factor, since products that are more consistently available are more likely to support turnover and reduce friction in the purchasing process. The analysis is conducted at the product level. Each observation corresponds to a product, and the model estimates how differences in product characteristics are associated with differences in stock availability. More specifically, the dependent variable is the percentage in stock, which measures how often a product is available over the period considered. This specification makes it possible to evaluate whether the products with higher demand, stronger B2B relevance, or different category characteristics also tend to show more stable availability.

### 8.1.1 Variables

The dependent variable is the percentage of in-stock time (*pct_in_stock_product*), which measures the extent to which a product is available during the observation period. A higher value indicates that the product was in stock more frequently, while a lower value indicates worse availability and a greater out-of-stock share. This variable is the main outcome of interest because it captures inventory reliability and the operational frictions that could be associated with customers' ability to purchase products through the B2B channel.

The key independent variables are *log_product_turnover*, *log_total_orders*, *log_orders_from_b2b_customers*, *average_orders_per_customer*, and *product_class_grouped*.

The variable *log_product_turnover* is the logarithm of product turnover. Product turnover captures the economic importance of each product, reflecting the total sales revenue associated with it. A logarithmic transformation is applied because turnover is typically right skewed, with a small number of products generating very high values. Using the logarithmic form reduces the influence of extreme observations and makes the variable more suitable for regression analysis. In this context, a positive coefficient suggests that products with higher commercial value are more likely to be in stock, whereas a negative coefficient would indicate greater exposure to stockout risk.

The variable *log_total_orders* measures the logarithm of the total number of orders for each product. It serves as a proxy for overall demand intensity. The logarithmic transformation again helps reduce skewness and makes the variable easier to interpret. A negative coefficient would be expected, as products with a higher total number of orders are likely to be associated with lower in-stock availability due to greater exposure to stockouts.

The variable *log_orders_from_b2b_customers* measures the logarithm of the number of orders placed by B2B customers. It captures the extent to which a product is relevant for digitally active customers. Including this variable makes it possible to assess whether products that are more important for B2B usage are also more likely to face stock issues.

The variable *average_orders_per_customer* captures customer concentration patterns by measuring the average number of orders per customer for each product. Products that are purchased repeatedly by the same customers may follow different stock dynamics than products ordered by a broader customer base. This variable helps control for structural differences in demand concentration.

The categorical variable *product_class_grouped* identifies the product class or family and is transformed into dummy variables through one-hot encoding. This allows the model to control for structural differences across product categories, since different classes may have different replenishment patterns, demand volatility, or stockout risk.

The one-hot encoding transformation creates a binary variable for each product class except the reference category *HV_Energy_Infrastructure*. This allows the model to estimate separate intercept shifts for each product class while keeping a single coefficient for the other explanatory variables. In this way, the regression captures heterogeneity across product categories without requiring separate models for each one.

### 8.1.2 Model estimation procedure

To estimate these relationships, the chapter adopts an ordinary least squares regression at the product level. This specification is appropriate because the dependent variable, percentage in stock, is continuous and can therefore be modeled as a function of observable product characteristics. The coefficients are interpreted as the marginal association between each explanatory variable and product availability, holding the other variables constant.

First, variables with skewed distributions were transformed using logarithms. This applies especially to turnover, total orders, and orders from B2B customers. The transformation reduces the influence of extreme values and improves the linearity of relationships in the regression model.

Second, observations with missing values were removed from the analysis to maintain consistency across the sample. This ensures that the model is estimated only on complete observations, although it may reduce the number of products included in the final dataset.

Third, the *product_class_grouped* variable was converted into dummy variables so that the model could account for differences across product categories. This one-hot encoding approach allows the regression to capture structural heterogeneity while maintaining a coherent overall specification.

The model is estimated using Ordinary Least Squares regression, which is a standard method for analyzing linear relationships between continuous variables. OLS minimizes the sum of squared residuals and provides coefficients that are straightforward to interpret under the assumption of a linear relationship between predictors and the dependent variable (Hansen, 2022; Safaria & Nugroho, 2024).

This regression addresses an important operational issue in inventory management: why some products are more likely to experience stockouts than others. Understanding these patterns helps Prysmian identify which products are most vulnerable to availability problems and where stock management should be strengthened.

## 8.2 Estimation results and model performance

This section presents the results of the stock-level regression and evaluates the overall performance of the model. The aim is to understand which product characteristics are associated with a higher percentage in stock and to assess how well the specification explains variation in product availability across the portfolio.

### 8.2.1 Coefficients and main drivers

The regression results show that several variables are associated with the percentage of in-stock time.

| Variable | Coefficient | Standard Error | t-statistic | P-value |
|---|---|---|---|---|
| Intercept | 0.872 | 0.108 | 8.080 | 0.000*** |
| log_product_turnover | 0.131 | 0.011 | 11.743 | 0.000*** |
| log_orders_from_b2b_customers | 0.155 | 0.046 | 3.351 | 0.001*** |
| average_orders_per_customer | -0.068 | 0.022 | -3.100 | 0.002** |
| log_total_orders | -0.301 | 0.049 | -6.141 | 0.000*** |
| product_class_grouped_Project_Customized | 0.164 | 0.180 | 0.910 | 0.363 |
| product_class_grouped_Standard_Installer | 0.013 | 0.082 | 0.161 | 0.872 |

**Notes**: Significance levels: $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

*Table 8.1 – Stock-level OLS regression coefficients showing the main determinants of in-stock rate.*

The Table 8.1 shows that *log_orders_from_b2b_customers* and *log_product_turnover* are the statistically significant positive variables associated with higher in-stock availability. The coefficient for *Project_Customized* is positive but not statistically significant and should not be interpreted as evidence of a product-class effect.

The variable *log_orders_from_b2b_customers* also has a positive coefficient of 0.155 and is statistically significant (p = 0.00088). This means that a one-unit increase in the logarithm of B2B orders is associated with an increase of about 15.5 percentage points in in-stock availability. Since the variable is log-transformed, a one-unit increase corresponds approximately to a 2.7 times increase in the number of B2B orders. This suggests that products with substantially higher B2B demand tend to have higher availability. One possible explanation is that these products are strategically important and therefore receive closer stock management. Another explanation is that products relevant for B2B customers may be replenished more systematically because their demand is more visible.

The coefficient for *log_product_turnover* is also positive, at 0.131, and is the most statistically significant variable in the model ($p = 1.46\times10^{-27}$). This means that a one-unit increase in the logarithm of product turnover is associated with about 13.1 percentage points higher in-stock availability. A possible explanation is that high-value or high-importance products are given greater operational attention, which helps reduce stockout risk.

The category *product_class_grouped_Standard_Installer* has a small positive coefficient of 0.0132, which suggests only a limited association with in-stock availability relative to the reference category *HV_Energy_Infrastructure*. This coefficient is not statistically significant (p = 0.872).

The variable *average_orders_per_customer* has a negative coefficient of -0.068 and is statistically significant (p = 0.0021). This means that products bought more frequently by the same customers are associated with about 6.8 percentage points lower in-stock availability. This indicates that products with more concentrated repeat demand may be more exposed to availability pressure. A possible interpretation is that repeated purchasing by the same customers can create more persistent demand on stock levels.

Finally, *log_total_orders* has a negative coefficient of -0.301, which is the largest negative effect in the model and is strongly statistically significant ($p = 1.98 \times 10^{-9}$). This means that a one-unit increase in the logarithm of total orders is associated with about 30.1 percentage points lower in-stock availability. This result suggests that, once the other variables are controlled for, products with substantially higher total demand may be more difficult to keep available, likely because their demand puts greater pressure on inventory. One interpretation is that these products are more exposed to stock constraints because they are ordered more intensively.

Overall, the results show that in-stock availability is associated with product turnover, B2B demand intensity, total-order intensity, and customer concentration, while product-class effects are not statistically significant in this specification.

### 8.2.2 Model evaluation and performance metrics

The performance of the regression model is evaluated using $R^2$ and RMSE, which are appropriate metrics for continuous outcomes in OLS regression, as shown in Table 8.2. These metrics provide complementary information about the model's explanatory power and predictive accuracy.

| Metric | Value |
|---|---|
| R-squared | 0.250 |
| RMSE | 0.302 |
| MAE | 0.235 |

*Table 8.2 - Test-set performance of the stock-level regression model, with $R^2$, RMSE and MAE reported as evaluation metrics.*

The $R^2$ is 0.25, which means the model explains about 25% of the variation in stock levels across products. This is a moderate result, because it shows that the variables in the model do capture some important patterns, but stockouts are also influenced by many other factors that are not included here, such as replenishment rules, supplier delays, seasonal demand, or promotional activity.

The RMSE is 0.302, which shows the average size of the prediction error. In practice, this means that the model gives a useful general picture of stockout behavior, but it is not precise enough to predict the exact in-stock percentage for every product. This means that this model is weak for product-level predictions and it is useful only for exploratory prioritization, so many products will still be overestimated or underestimated.

Taken together, the model is useful for identifying trends, not for making exact forecasts. It helps show which products are more exposed to stockouts and which ones are less at risk, so it can support better inventory prioritization.

The most important result is that *log_orders_from_b2b_customers* and *log_product_turnover* are associated with higher in-stock availability (both statistically significant), while *log_total_orders* and *average_orders_per_customer* are associated with lower in-stock availability (both statistically significant). The *product_class_grouped_Project_Customized* coefficient is positive but not statistically significant. This means that products that are commercially important and strongly linked to B2B demand tend to be better available, whereas products with broader and more regular order patterns appear to be slightly less stable in stock. As a result, these findings can help Prysmian identify which products require closer monitoring of stock availability and more careful replenishment.

## 8.3 Relationship between stock issues and B2B ordering behavior

This section integrates the stock-level findings from Section 8.2 with the order-level analysis from Chapter 6 to provide a comprehensive understanding of how stock availability is correlated with both sales performance and B2B adoption. The two analyses reveal that stock availability is a critical enabler of digital channel usage through two complementary mechanisms.

The order-level analysis suggests that stock availability is associated with channel choice, while the stock-level analysis helps identify which products are more exposed to availability risk.

When products are out of stock, customers usually have three possible reactions. They may postpone the purchase, which leads to delayed turnover and may also reduce customer satisfaction. They may switch to offline ordering, for example by contacting sales representatives to check availability or arrange special shipments. Alternatively, they may replace the product with another one, which can change the turnover value of the order and may also weaken customer satisfaction.

When a product is available in stock, customers are more likely to order it through the B2B channel rather than offline, because they face less uncertainty and can complete the purchase more easily on their own. Moreover, stock availability may support immediate purchasing because customers can buy without delaying the order or turning to alternatives. However, the analysis does not estimate a causal effect of stock availability on turnover. In this sense, improving stock availability can both raise overall sales and increase the share of transactions handled through B2B, which in turn reinforces digital adoption.

The B2B platform is designed as a self-service channel, so product availability is essential for its success. When products are in stock, customers do not need to contact sales representatives to check availability or solve stock issues, which reduces friction and supports online ordering. For this reason, the Spanish Business team should prioritize products that are especially important for B2B customers, ensure that critical products remain consistently available, and monitor stockouts carefully to identify where improvements would have the greatest impact on offline ordering (Rosales, Whipple, & Blackhurst, 2018).

## 8.4 Implications for stock allocation

The stock-level regression shows that stock issues are not only an operational problem, but also something that can affect customer behavior and the use of the B2B channel. The results suggest that stock allocation should be managed more selectively, with special attention to the products that are most exposed to availability problems.

The product-class coefficients are not statistically significant. Therefore, the model does not provide reliable evidence that *Project_Customized* or *Standard_Installer* products differ from the reference category in terms of in-stock availability. Inventory implications should therefore be based on the significant demand-related variables, product turnover, B2B customer orders, total orders, and average orders per customer, rather than on non-significant product-class coefficients.

A first implication is that products with higher turnover and stronger B2B demand should receive closer inventory control. The results show that *log_product_turnover* and *log_orders_from_b2b_customers* are both associated with higher in-stock availability. This means that the products that matter most commercially, and that are also important for digital customers, tend to be better supported from an inventory perspective.

In the current specification, *log_total_orders* has a negative and strongly significant coefficient (-0.301), suggesting that frequently ordered products may be more difficult to keep available because demand is more intense and puts greater pressure on inventory (Evans, 2023).

Another important implication is that stock allocation should support B2B adoption, not only sales performance. If products that are important for B2B customers are often unavailable, customers may move back to offline ordering or delay their purchases, which may reduce their trust in the B2B platform. For this reason, inventory decisions should consider not only turnover, but also the relevance of each product for online customers.

Finally, stock allocation should be reviewed regularly because demand patterns can change over time. A product that is stable today may become critical later, and a product with low exposure now may become more vulnerable if demand increases. Continuous monitoring of stock availability, B2B order patterns, and product turnover can help Prysmian adjust inventory policies before issues become more serious.

Overall, the findings suggest that Prysmian should move from a general stock policy to a more selective one. The stock-level model should not be interpreted as evidence that stock availability causes higher turnover. Since in-stock availability is the dependent variable, the model shows that higher-turnover and B2B-relevant products tend to display higher availability. The causal effect of stock availability on turnover would require a different design, with turnover as the outcome and lagged or exogenous variation in stock availability as the explanatory variable. This more focused approach can help reduce stock problems and better support digital adoption.

## 8.5 Limitations

Despite the meaningful and economically interpretable results, several important limitations must be considered when interpreting the findings from the stock-level regression analysis. These limitations affect the representativeness, causal interpretation, and generalizability of the results, and they define the scope within which the strategic implications for stock allocation should be applied. Acknowledging these limitations is essential to ensure that inventory management decisions are grounded in a realistic understanding of the analysis’s strengths and constraints.

The most significant limitation is the substantial reduction of the dataset due to missing stock information at the product level. Stock information is available for only 582 products out of an initial sample of 5,261, representing a decrease of approximately 89% in the number of observations used in the regression. This reduction may affect the representativeness of the sample and limit the generalizability of the results to the full product portfolio. Additionally, the 582 products may not adequately represent all product classes or segments, potentially missing important heterogeneity in stock availability across different product types. The smaller sample size also reduces the statistical power of the analysis, making it more difficult to detect smaller effects or to estimate coefficients with high precision. As a result, the findings should be interpreted cautiously when applied beyond the 582-product sample.

Moreover, there is a potential selection bias: the 582 products with stock data may not be randomly selected from the full portfolio. Products with stock information might systematically differ from those without in ways that affect stock availability mechanisms, for example, they could be higher-turnover products, products from certain categories, or products that managers considered more important. This selection bias could affect the estimated coefficients if the stock availability mechanisms differ between products with and without stock data, and without information on why stock data is missing for 4,679 products, it is difficult to assess the direction and magnitude of this bias.

A second important limitation is that the results identify conditional associations rather than causal effects. The model estimates the relationship between product characteristics and stock availability, but it does not prove that one factor directly causes higher or lower availability. For example, products with higher turnover or stronger B2B demand may be more exposed to stock issues because they are sold more frequently and inventory may not be replenished quickly enough. In this case, lower stock availability may be a consequence of strong demand rather than its cause. Without more advanced methods such as instrumental variables, lagged variables, or panel data techniques with fixed effects, the estimated relationships should therefore be interpreted as conditional associations rather than strictly causal effects (Leszczensky & Wolbring, 2022).

A third limitation is the presence of unobserved factors that may simultaneously influence both stock availability and product performance. These factors are not explicitly included in the model but may still play an important role in shaping inventory outcomes. Examples include pricing policies, promotions, supply chain delays, logistics constraints, product life cycle, seasonal demand patterns, and internal inventory allocation rules. Since these elements are not directly measured in the regression, they may partly explain the variation in stock availability and may bias the estimated effects of the included variables.

Another limitation is the moderate explanatory power of the model. The reported $R^2$ indicates that the model explains only part of the variation in stock availability across products, which means that a relevant share of the observed differences remains unexplained. This is not unusual in operational data, but it implies that the model is more useful for identifying broad patterns than for making precise predictions for individual products.

For these reasons, the results should be used as a guideline rather than as a fixed rule. They can help Prysmian identify which products are more exposed to stock availability problems and where inventory attention should be focused, but final stock decisions should still consider the Spanish team's judgments.

In summary, the main limitations of the stock-level analysis are the reduced sample size, the non-causal nature of the estimated relationships, the presence of omitted factors, and the moderate explanatory power of the model. Even so, the analysis still provides useful insight into which products are more likely to face stock availability issues and how inventory management can be improved.

# Chapter 9. Matching and Propensity Scores

The previous chapters identify several statistical patterns associated with B2B e-commerce adoption in Prysmian Spain, but they do not establish whether changes in these factors are associated with changes in ordering behavior. To address this limitation, this chapter uses propensity score matching to construct more comparable treated and control groups and to approximate causal reasoning.

The customer-level model examines which customer characteristics are associated with B2B actual usage, the order-level model analyzes which order characteristics are associated with offline ordering, and the product-level and stock-level analyses investigate how catalog inclusion and stock availability relate to digital channel usage. These analyses provide useful descriptive, explanatory, and predictive evidence. However, they do not fully answer a more demanding question that is central to management research: whether changing a managerial lever is likely to change an outcome of interest.

This distinction between correlation, prediction, and causation is especially important in business and management research. A predictive model may accurately identify customers who are likely to be B2B users, products that resemble current B2B products, or orders that are likely to remain offline. However, prediction alone does not necessarily indicate what should be done to change the situation.

For this reason, this chapter introduces a causal perspective through propensity score matching. The objective is not to claim definitive causal identification, because B2B actual usage by customers is not randomly assigned. Rather, the objective is to move beyond purely correlational and predictive evidence by constructing more credible comparisons between treated and untreated units. In this chapter, “treated” units are customers who are actually ordering on the B2B platform, while “control” units are comparable customers who are not. Matching is used to approximate the counterfactual logic of causal inference: what would the outcome of a B2B-active customer have looked like if it had not been active, compared with similar non-active customers?

This approach is theoretically and methodologically relevant for the thesis because the research question concerns not only which factors are associated with B2B e-commerce adoption, but also which levers firms can use to improve digital channel migration. B2B actual usage is an outcome shaped by managerial and behavioral factors, not a passive characteristic. Therefore, analyzing it only through correlations would leave the thesis unable to address the action-oriented dimension of the research question. By introducing matching, the analysis attempts to evaluate whether B2B usage is associated with differences in commercially relevant outcomes after reducing observable selection bias.

At the same time, the chapter remains cautious in its interpretation. Propensity score matching can reduce bias caused by observed differences between treated and control units, but it cannot eliminate all threats to causal inference. Omitted variables, reverse causality, endogeneity, and residual imbalance may still affect the results. Therefore, the results should be interpreted as matched comparisons that approximate causal evidence under specific assumptions, rather than as definitive proof of causal effects.

The chapter is organized into two sections.

Section 9.1 focuses on customer-level matching and examines whether B2B actual usage is associated with different customer turnover after matching B2B users and non-B2B users with similar observed characteristics.

Section 9.2 applies the same logic at the product level and evaluates whether B2B product availability is associated with higher turnover after matching B2B-enabled and non-enabled products on observable demand-related characteristics.

## 9.1 Customer-level matching: Turnover differences associated with B2B usage

This section evaluates whether B2B actual usage is associated with differences in customer turnover once B2B users are compared with non-users that are similar in observable characteristics. The analysis uses propensity score matching to reduce selection bias and to construct a more credible counterfactual benchmark than a simple comparison of average turnover between B2B users and non-users.

The matched-comparison question underlying this section can be formulated as follows: among customers with similar observable profiles, is being active on the B2B platform associated with higher or lower turnover? This question differs from the customer-level regression in Chapter 5.

The regression model identifies which customer characteristics are associated with the probability of B2B usage. By contrast, the matching analysis asks whether B2B usage itself is associated with a difference in outcome after accounting for the fact that B2B users are not randomly selected. This distinction matters for managerial decision-making. If B2B users have higher turnover in the raw data, this may simply reflect the fact that the company tends to activate or retain customers that are already larger or more commercially attractive. Conversely, if B2B users have lower turnover, this may reflect selection toward smaller but more standardized customers. In both cases, the raw difference would be misleading if interpreted as the effect of usage.

Matching therefore helps separate, as far as the available data allow, the relationship between B2B usage and turnover from the observable characteristics that influence both.

### 9.1.1 Methodology

To evaluate the relationship between B2B actual usage and customer turnover, a propensity score matching approach was implemented. The treatment variable is *b2b_active*, which indicates whether a customer is active on the B2B platform. The outcome variable is customer turnover. The propensity score represents the estimated probability that a customer is active on the B2B platform conditional on observable characteristics. These characteristics were selected because they may influence both the probability of B2B usage and the level of customer turnover. The variables include *log_total_orders* and *log_total_products* as proxies for customer size and purchasing intensity, *avg_products_per_order* as a proxy for order complexity, *ceec_orders* as a measure of prior digital exposure, and categorical controls such as *first_level_reduced* and *province_macro* to account for organizational and geographic heterogeneity.

This specification is motivated by the need to address omitted-variable bias from observable sources. Customers differ in size, purchasing breadth, ordering complexity, organizational group, and geographic context. If these factors are not controlled for, any comparison between B2B users and non-users may confuse the effect of B2B usage with pre-existing differences in customer structure. Matching attempts to reduce this problem by comparing B2B users with non-users that have similar probabilities of being active on the platform.

After estimating the propensity scores, a nearest-neighbor matching procedure was applied. The purpose of the matching procedure is to create a control group of non-users that resembles the user group in terms of the observed covariates. A balance check was then used to assess whether the matching procedure improved comparability between treated and control customers.

The initial matching procedure revealed imbalance in *ceec_orders*, B2B users differed from non-users in their prior use of another electronic channel. This imbalance is theoretically important because prior digital exposure may influence both the probability of B2B usage and subsequent turnover. To improve balance, *ceec_orders* was transformed into *log_ceec_orders,* reducing the influence of extreme values, and the matching caliper was tightened. These adjustments improved comparability between groups, although they did not eliminate all imbalance.

This methodological process reflects the broader objective of the chapter. The goal is not to eliminate all possible sources of bias, which is not possible with the available observational data, but to make the comparison more credible than a simple unconditional difference. In this sense, the matching design represents an attempt to address endogeneity and selection bias by controlling for observable factors that may jointly affect treatment assignment and the outcome.

### 9.1.2 Results

The balance check reported in Table 9.1 shows that the matching procedure achieves a reasonable level of comparability across several key variables.

| b2b_active | log_total_orders | log_total_products | avg_products_per_order | log_ceec_orders |
|---|---|---|---|---|
| 0 | 4.11 | 4.60 | 1.94 | 1.82 |
| 1 | 4.05 | 4.51 | 1.90 | 2.22 |

*Table 9.1 - Balance check for customer-level matching, comparing logarithm of the total number of orders and of the total products, average number of products per order, log of the number of CEEC orders between B2B-active and non-active customers.*

In the balance check in Table 9.1, the average values of *log_total_orders*, *log_total_products*, and *avg_products_per_order* are similar across B2B users and non-users after matching. This suggests that the matched groups are broadly comparable in terms of customer scale, purchasing breadth, and order complexity.

However, some imbalance remains in *log_ceec_orders,* which is higher among B2B users than among matched non-users. This residual difference indicates that B2B users may still be more digitally exposed or more digitally mature than their matched counterparts. As a result, the estimated effect should not be interpreted as a fully isolated causal effect of B2B usage. Rather, it should be interpreted as a matched association that is closer to causal evidence than a raw comparison, but still dependent on the assumption that the most relevant confounders have been adequately controlled.

| Measure | Value |
|---|---|
| Mean turnover (B2B) | 979.27 |
| Mean turnover (non-B2B) | 1752.53 |
| Matched turnover difference | -773.26 |

*Table 9.2 - Matched turnover difference in the matched sample for customer-level matching, showing mean turnover (k€) for B2B and non-B2B customers and the negative matched turnover difference associated with B2B actual usage.*

The estimated matched turnover difference reported in Table 9.2 is negative. In the matched sample, B2B users show lower average turnover than comparable non-users. This result should not be interpreted as evidence that B2B usage causes customer turnover to decline. A more appropriate interpretation is that, after matching on observable characteristics, B2B users in this sample are not associated with higher turnover than comparable non-users.

This finding is theoretically meaningful because it challenges a simplistic view of digital usage as an automatic performance lever. The negative estimate suggests that B2B usage may be a necessary but insufficient condition for commercial impact.

A possible interpretation is that customers who actually use the B2B platform may have been selected because they are more standardized or more compatible with the platform, rather than because they generate the highest absolute turnover.

From a methodological perspective, this result illustrates why causal thinking is valuable. The matching analysis asks whether B2B usage is associated with a different outcome once comparable customers are considered. Even though the answer remains uncertain, the question is more aligned with managerial decision-making than a purely descriptive comparison.

Overall, the customer-level matching analysis suggests that B2B usage should not be treated as an isolated causal lever. Its commercial effect is likely conditional on other mechanisms, including catalog fit, product availability, order complexity, customer habits, and organizational support.

## 9.2 Product-level matching: Turnover differences associated with B2B availability

This section applies the same matching logic at the product level. The objective is to evaluate whether products available in the B2B catalog are associated with higher turnover than comparable products not available in the catalog. While the customer-level analysis focuses on access to the platform, the product-level analysis focuses on assortment availability as a potential lever of digital adoption and commercial performance.

A simple comparison between B2B and non-B2B products may be misleading because products are not randomly included in the catalog. Products may be selected for B2B because they are standardized, commercially important, frequently ordered, or more suitable for self-service purchasing. Therefore, higher turnover among B2B products could reflect pre-existing product characteristics rather than the effect of catalog inclusion itself. Matching attempts to reduce this problem by comparing B2B-enabled products with non-enabled products that are similar in observable demand-related characteristics.

This section is particularly relevant to the research question because assortment completeness is one of the central theoretical mechanisms of the thesis. If customers cannot find the products they need in the B2B catalog, digital ordering becomes less feasible. Product availability in the catalog is therefore not just a descriptive feature of the platform; it is a potential intervention that may influence whether customers can shift from offline to online ordering.

### 9.2.1 Methodology

The treatment variable in the product-level matching analysis is the binary indicator of whether a product is available in the B2B catalog (*b2b_enabled_product*). The outcome variable is product turnover. The goal is to compare B2B enabled and non-enabled products that are similar in terms of observable product demand and usage patterns.

The propensity score was estimated using variables that capture product-level demand characteristics: *percentage_orders, total_orders*, and *average_orders_per_customer*. These variables were selected because they may influence both the likelihood that a product is included in the B2B catalog and the turnover generated by the product. For example, products with broader order penetration or repeated demand may be more likely to be considered suitable for digital inclusion, and they may also generate higher turnover regardless of B2B availability.

After estimating the propensity scores, a nearest-neighbor matching algorithm with a caliper was applied. The caliper restricts matches to products with sufficiently similar propensity scores, reducing the risk of comparing products that are too different in their underlying probability of treatment. The final matched sample includes 883 B2B-enabled products and 883 non-enabled products, indicating substantial overlap between the two groups.

The balance check assesses whether the matching procedure creates comparable treated and control groups.

### 9.2.2 Results

As shown in Table 9.3, *percentage_orders* is almost identical across B2B and non-B2B products after matching, which suggests good comparability on this dimension.

| b2b_enabled_product | percentage_orders | total_orders | average_orders_per_customer |
|---|---|---|---|
| 0 | 0.42 | 287.16 | 3.26 |
| 1 | 0.42 | 219.06 | 2.11 |

*Table 9.3 - Balance check for product-level matching, comparing percentage of orders, total number of orders, and average number of orders per customer between enabled and non-enabled products.*

However, some residual differences remain in *total_orders* and *average_orders_per_customer*. Therefore, the product-level matching improves comparability but does not eliminate observable imbalance, and the result should be interpreted as a matched association within the matched sample rather than as fully generalizable causal evidence.

The relatively large matched sample size, with 883 treated and 883 control products, improves the statistical reliability of the results and supports a more generalizable interpretation compared to the customer-level analysis.

| Measure | Value |
|---|---|
| Mean turnover (B2B) | 567.89 |
| Mean turnover (non-B2B) | 137.91 |
| Matched turnover difference | 429.98 |

*Table 9.4 - Matched turnover difference in the matched sample for product-level matching, showing mean turnover (k€) for B2B and non-B2B products and the positive matched turnover difference associated with B2B enablement.*

The product-level matching results reported in Table 9.4 show a positive matched turnover difference. In the matched sample, B2B-enabled products generate substantially higher average turnover than comparable non-enabled products. The estimated difference is approximately 429.98 thousand euros, with mean turnover of approximately 567.89 thousand euros for B2B-enabled products and 137.91 thousand euros for non-enabled products.

This result is consistent with the idea that B2B product availability may be commercially valuable. One possible interpretation is that making products available in the B2B catalog reduces transaction friction, improves accessibility, and supports sales among customers who are already close to digital ordering. This interpretation is aligned with the broader theoretical argument that assortment completeness is a key condition for B2B e-commerce adoption. If products are available online, customers have a greater opportunity to complete orders through the digital channel; if products are absent from the catalog, online ordering is structurally constrained.

However, the estimate should still be interpreted cautiously. The positive difference does not definitively prove that B2B catalog inclusion caused the higher turnover. Reverse causality remains possible: products may have been added to the B2B catalog precisely because they were already commercially attractive, standardized, or strategically important. Omitted variables may also remain, including pricing policies, product margins, product class, promotional activity, sales priorities, technical feasibility, and internal catalog decisions. Therefore, the result should be described as evidence consistent with a positive effect of B2B availability, not as conclusive proof of such an effect (Leszczensky & Wolbring, 2022).

Despite this caution, the product-level result is theoretically important. It suggests that product availability may be a more direct lever of digital channel migration than customer enablement alone. This distinction matters because B2B adoption occurs through concrete transactions. A customer may be willing to use the platform, but if the required products are not available, the order cannot be completed digitally. Product-level enablement therefore operates closer to the transaction itself and may have a more immediate relationship with turnover.

From a methodological perspective, the product-level matching analysis also demonstrates the value of moving beyond prediction. The product-level logistic regression in Chapter 7 identifies which products resemble current B2B catalog products. However, the matching analysis asks a different question: among products with similar observable demand characteristics, are those available in B2B associated with higher turnover? This question is closer to the type of reasoning required for managerial action, because it concerns the potential outcome of changing catalog availability.

The results offer cautious but meaningful evidence that B2B product availability is associated with stronger commercial outcomes, even after accounting for selected observable characteristics. Although the analysis does not remove all causal ambiguity, it suggests that catalog inclusion is not merely a technical feature of the platform. Instead, it may be one of the mechanisms through which B2B e-commerce generates commercial value.

Overall, this chapter reinforces the thesis argument that B2B e-commerce adoption is a multi-level and conditional process. B2B actual usage, catalog availability, and operational reliability should not be viewed as isolated drivers, since their effects depend on how they interact. By approximating causal inference through matching, the chapter strengthens both the theoretical and managerial contribution of the thesis, offering a more credible basis for identifying actionable drivers of digital channel migration.

# Chapter 10. Customer, Product, and Stock Prioritization

This chapter identifies the main opportunities to increase B2B adoption in Prysmian Spain by prioritizing the customers, products, and stock conditions with the highest potential effect. Building on the analyses developed in the previous chapters, the purpose of this chapter is to move from general evidence to a more operational view of where B2B activation can generate the largest turnover gains and reduce the main barriers to B2B usage.

The chapter is organized into five sections.

Section 10.1 focuses on new customers not yet enabled for B2B, based on their purchasing behavior and on the share of products already available in the B2B catalog. A sensitivity analysis is then used to define the threshold for identifying potential B2B customers to isolate customers whose purchasing patterns already match the B2B offer and who therefore appear to be the most natural candidates for activation. However, because the previous analyses show that enablement and actual usage are not the same thing, this section should be read as identifying potential candidates for activation, not guaranteed B2B users.

Section 10.2 examines the behavior of customers who are already enabled for B2B but still place part of their orders offline. By comparing offline and B2B orders, this part of the chapter helps explain why enabled customers do not always use the digital channel, and whether the reason lies in product availability, stock issues, habitual behavior, or other commercial factors.

Section 10.3 looks at the composition of offline orders to understand whether these orders are driven by products already available in B2B, by products missing from the catalog, or by a combination of both. This provides a clearer picture of one of the main drivers behind offline purchasing among B2B-enabled customers, namely missing products from the B2B catalog.

Section 10.4 turns to product prioritization. Here, the analysis identifies products that are not yet available in the B2B catalog but are already purchased offline by B2B-enabled customers and estimates the turnover that could potentially be recovered by adding them to the digital channel. This section therefore highlights the products with the greatest potential for catalog expansion.

Section 10.5 focuses on stock availability and stock patterns. In this part, the chapter assesses which B2B products are most frequently out of stock. A sensitivity analysis is also used to define the threshold for classifying products as critical out-of-stock products.

## 10.1 New customers: non-enabled customers and addressable turnover

The focus is therefore on identifying non-enabled customers who are commercially compatible with the current B2B offer and therefore may be suitable candidates for activation. The analysis focuses on customers who are not currently enabled for B2B and examines the structure of their orders to assess whether they could be enabled to the online channel.

The data used for this section come from the Cockpit dataset and include only customers not enabled for B2B. In total, 1,049 customers were identified as not enabled for B2B. Customers who always place orders through other digital channels were excluded. They are not the focus of this analysis, since these customers already use other online channels and do not rely on manual ordering. The objective is instead to identify customers with a high share of manual orders, who could be migrated to B2B. After this exclusion, the final sample comprises 1,032 non-B2B customers.

For each non-B2B customer, two main indicators were calculated. The first is the percentage of purchased products available in the B2B catalog. A high percentage suggests that the customer mainly buys products that are already compatible with the B2B platform and therefore represents a potential B2B customer. The second indicator is the turnover generated by products available in the B2B catalog. A high value for this measure indicates a potentially relevant turnover amount that could be shifted to the online channel if the customer were enabled.

In addition, the percentage of orders placed through Electronic E-Commerce in Spain and the percentage of manual orders were calculated for each customer. This step is important because the analysis does not focus on customers who already rely heavily on the Electronic E-Commerce. Instead, the objective is to identify customers with a high share of manual orders, who are more likely to benefit from a migration to the B2B platform.

Potential new B2B customers were then identified as those purchasing a high share of products already available on B2B, using a threshold of at least 95% products already available in the B2B catalog. Based on this criterion, 183 non-enabled customers were identified as potential B2B candidates. Together, these customers generate approximately 31 million euros in turnover. Within this group, 95 customers purchase only products that are available on B2B, generating approximately 8 million euros in turnover.

From this subset, 17 top potential B2B customers were selected (Table 10.1). These customers purchase only products available on the B2B platform and each of them generates more than 150,000 euros in turnover. As a result, they represent a particularly interesting group for further investigation, since they combine a very high alignment with the B2B catalog and a significant commercial value.

| Customer Code | First Level | Billing City | Sales Owner | % B2B | Turnover prod B2B (k€) |
|---|---|---|---|---|---|
| Customer_1 | FEGIME INTERNATIONAL | Burgos | Joan Ferran | 100 | 767 |
| Customer_2 | FEGIME INTERNATIONAL | Burgos | Joan Ferran | 100 | 568 |
| Customer_3 | FEGIME INTERNATIONAL | Burgos | Joan Ferran | 100 | 477 |
| Customer_4 | FEGIME INTERNATIONAL | Burgos | Joan Ferran | 100 | 470 |
| Customer_5 | GETELESA | Malaga | Joan Ferran | 100 | 469 |
| Customer_6 | FEGIME INTERNATIONAL | Burgos | Joan Ferran | 100 | 378 |
| Customer_7 | FEGIME INTERNATIONAL | Burgos | Joan Ferran | 100 | 298 |
| Customer_8 | FEGIME INTERNATIONAL | Burgos | Joan Ferran | 100 | 263 |
| Customer_9 | GETELESA | Malaga | Joan Ferran | 100 | 227 |
| Customer_10 | FEGIME INTERNATIONAL | Burgos | Joan Ferran | 100 | 215 |
| Customer_11 | FEGIME INTERNATIONAL | Burgos | Joan Ferran | 100 | 200 |
| Customer_12 | FEGIME INTERNATIONAL | Burgos | Joan Ferran | 100 | 197 |
| Customer_13 | ELECNOR | Alicante | Joan Ferran | 100 | 186 |
| Customer_14 | FEGIME INTERNATIONAL | Burgos | Joan Ferran | 100 | 177 |
| Customer_15 | IMELCO | Espulgues de Llobregat | Joan Ferran | 100 | 175 |
| Customer_16 | GRUPO SONEPAR | Vigo | Joan Ferran | 100 | 175 |
| Customer_17 | FEGIME INTERNATIONAL | Burgos | Joan Ferran | 100 | 151 |

*Table 10.1 - Top non B2B-enabled customers ranked by turnover from products available in B2B catalog, showing customer code, company, geographic location, sales owner, percentage of B2B orders, and turnover from products available in B2B catalog (k€).*

Overall, these 17 customers account for approximately 5.4 million euros in potentially B2B-addressable offline turnover among customers not currently enabled for the platform. This finding suggests that, for these customers, the main barrier to online adoption may not be product availability or stock issues, but rather the absence of B2B activation. For this reason, an internal validation is recommended to understand why these customers have not yet been enabled and whether they could be enabled.

### 10.1.1 Sensitivity analysis: threshold for potential B2B customers

A sensitivity analysis was conducted to determine the most appropriate threshold for identifying potential B2B customers. Based on the chart in Figure 10.1, a threshold of 95% of B2B products purchased was selected. This means that a customer is considered a potential B2B candidate if at least 95% of the products purchased by the customer are already available on the B2B platform.

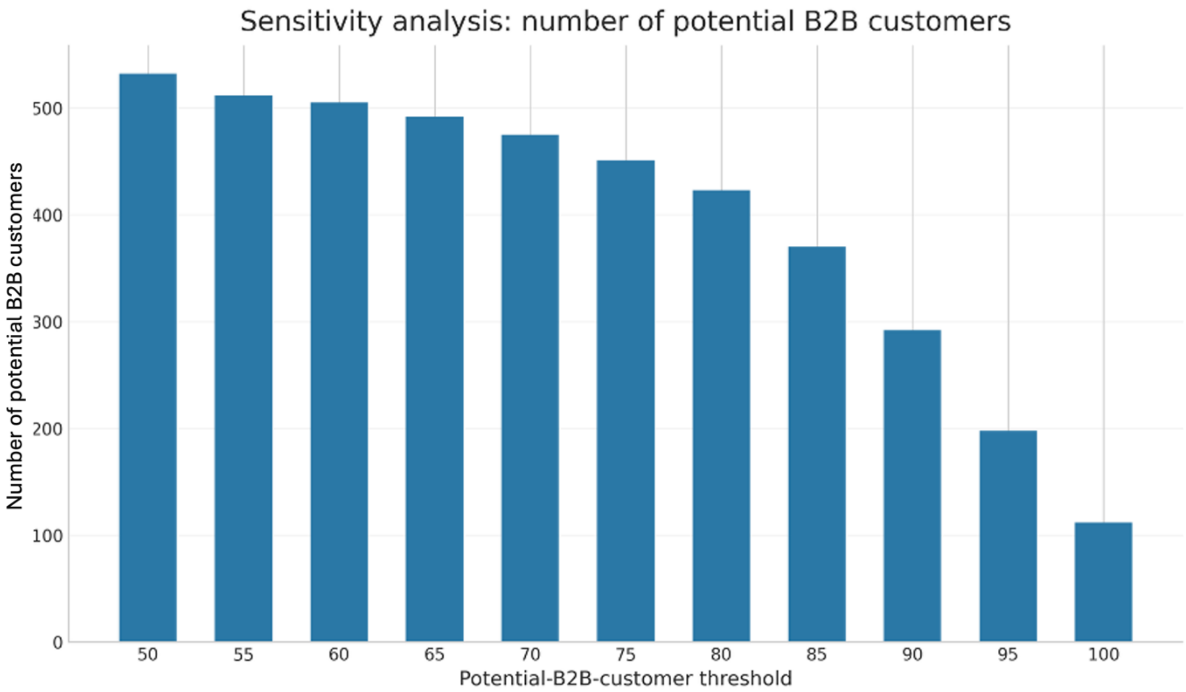


*Figure 10.1 - Histogram of the number of customers across different threshold values, with the 95% threshold selected as the criterion for identifying potential B2B customers.*

At this threshold level, the number of identified customers (183 customers) is significantly reduced compared to lower thresholds, but the selected group shows a much stronger alignment with the B2B catalog. Their purchasing behavior suggests that they already rely almost entirely on products that can be ordered through the online channel, which makes them strong candidates for B2B enablement. In this sense, the threshold helps isolate customers for whom the lack of enablement is likely to be an important barrier preventing a shift to the B2B platform.

Using a lower threshold would increase the number of potential customers, but it would also include more mixed purchasing patterns, making B2B adoption less straightforward and more dependent on additional actions such as catalog expansion. By contrast, the 95% threshold focuses on customers whose product mix is already highly compatible with the existing B2B offering and therefore offers a more precise basis for identifying near-term enablement opportunities.

## 10.2 Enabled customers' behavior: offline vs B2B orders

The objective of this analysis is to examine the behavior of customers who are already enabled for the B2B platform, in order to understand why these customers do not always place their orders through the B2B channel.

The analysis focuses on enabled customers who place part or most of their orders offline, investigating the structure of their orders to identify patterns that may explain the limited use of the online channel.

The data used for this section come from the Cockpit dataset and include only B2B-enabled customers. Out of 1,446 Spanish customers, 404 were identified as B2B-enabled before excluding customers using only the Spanish Electronic E-Commerce channel. After this exclusion, the analytical enabled sample consists of 397 customers, of whom 119 placed at least one B2B order in 2025, while 278 customers have never used the B2B platform. If only customers with positive turnover are considered, the number of customers who have never used B2B becomes 276. This gap between enablement and actual usage highlights that the enablement for the platform is not sufficient to ensure migration to the online channel.

For each B2B-enabled customer, two main indicators were calculated: the percentage of orders placed through B2B and the percentage of orders placed offline (non-B2B). These indicators provide a direct measure of how much each enabled customer relies on the online channel compared with traditional channels.

| Customer Code | % non-B2B orders | % CEEC orders | % manual orders | Turnover from non-B2B orders (K€) |
|---|---|---|---|---|
| Customer_1 | 100 | 0 | 100 | 15181 |
| Customer_2 | 100 | 73 | 27 | 5722 |
| Customer_3 | 100 | 77 | 23 | 4305 |
| Customer_4 | 100 | 0 | 100 | 4249 |
| Customer_5 | 100 | 0 | 99 | 3376 |
| Customer_6 | 100 | 0 | 99 | 3095 |
| Customer_7 | 100 | 8 | 92 | 2491 |
| Customer_8 | 100 | 0 | 84 | 2412 |
| Customer_9 | 100 | 85 | 15 | 2213 |
| Customer_10 | 100 | 0 | 100 | 2097 |
| Customer_11 | 100 | 1 | 99 | 2085 |
| Customer_12 | 100 | 64 | 35 | 1876 |
| Customer_13 | 100 | 78 | 22 | 1873 |
| Customer_14 | 100 | 33 | 66 | 1809 |
| Customer_15 | 100 | 0 | 96 | 1692 |
| Customer_16 | 100 | 57 | 43 | 1554 |
| Customer_17 | 100 | 0 | 88 | 1528 |
| Customer_18 | 100 | 0 | 92 | 1524 |

*Table 10.2 - Top B2B-enabled customers ranked by non-B2B turnover, showing customer code, percentage of orders non B2B, Electronic E-Commerce and manual, and turnover from non-B2B orders (k€).*

Table 10.2 shows 18 B2B-enabled customers who do not place any orders through B2B, including their percentage of non-B2B orders, Electronic E-Commerce orders, manual orders, and associated offline turnover. Each of these customers generates more than 1.5 million euros in non-B2B turnover. A low share of B2B orders suggests the need for deeper investigation into order characteristics, such as stock availability issues, products not included in the B2B catalog, or the need for direct interaction with sales personnel.

## 10.3 Order composition

The objective of this analysis is to understand the composition of offline orders placed by B2B-enabled customers, to identify the main drivers behind offline purchasing.

The analysis focuses specifically on offline orders from B2B-enabled customers, examining their composition in terms of product availability in the B2B catalog.

The data used for this section come from the Cockpit dataset and include only customers who are enabled to purchase through the B2B platform, excluding those who place orders exclusively through the Spanish Electronic E-Commerce channel. A total of 397 B2B-enabled customers were selected for the analysis, and only offline orders were considered.

Among the total orders, 22,546 orders were identified. Of these, 2,037 orders were placed via B2B (online), 15,525 orders were placed manually (offline) and the remaining 4,984 were placed via other channels. The composition analysis then focuses on the 15,525 manual offline orders.

For each offline order, the composition was analyzed based on the availability of products in the B2B catalog. Three categories were defined: orders including only products available in the B2B catalog (100% B2B products), orders including both available and non-available products (mixed orders), and orders including only products not available in the B2B catalog (100% non-B2B products). The output of the analysis is a table (Table 10.3) that classifies the offline orders according to these three categories.

| Order Type | Percentage |
|---|---|
| Only non-B2B products | 49% |
| Only B2B products | 49% |
| Mixed products | 2% |

*Table 10.3 - Distribution of order types by product composition, showing the share of orders containing only non-B2B products, only B2B products, and both B2B and non-B2B products.*

The analysis of offline orders placed by B2B-enabled customers shows that 49% of offline orders include only products already available in the B2B catalog (Table 10.3). These offline orders are fully compatible with the B2B catalog from a product-coverage perspective, but further analysis is required to determine whether stock availability, pricing, customer habits, or sales support prevented digital ordering.

Another 49% of offline orders include only products that are not available in the B2B catalog. In these cases, product availability is clearly the main barrier preventing online ordering. When required products are not available on the B2B platform, customers are forced to place orders offline. Expanding the B2B catalog would therefore directly address this issue and could potentially capture a significant portion of this offline turnover.

Only 2% of offline orders include a mix of available and non-available products. These cases could be resolved by enabling the missing products in the B2B catalog, allowing customers to place complete orders online. However, given the very small share of such orders, the potential business impact appears limited and may not justify a significant investment in catalog expansion for this specific segment.

Overall, the analysis suggests that product availability is a major driver of offline purchasing among B2B-enabled customers, although the fact that 49% of offline orders involve only products already available in the B2B catalog indicates that other factors also play an important role.

## 10.4 Product catalog coverage: B2B vs non-B2B products and addressable turnover

The objective of this analysis is to identify which products could be added to the B2B catalog in order to increase B2B orders. The analysis focuses on products that are not available in the B2B catalog but are purchased offline by customers already enabled for B2B. These products are then evaluated to identify those with the highest potential for inclusion in the B2B catalog. By doing so, the analysis provides a more granular view of catalog gaps and their potential commercial impact.

The data used for this section come from the Cockpit dataset and include only customers who are enabled on the B2B platform, excluding those who place orders exclusively through the Spanish Electronic E-Commerce channel. A total of 397 B2B-enabled customers were selected for the analysis. Only offline orders placed by these customers were considered, as the goal is to understand which products drive offline purchasing among customers who already have access to the online channel.

The number of offline orders placed by B2B-enabled customers is 15,525. These orders include 1,097 distinct products. Among these offline orders, only mixed orders and orders containing only non-B2B products were taken into consideration for the analysis, in order to identify products that are currently not available in the B2B catalog but are still purchased by B2B-enabled customers. Orders that contain only products already available in the B2B catalog were excluded from this specific analysis, since they do not directly reflect catalog gaps.

The results show that 7,800 orders include either only non-B2B products or a mix of B2B and non-B2B products. Among these orders, there are 264 products with positive turnover that are purchased offline by B2B-enabled customers and are not available in the B2B catalog. These products account for a total turnover of around 9 million euros. They were further analyzed according to turnover, number of times ordered, and product class, in order to identify which products are the most relevant from a commercial perspective.

The output of the analysis is a table (Table 10.4) that includes Product ID, product turnover, number of times the product was ordered (line items), and product class. Products with high turnover and high order frequency represent the best candidates to be added to the B2B catalog, as they are repeatedly purchased by customers and contribute significantly to revenue.

From this set, 13 products were identified as high-priority candidates for B2B inclusion. These are the products with the highest number of line items.

| Product ID | Turnover (k€) | Line Items | Product Class |
|---|---|---|---|
| Product_1 | 333.9 | 3249 | Multimedia Solutions |
| Product_2 | 158.4 | 961 | Multimedia Solutions |
| Product_3 | 273.9 | 900 | Trade and Installers |
| Product_4 | 28 | 784 | Multimedia Solutions |
| Product_5 | 13 | 441 | NWC LV/MV |
| Product_6 | 38.4 | 225 | NWC LV/MV |
| Product_7 | 57.5 | 196 | Multimedia Solutions |
| Product_8 | 21.6 | 144 | Multimedia Solutions |
| Product_9 | 87.7 | 100 | Multimedia Solutions |
| Product_10 | 11.6 | 100 | Multimedia Solutions |
| Product_11 | 122.5 | 81 | Trade and Installers |
| Product_12 | 11.6 | 81 | NWC LV/MV |
| Product_13 | 2.6 | 81 | NWC LV/MV |

*Table 10.4 - Top products by line items, showing product ID, turnover, line items and product class.*

These top 13 products in Table 10.4, currently not available on the B2B platform but purchased offline by B2B-enabled customers, represent approximately 1.16 million euros in B2B-addressable offline turnover associated with products currently absent from the digital catalog. An internal validation is recommended to understand why these products have not yet been included in the B2B catalog, whether due to technical constraints, product classification issues, or strategic decisions about which product lines to prioritize.

A useful point to consider is that the current structure of the B2B catalog is still concentrated on the most standardized product families, called Trade and Installers (T&I), while other product areas remain outside the platform and may represent additional opportunities for expansion, such as Network components (NWC) and multimedia solutions.

| Product Class | Count |
|---|---|
| Trade and Installers | 125 |
| NWC | 75 |
| Multimedia Solutions | 46 |
| Others | 18 |

*Table 10.5 - Distribution of products by class, showing the number of products classified as Trade & Installers, NWC, Multimedia Solutions, and Other.*

Table 10.5 shows that among the 264 candidate catalog-gap products, Trade and Installers is the largest class, followed by NWC and Multimedia Solutions. A potential improvement could be to introduce a dedicated section for Multimedia Solutions products within the B2B platform, which could support higher B2B usage and help capture a segment of demand that is currently served mainly through offline channels.

### 10.4.1 Limitations

This analysis has some important limitations. Offline orders containing only non-B2B products or mixed products account for approximately 51% of total offline orders among B2B-enabled customers. As a result, expanding the B2B product catalog alone would not fully solve the issue of B2B-enabled customers placing orders offline, since a large share of offline orders involves products that are already available in the B2B catalog. However, the products identified in this analysis represent those with the highest potential business impact in terms of order frequency and turnover and therefore offer a focused and evidence-based basis for catalog expansion decisions. By prioritizing these high-impact products, Prysmian can improve the return on investment in catalog development while addressing one of the main operational barriers to B2B adoption.

## 10.5 Stock availability: stockouts and their association with B2B orders

The objective of this analysis is to identify which B2B products show the highest risk of stockouts and therefore could benefit from increased stock levels to reduce out-of-stock events and capture additional online demand. The analysis focuses on products that are available in the B2B catalog but are still purchased offline by B2B-enabled customers. These products are then analyzed to assess their stock availability and to identify which ones are most frequently unavailable, suggesting that stock constraints may be driving customers toward offline channels.

The data used for this section come from the Cockpit dataset and include only customers who are enabled to purchase through the B2B platform, excluding those who place orders exclusively through the Electronic E-Commerce channel. A total of 397 B2B-enabled customers were selected for the analysis. Only offline orders were considered, to identify products that are available in B2B but are still purchased outside the digital channel.

The number of offline orders is 15,525, and these orders include 1,097 distinct products. To examine stock availability for B2B products purchased offline, the analysis considers mixed orders and orders containing products available in B2B. The number of relevant orders, including mixed orders and orders containing B2B-available products, is 7,725. Among these orders, there are 623 products purchased offline by B2B customers that are also available in B2B, representing 45.4 million euros in turnover.

For each product, the following indicators were calculated: the percentage of times the product is out of stock, the percentage of times the product is in stock, and the share of mixed availability cases. Products with a high percentage of out-of-stock occurrences should be prioritized for stock increases, as their unavailability may discourage customers from using the B2B channel.

The results show that 33 additional products, excluding always-out-of-stock items, are out of stock more than 75% of the time, and 28 products are always out of stock. In total, 61 products are classified as critical out-of-stock items. Among these 61 products, 14 represent the highest-priority candidates for stock improvement, as they are out of stock at least 75% of the time and have a total turnover greater than or equal to approximately 600,000 euros.

| Product ID | Share out of stock | Total Turnover (k€) | Product Class |
|---|---|---|---|
| Product_1 | 0.75 | 3006.50 | Trade and Installers |
| Product_2 | 0.79 | 2195.64 | Trade and Installers |
| Product_3 | 0.87 | 1776.66 | Trade and Installers |
| Product_4 | 0.9 | 1766.27 | Trade and Installers |
| Product_5 | 0.95 | 1241.61 | Trade and Installers |
| Product_6 | 0.79 | 1006.74 | Trade and Installers |
| Product_7 | 0.95 | 980.88 | Trade and Installers |
| Product_8 | 0.84 | 954.16 | Trade and Installers |
| Product_9 | 0.79 | 777.81 | Trade and Installers |
| Product_10 | 0.95 | 706.05 | Trade and Installers |
| Product_11 | 0.78 | 692.10 | Trade and Installers |
| Product_12 | 0.83 | 683.41 | Renewables, Trade and Installers |
| Product_13 | 0.8 | 645.85 | Trade and Installers |
| Product_14 | 0.75 | 599.01 | Trade and Installers |

*Table 10.6 - Products ranked by total turnover (k€), showing product ID, share of times out of stock, total turnover and product class.*

These top 14 products in Table 10.6 account for approximately 17 million euros in potential B2B-addressable offline turnover.

| Product Class | Count |
|---|---|
| Trade and Installers | 54 |
| Power Distribution, Trade and Installers | 5 |
| Renewables, Trade and Installers | 2 |

*Table 10.7 - Distribution of products by class, showing the number of products classified as Trade & Installers, Power distribution/Trade & Installers and Renewables/Trade & Installers.*

As seen in Table 10.7, most of these products belong to the Trade and Installers category, which represents the core of the B2B catalog. This is consistent with the product-level analysis (Chapter 7), as standard and simpler products are generally easier to manage operationally and better suited to a self-service channel structure.

### 10.5.1 Sensitivity analysis for identifying out-of-stock products

A sensitivity analysis was conducted to determine the appropriate threshold for classifying products as "out of stock." The analysis tested different thresholds based on the share of out-of-stock occurrences: a threshold of 87.5% identifies 8 products, 75% identifies 33 products, 62.5% identifies 72 products, 50% identifies 133 products, and 25% identifies 236 products.

Very high thresholds, such as 87.5%, identify only a small number of extreme cases, focusing on products that are almost always unavailable, but they represent very few products, so the overall impact would be very limited. Lower thresholds, such as 62.5% or less, significantly increase the number of products classified as out of stock, but they also include products with occasional or non-structural stock issues, where increasing stock may not be justified from a business perspective.

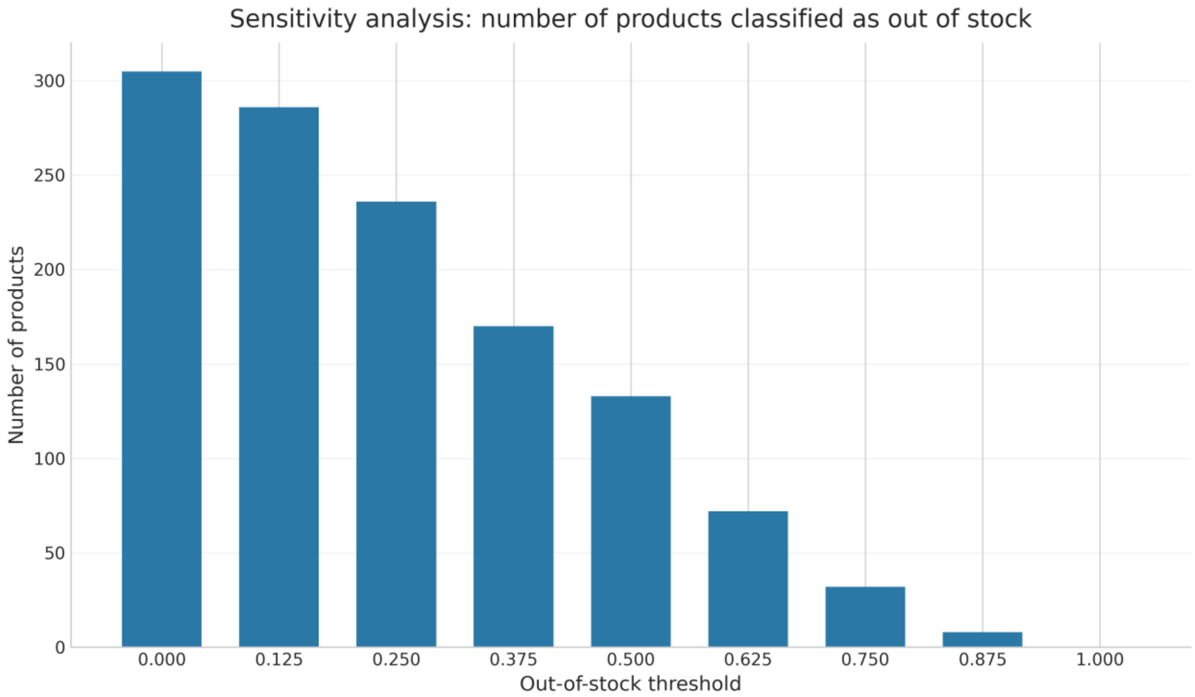


*Figure 10.2 - Histogram of the number of products classified as out of stock across different threshold values, with the 75% threshold selected as the criterion for identifying out of stock products.*

The results in Figure 10.2 suggest that a 75% threshold offers the best compromise between identifying products that are frequently unavailable and keeping the number of flagged items manageable. In this way, the threshold is used to isolate products for which stock improvements are most likely to reduce offline orders and recover additional B2B turnover. It is important to note that the 28 products classified as "always out of stock" are excluded from the graph, so the threshold is applied only to the remaining products. By concentrating on these critical items, Prysmian can prioritize inventory investments where they are most likely to have a measurable impact on B2B adoption.

# Chapter 11. Discussion, Managerial Implications and Limitations

This chapter brings together the main findings of the thesis and interprets them considering the theoretical discussion presented in the initial chapters. After the empirical analyses conducted at the customer, order, product, and stock levels, the objective is to move from isolated results to an understanding of the determinants of B2B e-commerce adoption in Prysmian Spain, in order to derive concrete managerial actions and actionable suggestions.

The chapter is structured in three parts.

Section 11.1 discusses the empirical findings in relation to the existing literature on digital adoption, channel migration, product assortment, and stock availability, with the aim of providing a more comprehensive answer to the research question.

Section 11.2 translates these findings into managerial implications, identifying the main actions that Prysmian Spain can take to improve B2B adoption through more targeted customer selection for B2B usage, more effective catalog development, better stock allocation, and potential future AI-based implementations.

Section 11.3 addresses the limitations of the analysis, discussing the main methodological and data-related constraints that should be considered when interpreting the results and outlining possible directions for future research.

By combining theoretical interpretation and practical recommendations, this chapter seeks to show that B2B adoption is not driven by a single factor, but by the interaction between customer behavior, product compatibility, and operational execution.

## 11.1 Discussion

This thesis set out to investigate how targeted customer activation, assortment completeness, and stock availability are associated with the likelihood of B2B e-commerce ordering in a restricted industrial platform context. The empirical case of Prysmian Spain provides a useful setting for examining this question because it combines three conditions that are theoretically relevant for B2B digital adoption: partial e-commerce usage, incomplete catalog coverage, and operational constraints related to product availability. Rather than treating the case only as a company-specific problem, the findings can be interpreted as evidence of a broader theoretical phenomenon: in industrial B2B markets, digital channel migration is not a simple consequence of platform access, but a multi-level process shaped by the interaction between customer readiness, product suitability, and operational reliability.

### 11.1.1 B2B adoption as a multi-level process

From a methodological perspective, the thesis also highlights why a purely predictive approach would be insufficient for studying B2B channel migration. A model may identify which customers are more likely to be B2B users or which products resemble current B2B catalog items, but prediction alone does not tell managers which actions are likely to change adoption. The attempt to use propensity score matching therefore reflects a more action-oriented logic: it asks whether managerial levers are associated with different outcomes when treated and untreated units are made more comparable. Although the available data do not allow definitive causal claims, this attempt is theoretically important because it shifts the analysis from “who is likely to adopt?” to “which intervention may alter adoption conditions?”

Taken together, the findings suggest that B2B e-commerce adoption in industrial markets should be theorized as a multi-level alignment process. At the customer level, adoption depends on actual usage, organizational readiness, purchasing routines, and willingness to change. At the order level, it depends on whether the specific transaction can be completed digitally. At the product level, it depends on whether the items are sufficiently standardized and included in the catalog. At the operational level, it depends on whether stock availability makes the digital channel reliable. This multi-level perspective helps explain why adoption remains partial even when a platform is already available. Barriers do not operate independently; rather, they reinforce one another. A customer may be selected for B2B but not behaviorally ready; a product may be suitable but not available; an order may contain mostly digital products but still require offline handling because of one missing item or stock uncertainty.

In practical terms, this means that the research question is answered by showing that B2B e-commerce adoption is shaped by the interaction between customer B2B usage, assortment completeness, and stock availability. The case demonstrates that adoption cannot be reduced to one variable, such as whether the customer has access to the platform. Instead, actual B2B ordering emerges when access, product fit, transaction feasibility, and operational reliability are aligned. This interpretation contributes to the literature by shifting attention from adoption as a binary customer-level decision to adoption as a layered process involving customers, orders, products, and inventory conditions.

Building on this interpretation, the theoretical contribution of this study lies in using the Prysmian case to show how digital channel migration in industrial B2B markets depends on multi-level alignment. Beyond the specific case of Prysmian Spain, the findings suggest that B2B e-commerce adoption in industrial markets should be understood as a multi-level alignment process rather than as a binary adoption decision. Formal platform access is only the first condition for digital migration. Actual usage emerges when customer purchasing routines, product standardization, catalog completeness, order-level feasibility, and stock reliability are simultaneously aligned. This interpretation extends existing adoption research by showing that resistance to digital channels is not only behavioral or technological, but also operational. Customers may fail to migrate online not because they reject the platform, but because the transaction they need to perform is not yet fully supported by the platform’s product, stock, pricing, or service configuration.

### 11.1.2 Customer B2B usage and behavioral adoption

At the customer level, the most visible dimension of this multi-level process is the gap between formal access and actual usage.

The first theoretical implication concerns the distinction between formal access and actual usage. The literature reviewed in Chapter 1 suggests that digital adoption should not be understood as an immediate switch from offline to online channels, but as a gradual process shaped by habits, perceived usefulness, trust, and organizational routines (Yang et al., 2011; Talwar et al., 2021; Ghosh & Dash, 2023). The empirical findings support this view. In the Prysmian case, a customer may be formally selected for the B2B platform but still continue to place most or all orders offline. This distinction is important because it shows that platform access is only a necessary condition for adoption, not a sufficient one. A broader implication is that studies of B2B e-commerce adoption should avoid measuring adoption only through access or registration. In restricted B2B platforms, adoption is better understood as a behavioral outcome that emerges only when customers repeatedly use the digital channel for actual transactions.

The customer-level results in Chapter 5 reinforce this interpretation. The analysis shows that B2B usage is mainly associated with organizational group and, to a lesser extent, prior Electronic E-Commerce exposure. Geography does not emerge as a robust determinant, while customer scale, purchasing breadth, and order complexity should not be presented as the main confirmed drivers in the final model. This suggests that B2B digital adoption is not evenly distributed across the customer base, but is shaped by pre-existing commercial and organizational structures. This finding is consistent with the literature emphasizing that adoption depends on organizational readiness, established routines, and the fit between customer processes and digital service models (Wang & Ahmed, 2009; Ballerini et al., 2023). A possible theoretical interpretation is that customer selection for B2B markets reflects not only technological readiness, but also the supplier's assessment of which customer relationships are sufficiently standardized, scalable, and operationally compatible with self-service ordering. In this sense, adoption is partly co-produced by the customer and the supplier: the customer must be willing and able to use the channel, but the supplier also defines who can use it and under which conditions.

The clustering analysis in Chapter 4 adds a further theoretical insight by showing that B2B customers are not homogeneous. The distinction between B2B Champions, Hidden Giants, and Non-B2B Fit customers suggests that digital adoption depends on the alignment between catalog fit and actual digital usage. B2B Champions represent the expected adoption pattern: high product compatibility and high platform usage. Non-B2B Fit customers illustrate a structural barrier: low usage appears linked to a mismatch between their purchasing needs and the current digital assortment. The most theoretically interesting group is the Hidden Giants segment, because these customers show high compatibility with the existing B2B catalog but limited B2B usage. This pattern suggests that assortment fit alone does not explain adoption. Even when the platform can technically support the customer's purchasing needs, behavioral inertia, internal procurement routines, sales relationships, perceived risk, or lack of habit may still prevent migration to the online channel. This is consistent with innovation resistance theory, which emphasizes that customers may continue relying on familiar channels even when a digital alternative is available (Talwar et al., 2021; Dimitrova et al., 2022).

### 11.1.3 Assortment, product fit, and operational reliability

A second implication concerns assortment completeness as a condition for digital adoption. As Chapter 1 shows, limited assortment and missing products can interrupt the customer journey and reduce online ordering (Bijmolt et al., 2021; Derhami et al., 2021). The order-level and product-level analyses in Chapter 6 and Chapter 7 provide empirical support for this mechanism. Orders containing products outside the B2B catalog are more likely to remain offline, and products included in the B2B catalog are more likely to be standardized and recurrent. This suggests that digital migration in B2B settings depends not only on customer willingness, but also on whether the platform can represent the customer's actual purchasing basket. In theoretical terms, assortment completeness functions as a feasibility condition: if the required products are missing from the digital catalog, adoption cannot occur at the order level, regardless of customer readiness.

However, the findings also show that assortment completeness has limits as an explanation. A substantial share of offline orders placed by enabled customers contains products already available in the B2B catalog. This pattern is theoretically important because it separates structural barriers from behavioral or organizational barriers. When products are missing from the catalog, offline ordering can be explained by technical infeasibility. When products are available, but customers still order offline, the explanation must lie elsewhere: in customer habits, the perceived need for human support, pricing or negotiation issues, urgency, trust, or internal routines. Therefore, the thesis contributes to the literature by showing that catalog completeness is necessary but not sufficient. It removes one barrier to digital ordering, but it does not automatically generate behavioral migration.

The product-level analysis further develops this point by highlighting the role of product standardization. Standardized products are substantially more likely to be included in the B2B catalog, while customized or technically complex products are less likely to be enabled. This finding is consistent with the theoretical expectation that digital self-service channels are better suited to repeatable, well-defined, and low-complexity transactions. In industrial B2B contexts, many products require technical configuration, project-specific adaptation, or commercial negotiation. These characteristics make full digitalization more difficult. A possible theoretical proposition emerging from this result is that product standardization moderates the relationship between platform availability and adoption: the more standardized the product, the more likely it is that digital self-service can substitute for offline interaction; the more customized the product, the more likely it is that offline or assisted channels remain necessary. This interpretation also links the findings to omnichannel literature, which emphasizes that online and offline channels often perform complementary rather than fully substitutive roles (Chou et al., 2025; Rahman et al., 2025).

A third theoretical implication concerns stock availability and operational reliability. Chapter 1 argues that trust and reliability are important drivers of digital adoption, especially when customers must believe that the online channel can support their purchasing needs. The empirical results suggest that stock availability is one of the operational conditions through which reliability becomes visible to customers. At the order level, higher stock availability is associated with a lower likelihood of offline ordering. At the stock level, products with higher turnover and stronger B2B-order intensity tend to show higher in-stock availability, while products with higher total-order intensity and more concentrated repeat demand show lower availability. Taken together, these findings suggest that stock availability may influence digital adoption not only by determining whether a product can be purchased, but also by shaping the perceived dependability of the online channel.

This result extends the literature on B2B adoption by connecting operational reliability to channel choice. In many adoption studies, trust is treated as a psychological or relational construct. In this thesis, trust is not directly measured, but the empirical evidence suggests a possible operational foundation for trust: customers may be more willing to use a digital channel when it reliably shows available products and supports order completion. Conversely, if customers repeatedly encounter unavailable products or uncertainty about fulfillment, they may return to offline channels where sales representatives can provide reassurance, alternatives, or special handling. This does not prove that stockouts cause distrust, but it suggests a plausible mechanism: operational unreliability may reinforce customer resistance by making the digital channel appear less dependable than traditional interaction.

This operational discussion also extends to the characteristics of the order itself; in fact, the order-level findings in Chapter 6 also show that higher-value orders are more likely to remain offline. This result has broader theoretical relevance because it suggests that the economic and relational importance of a transaction may shape channel choice. In B2B markets, larger orders may involve negotiation, customer-specific pricing, credit conditions, delivery coordination, or technical verification. Even when the digital platform exists, customers and suppliers may prefer offline interaction for transactions perceived as commercially sensitive or complex. This finding supports the view that B2B channel migration is selective rather than universal. Digital platforms may first absorb standardized, recurrent, and lower-complexity transactions, while offline channels remain important for high-value or high-complexity exchanges. The broader implication is that successful B2B digitalization may not mean eliminating offline channels but reallocating different transaction types to the channels best suited to handle them.

### 11.1.4 Theoretical propositions and contribution

The matching analysis in Chapter 9 adds an important caution to the theoretical interpretation. Although product-level matching suggests that B2B-enabled products are associated with higher turnover, and customer-level matching shows differences between B2B users and non-B2B users, these results should not be interpreted as definitive causal effects. The observational nature of the data means that B2B usage and catalog inclusion may reflect pre-existing differences rather than the independent effect of the B2B platform. For example, products may be added to the catalog because they are already commercially attractive, standardized, or relevant to digitally active customers. Similarly, customers may be B2B users because they already fit the supplier’s digital strategy. Therefore, the matching results are best understood as evidence of structured associations consistent with the theoretical framework.

Based on the empirical patterns observed in the case, four theoretical propositions can be derived. First, in restricted B2B platforms, formal enablement is unlikely to generate actual adoption unless it is accompanied by catalog compatibility and operational reliability. Second, assortment completeness acts as a structural precondition for online ordering, but it does not eliminate behavioral resistance or relational dependence on offline channels. Third, product standardization moderates the feasibility of digital self-service: standardized products are more likely to migrate online, while customized or high-complexity products preserve the role of assisted channels. Fourth, stock availability may operate as an operational source of trust, because customers are more likely to rely on self-service channels when fulfillment appears dependable. These propositions should be interpreted as theoretically grounded implications of the empirical findings rather than as definitive causal laws.

## 11.2 Managerial implications

The managerial implications should be interpreted as applications of the theoretical mechanism identified in the thesis: B2B actual usage is more likely when customer readiness, catalog feasibility, and operational reliability are jointly aligned. Therefore, the recommendations below are not isolated consulting suggestions, but practical translations of the broader multi-level adoption framework.

A first implication concerns customer prioritization. The customer-level analysis indicates that some customer groups are much more likely to be B2B active users than others, while the segmentation exercise reveals a meaningful group of "Hidden Giants," which are customers with a strong fit for the B2B offer but weak platform usage. These customers should be treated as a priority target because they represent latent digital potential already present in the current base.

A targeted activation plan could focus on explaining platform value, reducing onboarding friction, and monitoring actual usage after enablement. In addition, Prysmian could introduce specific training sessions for both customers and sales teams, provide simple step-by-step guidance materials, and offer regular follow-up meetings to help users become more confident with the platform. Internal trainings for account managers would also be important, so that they can better promote the platform, answer customer questions, and identify the most suitable customers for activation. Over time, these actions could be reinforced with usage reviews, feedback collection, and refresher training to make sure adoption remains consistent.

A second implication concerns the commercial logic of catalog expansion. The product-level results show that standardized and recurring products are more suitable for digital enablement, while more complex or customized items are less likely to be included in the B2B catalog. This suggests that Prysmian should continue prioritizing products with high repeatability, broad demand, and low configuration complexity when expanding the platform. At the same time, highly specialized product families should not be excluded from the strategic discussion, but they may require alternative digital support tools, such as guided configuration, assisted selling, and AI-based assistant tools to support customization requests, rather than full catalog automation. In this way, AI could help customers specify their needs more clearly, suggest suitable product options, and support sales teams in handling more complex or tailored requests.

A third implication relates to stock management. The stock-level analysis shows that availability is strongly associated with turnover and channel choice, meaning that stockouts are associated with substantial offline turnover and may constrain B2B migration. When products are unavailable, customers may revert to manual ordering, which weakens the credibility of the digital channel.

For this reason, Prysmian should treat inventory reliability as a core driver of B2B adoption. If customers repeatedly find that the products they need are unavailable, they are less likely to trust the platform and more likely to return to manual ordering, which weakens digital adoption over time. In practice, this means improving replenishment planning for high-impact products, monitoring out-of-stock risk more closely, and aligning inventory priorities with the products that are most relevant for digital adoption. Prysmian could also introduce more frequent stock reviews for strategic product families and add a feature that allows customers to request an alert when a product is out of stock, so they can be notified as soon as it becomes available again. By doing so, the company would not only reduce lost sales, but also strengthen the reliability of the B2B channel and make it a more credible and attractive option for customers.

Finally, the thesis suggests that Prysmian should adopt a segmented management approach. Not all customers, products, or situations need the same approach. High-potential customers with a good fit for the platform should get support to start using it, fast-selling standard products should be added to the catalog, and important products that often run out should be managed more carefully to avoid stock problems. This more selective approach would allow the company to concentrate resources where they can generate the greatest impact on adoption, turnover, and customer experience.

## 11.3 Limitations

This thesis provides an empirical analysis of the barriers to B2B e-commerce adoption in Prysmian Spain by combining customer-, order-, product-, and stock-level data. However, the findings should be interpreted within several methodological and data-related limitations. These limitations do not invalidate the empirical contribution of the study, but they define the scope of what the analysis can and cannot demonstrate.

First, the study is based on a single-firm and single-country case study. The empirical setting is Prysmian Spain, and the analysis reflects the specific structure of its customer base, product portfolio, B2B platform, internal systems, and commercial organization during the period analyzed. This provides a rich and practically relevant context for studying B2B channel migration, but it limits the external generalizability of the results. The conclusions should therefore be interpreted as evidence from a specific industrial B2B environment rather than as universal findings applicable to all B2B e-commerce platforms. Replication across other countries, business units, product categories, or firms would be necessary to assess whether the same patterns hold in different organizational and market contexts.

Second, the analysis relies on internal operational data extracted from SAP, Salesforce, and related company systems. These data provide detailed information on customers, orders, products, catalog status, turnover, and stock-related indicators, but they also contain some inconsistencies and measurement constraints. For example, the thesis identifies discrepancies between Salesforce and Cockpit regarding the number of active B2B customers and minor mismatches between product lists across systems. Although conservative choices were made to use the most consistent observations, these differences show that the empirical results depend partly on data reconciliation decisions. As a result, the findings should be interpreted as robust operational approximations rather than as a perfectly complete representation of all customer and product activity.

Third, the concept of B2B adoption is only partially observable in the available data. The thesis distinguishes between formal enablement and actual usage, but the available variables cannot capture the full customer adoption process. A customer may be enabled for the platform but never place a B2B order, while another customer may place only one B2B order and still be classified as active. Similarly, the data do not include platform logins, failed ordering attempts, abandoned carts, search behavior, customer satisfaction, training participation, or direct measures of perceived usefulness, perceived ease of use, trust, or resistance to change. Therefore, the analysis can identify observable adoption outcomes, such as B2B orders, B2B turnover share, and offline order persistence, but it cannot directly measure the psychological or experiential reasons behind customer behavior.

Fourth, several theoretical mechanisms discussed in the literature are represented only through indirect proxies. Customer habits and resistance to change are inferred from continued offline ordering, especially among enabled customers whose products are already available in the B2B catalog. Trust and reliability are approximated mainly through stock availability and catalog completeness. Product suitability is proxied through product class and B2B catalog inclusion. These proxies are useful for operational analysis, but they do not fully capture the underlying theoretical constructs. For example, offline ordering may reflect customer habits, but it may also reflect negotiated prices, project-specific requirements, urgent delivery needs, technical assistance, sales-owner influence, or contractual conditions that are not observed in the dataset. Consequently, behavioral explanations should be interpreted as plausible interpretations rather than directly measured causes.

Fifth, the empirical models identify associations rather than definitive causal effects. The logistic regression models at the customer, order, and product levels estimate relationships between observable characteristics and outcomes such as B2B enablement, offline ordering, and catalog inclusion. However, these models are estimated using observational data, and the explanatory variables are not randomly assigned. The results therefore show which factors are associated with B2B-related outcomes, but they do not prove that these factors causally determine adoption. This limitation is particularly important because some variables may be both causes and consequences of digitalization. For example, products included in the B2B catalog may generate more B2B demand because they are available online, but they may also have been selected for inclusion precisely because they were already commercially attractive or operationally suitable.

Sixth, the matching analysis (Chapter 9) improves comparability between B2B-user and non-B2B-user units but does not eliminate all sources of selection bias. Propensity score matching reduces differences between treated and control observations based on observed covariates, but it cannot control for unobserved factors such as sales strategy, customer willingness, procurement maturity, pricing conditions, internal account priorities, product margins, or commercial relationship strength. The remaining imbalance in some matched variables also suggests that the treated and control groups are not perfectly equivalent. For this reason, the matching results should be interpreted as matched associations, or as evidence consistent with possible causal relationships, rather than as definitive causal estimates.

Seventh, the stock analysis (Chapter 8) is subject to important data limitations. Stock information is available only for a restricted subset of products, which reduces the representativeness of the stock-level regression relative to the full product portfolio. In addition, the stock indicators used in the prioritization analysis do not fully overlap with the offline-order dataset. As a result, B2B stock data are used as a proxy for availability conditions affecting offline orders. This is useful for identifying potential stock-risk products, but it does not provide a precise historical measure of stock status at the exact moment when each offline order was placed. Moreover, the classification of products as out of stock is based on Salesforce categories, where “to be ordered” may reflect an actual stockout, but may also indicate a planned make-to-order or nonstocked product policy. Therefore, recommendations related to increasing stock levels should be validated operationally before implementation.

Eighth, some methodological choices are necessarily simplified for analytical and managerial clarity. The K-means segmentation (Chapter 4) uses two main behavioral indicators: B2B turnover share and the share of purchased products available in the B2B catalog. This produces interpretable segments, but it does not capture all dimensions of customer value, profitability, purchasing complexity, or relationship importance. Similarly, thresholds used in the prioritization chapter, such as the 95% catalog-fit threshold for potential customers and the 75% stockout threshold for critical products, are useful screening rules but should not be interpreted as statistically optimal cutoffs. They provide a practical way to identify manageable target groups, but alternative thresholds may lead to different prioritization lists.

Ninth, the analysis focuses primarily on turnover and order behavior, while profitability and implementation costs are not observed. This is relevant for the managerial recommendations. A customer, product, or stock intervention may appear attractive in terms of turnover, but its true business value also depends on margins, inventory carrying costs, service costs, product complexity, pricing agreements, customer-specific discounts, and the cost of onboarding or catalog maintenance. For this reason, the prioritization outputs should be used as screening tools for managerial review rather than as automatic implementation decisions. Before activating customers, adding products to the catalog, or increasing inventory, Prysmian should validate technical feasibility, margin contribution, contractual conditions, and operational cost.

Finally, the thesis does not fully capture qualitative organizational factors that may influence B2B adoption. Although internal documents and informal managerial insights provide useful context, the study does not include a formal interview protocol, customer survey, or systematic qualitative coding process. As a result, factors such as customer training, platform usability, sales-team incentives, internal resistance, perceived channel conflict, and customer trust in the platform cannot be assessed directly. Future research could strengthen the analysis by combining transaction-level data with customer interviews, sales-team interviews, platform usage logs, and longitudinal adoption data.

Overall, these limitations mean that the thesis should not be read as proving that actual B2B usage, catalog inclusion, or stock availability alone cause B2B adoption. Rather, the contribution of the thesis is to show, using detailed internal data, where the strongest observable frictions are located and how they are associated with continued reliance on offline channels. The analysis is therefore most valuable as a diagnostic and prioritization framework. It provides evidence-based guidance for identifying high-potential customers, relevant catalog gaps, and stock-risk products, while recognizing that final managerial decisions require additional validation from commercial, operational, and financial perspectives.

# Conclusion

This thesis set out to understand why B2B e-commerce adoption remains partial in Prysmian Spain and which levers can help increase the use of the digital B2B channel. The evidence shows that adoption is not explained by a single factor, but by the interaction between B2B usage, product assortment, and stock availability. More specifically, customer decisions to order through B2B depend not only on whether customers are B2B users, but also on whether the platform fits their purchasing behavior and can support their actual needs.

The literature showed that the shift from offline to online channels is shaped by both barriers and drivers. Customers may resist digital channels because of established routines, trust issues, product mismatch, or operational uncertainty, while adoption is more likely when the platform is convenient, complete, reliable, and well aligned with customer requirements. This framework was useful for interpreting the empirical results, because it made clear that B2B migration should be understood as a gradual behavioral process rather than an immediate switch from one channel to another.

The empirical analysis confirmed that Prysmian Spain already has an active B2B platform, but actual usage is still limited. The descriptive analysis showed a clear gap between enablement and usage: many customers are enabled, but only a smaller share actively places orders through B2B. A similar pattern appears at product level, where the catalog covers only part of the assortment sold in Spain, and at order level, where offline orders remain dominant. These results suggest that the platform has potential, but that its current coverage and adoption are still insufficient to substantially reduce offline ordering.

The following tables, Table C.1 and Table C.2, provide an explicit assessment of the research objectives and hypotheses, linking each of them to the empirical findings and to the resulting managerial implications.

| Objective | Assessment and main findings | Managerial implications |
|---|---|---|
| O1: Measure the gap between B2B platform enablement and actual B2B usage. | **Achieved.** The analysis shows a clear gap between formal enablement and actual platform usage. Although a relevant number of customers are enabled for the B2B platform, only a smaller subset actively places orders through it, and B2B still accounts for a limited share of total orders and turnover. | Industrial B2B companies should not measure adoption only through platform access. They should monitor the full adoption funnel, including enablement, activation, recurrent usage, B2B turnover share, and offline persistence. |
| O2: Segment B2B-enabled customers according to digital purchasing behavior and catalog compatibility. | **Achieved.** The k-means segmentation identified three customer groups: B2B Champions, Hidden Giants, and Non-B2B Fit customers. This shows that enabled customers are not homogeneous and that high catalog compatibility does not automatically translate into actual B2B usage. | The Hidden Giants segment should be prioritized because these customers already show strong compatibility with the existing catalog but still require onboarding, training, commercial follow-up, or behavioral activation to increase digital usage. |
| O3: Identify order-level factors associated with continued offline ordering among enabled customers. | **Achieved.** The order-level analysis shows that higher order value is associated with a greater likelihood of offline ordering, while higher stock availability is associated with a lower likelihood of offline ordering. This indicates that offline ordering is not explained only by customer resistance, but also by transaction characteristics and operational feasibility. | Companies should improve the platform's ability to support high-value orders through customer-specific pricing, assisted checkout, validation workflows, or dedicated commercial support. At the same time, they should improve stock reliability for products frequently involved in B2B-relevant transactions. |
| O4: Examine product characteristics associated with B2B catalog inclusion and identify high-demand offline products for catalog expansion. | **Achieved.** The product-level analysis shows that standardized products are much more likely to be included in the B2B catalog, while customized or project-based products are less compatible with the current self-service model. The prioritization analysis also identified offline products purchased by enabled customers that could be considered for future catalog expansion. | Catalog expansion should prioritize standardized, recurrent, and commercially relevant products. More complex products should not necessarily be forced into full self-service ordering, but may require assisted digital tools, guided configuration, or sales-supported workflows. |
| O5: Assess the role of stock availability and stock-related constraints in B2B ordering behavior. | **Achieved.** The results indicate that stock availability is an important operational condition for digital adoption. Lower stock availability is associated with greater offline ordering, and several high-turnover products show recurrent stock constraints. | Stock allocation should be treated as part of the B2B adoption strategy. Products that are central to digital demand should receive closer replenishment monitoring, because stockouts may redirect customers toward manual channels and weaken trust in the platform. |
| O6: Develop a data-driven prioritization framework for customer activation, catalog expansion, and stock allocation. | **Achieved.** The thesis translates the empirical findings into a prioritization framework that identifies high-potential non-enabled customers, enabled customers with persistent offline behavior, non-B2B products with high offline demand, and B2B products affected by stock constraints. | B2B growth should be pursued through coordinated prioritization. Companies should activate the right customers, enable the right products, and ensure the availability of products that matter most for digital transactions. |

*Table C.1 - Assessment of research objectives, summarizing the achievement of each objective, the corresponding empirical evidence, and the main managerial implications for B2B e-commerce adoption.*

| Hypothesis | Assessment | Interpretation |
|---|---|---|
| H1: Higher catalog compatibility is positively associated with actual B2B platform usage. | Partially supported. | Catalog compatibility is important, as shown by the Hidden Giants segment and by the identification of non-enabled customers whose purchases are already highly compatible with the B2B catalog. However, compatibility alone is not sufficient, since many enabled customers still place offline orders even when the products are already available online. |
| H2: Higher stock availability is associated with a lower probability of offline ordering among B2B-enabled customers. | Supported. | The order-level results show that higher stock availability is associated with a lower likelihood of offline ordering. This confirms that inventory reliability is not only an operational issue, but also a condition that affects the credibility and usability of the B2B channel. |
| H3: Standardized and recurrently purchased products are more likely to be included in the B2B catalog. | Partially supported. | The product-level analysis strongly confirms the importance of standardization, since Standard & Installers products are substantially more likely to be included in the B2B catalog. However, order frequency and other demand-intensity variables are less consistently significant, suggesting that standardization is more important than frequency alone. |
| H4: Customers with stronger prior digital exposure and more digitally aligned organizational contexts are more likely to become active B2B users. | Partially supported. | Prior electronic ordering behavior is positively associated with B2B activity, but the marginal effect is small. Organizational affiliation explains a substantial part of the variation in B2B usage, suggesting that digital readiness is partly behavioral but also embedded in customer organization and procurement routines. |
| H5: Higher-value orders are more likely to be placed through offline channels. | Supported. | The order-level model shows that higher-value orders are more likely to remain offline. This suggests that economically important transactions may require negotiation, verification, pricing support, or additional coordination, which are not yet fully supported by the current B2B self-service logic. |

*Table C.2 - Assessment of research hypotheses, summarizing the level of empirical support for each hypothesis and the interpretation of the results in relation to B2B e-commerce adoption.*

Overall, the results show that B2B e-commerce adoption in industrial markets is not determined by platform access alone. Actual usage emerges when customer readiness, catalog compatibility, product standardization, transaction feasibility, and stock reliability are jointly aligned. This confirms that B2B adoption should be interpreted as a multi-level process rather than as a simple binary distinction between enabled and non-enabled customers.

From a practical perspective, the findings suggest that industrial B2B companies should prioritize three coordinated actions. First, they should activate high-potential customers whose purchasing behavior is already compatible with the digital catalog. Second, they should expand the catalog toward standardized and recurrently purchased products with strong offline demand. Third, they should improve stock reliability for products that are central to B2B transactions, since stockouts may reduce customer trust and redirect demand toward manual channels.

Future research could extend the analysis to longer time periods, additional countries, or other industrial B2B companies to test whether the same patterns hold in different contexts. Further research could also integrate platform usage logs, customer interviews, and sales-team perspectives to better capture behavioral barriers such as trust, habits, perceived usefulness, and resistance to change.

The broader contribution of this thesis is to show that B2B e-commerce adoption in industrial markets depends on whether the digital channel can support the actual transaction that the customer needs to perform. Successful digitalization requires not only enabling customers, but also making the platform operationally reliable, commercially suitable, and complete.

## References

Amornkitvikai, Y., Tham, S. Y., Harvie, C., & Buachoom, W. W. (2022). Barriers and factors affecting the e-commerce sustainability of Thai micro-, small- and medium-sized enterprises (MSMEs). Sustainability, 14(14), Article 8476.

Anand, D., Crecelius, A. T., Lawrence, J. M., & Scheer, L. K. (2025). A potential win-win-win strategy when pruning the direct channel: Motivating retailers to migrate to a designated intermediary. Journal of Retailing, 101(1), 25-39.

Ballerini, J., Herhausen, D., & Ferraris, A. (2023). How commitment and platform adoption drive the e-commerce performance of SMEs: A mixed-method inquiry into e-commerce affordances. International Journal of Information Management, 72, Article 102649.

Beatty, P., Reay, I., Dick, S., & Miller, J. (2007). P3P adoption on e-commerce web sites: A survey and analysis. IEEE Internet Computing, 11(2), 65-71.

Bian, Z., & Che, C. (2025). How AI overview of customer reviews influences consumer perceptions in e-commerce? Journal of Theoretical and Applied Electronic Commerce Research, 20(4), Article 315.

Bijmolt, T. H. A., Broekhuis, M., de Leeuw, S., Hirche, C., Rooderkerk, R. P., Sousa, R., & Zhu, S. X. (2021). Challenges at the marketing–operations interface in omni-channel retail environments. Journal of Business Research, 122, 864-874.

Boom-Cárcamo, E., Molina-Romero, S., Galindo-Angulo, C., & del Mar Restrepo, M. (2024). Barriers and strategies for digital marketing and smart delivery in urban courier companies in developing countries. Journal of the Knowledge Economy, 15(4), 19203-19232.

Bravo, R., Gonzalez Segura, M., Temowo, O., & Samaddar, S. (2022). How does a pandemic disrupt the benefits of e-commerce? A case study of small and medium enterprises in the US. Journal of Theoretical and Applied Electronic Commerce Research, 17(2), 522-557.

Chou, T.-H., Hung, Y. W., & Wang, Y.-C. (2025). Understanding the effects of service experience in online-to-offline services. Pacific Asia Journal of the Association for Information Systems, 17(2), Article 9.

Costa, P., & Rodrigues, H. (2024). The ever-changing business of e-commerce: Net benefits while designing a new platform for small companies. Review of Managerial Science, 18(9), 2507-2545.

Derhami, S., Montreuil, B., & Bau, G. (2021). Assessing product availability in omnichannel retail networks in the presence of on-demand inventory transshipment and product substitution. Omega, 102, Article 102315.

Dimitrova, I., Öhman, P., & Yazdanfar, D. (2022). Barriers to bank customers' intention to fully adopt digital payment methods. International Journal of Quality and Service Sciences, 14(5), 16-36.

Evans, C. (2023, October 25). Customer digital service development and channel migration foundational framework. IBRS.

Gao, J., Opute, A. P., Jawad, C., & Zhan, M. (2025). The influence of artificial intelligence chatbot problem solving on customers' continued usage intention in e-commerce platforms: An expectation-confirmation model approach. Journal of Business Research, 200, Article 115661.

George, A., Panattil, S. J., & Sajid, M. (2025). Beyond clicks and mortar: Exploring the influence of experience value on willingness to adopt phygital services. Services Marketing Quarterly. Advance online publication.

Ghosh, D., & Dash, S. (2023). Barriers and facilitators of B2B degree of digital use and brand engagement: An integration of technology and behavioral perspectives. Journal of Business & Industrial Marketing, 38(12), 2793-2810.

Hansen, B. E. (2022). A modern Gauss–Markov theorem. Econometrica, 90(3), 1283-1294.

Jain, A. K. (2010). Data clustering: 50 years beyond K-means. Pattern Recognition Letters, 31(8), 651-666.

Leszczensky, L., & Wolbring, T. (2022). How to deal with reverse causality using panel data? Recommendations for researchers based on a simulation study. Sociological Methods & Research, 51(2), 837-865.

Naunthong, P. (2022). Understanding characteristics of organizations for e-commerce adoption in the new normal era. In 2022 7th International Conference on Business and Industrial Research (ICBIR) (pp. 353-358). IEEE.

Papastamoulou, P., & Antonopoulos, N. (2025). Artificial intelligence in e-commerce: A comparative analysis of best practices across leading platforms. Systems, 13(9), Article 746.

Pérez Trujillo, M., & Aleán-Romero, A. (2024). The impact of ICT and e-commerce on employment in Colombia. Electronic Commerce Research and Applications, 65, Article 101378.

Plekhanov, D., Franke, H., & Netland, T. H. (2023). Digital transformation: A review and research agenda. European Management Journal, 41(6), 821-844.

Rahman, S. M., Carlson, J., Gudergan, S. P., Wetzels, M., & Grewal, D. (2025). How do omnichannel customer experiences affect customer engagement? Theory and empirical validation. Journal of Business Research, 189, Article 115196.

Rosales, C., Whipple, J. M., & Blackhurst, J. (2018). The impact of out-of-stocks and supply chain design on manufacturers: Insights from an agent-based model. Transportation Journal, 57(2), 137-162.

Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. Journal of Computational and Applied Mathematics, 20(1), 53-65.

Safaria, T., & Nugroho, B. A. (2024). Comparison of OLS Regression and Robust Regression Methods in Handling Outliers. Zero Educational Research Journal, 2(2), 112-121.

Sell, A., Walden, P., Jeansson, J., Lundqvist, S., & Marcusson, L. (2019). Go digital: B2C microenterprise channel expansions. Journal of Electronic Commerce Research, 20(2), 75-90.

Sharma, M., Gupta, M., & Joshi, S. (2020). Adoption barriers in engaging young consumers in omni-channel retailing. Young Consumers, 21(2), 193-210.

Shivakumar, S. K., & Suresh, P. V. (2014). Maximizing knowledge management returns in e-commerce. In 2014 International Conference on Computing for Sustainable Global Development (INDIACom) (pp. 545-550). IEEE.

Sivakumar, M., Parthasarathy, S., & Padmapriya, T. (2024). Trade-off between training and testing ratio in machine learning for medical image processing. PeerJ Computer Science, 10, Article e2245.

Sperandei, S. (2014). Understanding logistic regression analysis. Biochemia Medica, 24(1), 12-18.

Styvén, M. E., & Wallström, Å. (2019). Benefits and barriers for the use of digital channels among small tourism companies. Scandinavian Journal of Hospitality and Tourism, 19(1), 27-46.

Sujon, K. M., Hassan, R., Choi, K., & Samad, M. A. (2025). Accuracy, precision, recall, F1-score, or MCC? Empirical evidence from advanced statistics, ML, and XAI for evaluating business predictive models. Journal of Big Data, 12, Article 268.

Szumilas, M. (2010). Explaining odds ratios. Journal of the Canadian Academy of Child and Adolescent Psychiatry, 19(3), 227-229.

Talwar, S., Dhir, A., Scuotto, V., & Kaur, P. (2021). Barriers and paradoxical recommendation behaviour in online to offline (O2O) services: A convergent mixed-method study. Journal of Business Research, 131, 25-39.

Tchao, E. T., Diawuo, K., Aggor, C. S., & Kotey, S. D. (2017). Ghanaian consumers' online privacy concerns: Causes and its effects on e-commerce adoption. International Journal of Advanced Computer Science and Applications, 8(11).

Vrigazova, B. (2021). The proportion for splitting data into training and test set for the bootstrap in classification problems. Business Systems Research, 12(1), 228-242.

Wang, Y., & Ahmed, P. K. (2009). The moderating effect of the business strategic orientation on eCommerce adoption: Evidence from UK family run SMEs. The Journal of Strategic Information Systems, 18(1), 16-30.

Xiao, J., Xu, Z., Xiao, A., Wang, X., & Skare, M. (2024). Overcoming barriers and seizing opportunities in the innovative adoption of next-generation digital technologies. Journal of Innovation & Knowledge, 9(4), Article 100622.

Yang, S., Lu, Y., Zhao, L., & Gupta, S. (2011). Empirical investigation of customers' channel extension behavior: Perceptions shift toward the online channel. Computers in Human Behavior, 27(5), 1688-1696.

# Acknowledgments

Desidero ringraziare il Professor Rullani, il mio relatore, per la grande disponibilità dimostrata in ogni momento, a qualsiasi ora e in qualsiasi giorno. Il suo supporto costante e le energie che mi ha dedicato sono stati fondamentali per portare a termine questo lavoro.

Alle persone che hanno reso possibile la realizzazione di questa tesi, Bea, Maria Vittoria e Maria Teresa. Grazie per la fiducia che mi avete dato scegliendomi per lavorare con voi, per il supporto e la formazione che mi avete offerto durante questi mesi in Prysmian, per la pazienza e per il tempo che mi avete permesso di dedicare a questa tesi. Grazie perché mi avete fatta crescere tanto, sia dal punto di vista professionale che personale.

Alla mia famiglia, mamma, papà e Marghe, per avermi sempre sostenuta in ogni scelta, senza mai farmi mancare il vostro supporto. Grazie per avermi dato la possibilità di vivere l'esperienza Erasmus a Vienna, di studiare a Venezia per seguire il percorso che desideravo di più (anche se ancora non sapete come si chiama il corso in cui mi sto laureando), e di iniziare a dare forma al mio futuro a Milano. Grazie perché, anche se non è stato facile vedermi andare lontana (soprattutto per mamma e papà), mi avete sempre incoraggiata a farlo. Grazie per aver sempre messo il mio bene davanti a tutto, vi voglio bene.

Ai miei nonni, Marisa e Alberto, Ofelia e Bruno, i miei angeli custodi. Nel vostro piccolo mi avete sempre aiutata in questi anni di studi. Ogni volta che vi chiedevo di pregare per me prima di un esame, riuscivate a darmi forza e sicurezza. Avete reso le mie domeniche "casa" quando ero lontana, con il ragù, la crostata e i tortelli che avrei mangiato se fossi stata vicina a voi.

Alle mie amiche di Imola, in particolare Anna e Meggy. Siete sempre state casa, sia quando ero (e sono) lontana, sia quando eravate lontane voi. A volte vorrei avervi più vicine, per passare insieme le serate di quelle giornate no e per esserci nelle vostre; poi però mi rendo conto che, nonostante la distanza fisica, il nostro legame non cambia, anzi si rafforza sempre di più. Grazie per rispondere sempre alle mie chiamate improvvise, ormai ogni volta che vi chiamo pensate che sia in pericolo di vita…ma sto migliorando, promesso.

Alle mie amiche di Venezia, in particolare Salma, Marti, Fra, Benny, Sofi, Auri e Lisa. Tra una sessione e l'altra, tra risate e momenti di panico, avete reso tutto più leggero. Insieme riuscivamo a sdrammatizzare anche gli esami più pesanti e a trasformare i lavori di gruppo, che prima odiavo profondamente, in momenti divertenti. Grazie per tutte le ore passate insieme in biblioteca e per aver condiviso con me la vita a Venezia, fatta di pranzi e cene al Parlamento, di caffè al Gray, di spritz nei bacari e di tutti quei ricordi che porterò sempre con me. E soprattutto, grazie per l'amicizia che abbiamo costruito, che sono sicura porteremo con noi anche oltre Venezia.

Ao Tiago, a minha rocha, o meu porto seguro. Obrigada pela paciência que tiveste e por estares sempre presente. És um amigo, um namorado e um exemplo para mim. Obrigada por me incentivares a ir sempre mais longe, por me dares força nos momentos mais difíceis e por acreditares em mim, mesmo quando eu própria tenho dificuldade em fazê-lo. Já passámos por tantas fases e dificuldades, mas, apesar de tudo, hoje estamos aqui, finalmente perto um do outro, a construir o nosso futuro juntos. Sinto-me mesmo grata por te ter encontrado e por te ter na minha vida…talvez tenha sido destino acabar naquela residência, Burgenland 2, no mesmo terceiro andar…obrigada ao destino por me ter levado até lá.

Infine, un grazie speciale a tutte le persone che, in un modo o nell'altro, hanno fatto parte del mio percorso accademico. La vostra presenza ha contribuito a rendermi la persona che sono oggi, accompagnando la mia crescita sia personale che universitaria.

Chiara

# Dichiarazione del ricorso a strumenti di Intelligenza Artificiale (IA)

La sottoscritta Chiara Golinelli, matricola numero 905996, iscritta presso Ca' Foscari al corso di Laurea magistrale in Data Analytics for Business and Society, autrice della tesi di laurea "Customer, Product, and Stock Barriers to B2B E-Commerce Adoption: Evidence from Prysmian Spain's Digital Platform", relatore Francesco Rullani, consapevole delle sanzioni previste dall'art. 76 del Testo unico, D.P.R. 28/12/2000 n.445, e della decadenza dei benefici prevista dall'art. 75 del medesimo Testo unico in caso di dichiarazioni false o mendaci, sotto la propria personale responsabilità ai sensi degli artt. 46 e 47 del D.P.R. 28/12/2000 n. 445,

**DICHIARA**

di aver fatto ricorso a tecnologie di intelligenza artificiale generativa per supportare vari aspetti della ricerca e della scrittura della mia tesi di laurea.

In particolare, ho utilizzato i seguenti strumenti:

- ChatGPT Enterprise, versione utilizzata nel periodo di redazione della tesi, per la revisione e l'ottimizzazione di alcune parti testuali, al fine di migliorarne chiarezza, coerenza e fluidità, nonché per la riformulazione di frasi e paragrafi già elaborati da me.
- NotebookLM, impiegato per sintetizzare e riassumere le fonti consultate, nonché per supportare l'organizzazione della bibliografia e la gestione delle informazioni provenienti dalla letteratura.
- Copilot 365, utilizzato come supporto alla scrittura del codice per le analisi quantitative.

L'utilizzo di tali strumenti ha riguardato in particolare le seguenti sezioni/capitoli dell'elaborato: revisione linguistica e stilistica di tutta la tesi, revisione della parte metodologica, supporto alla scrittura e correzione del codice per le analisi, sintesi delle fonti e predisposizione della bibliografia.

La correttezza, la qualità e l'attendibilità dei contenuti generati o rielaborati con l'ausilio di tali strumenti sono state verificate personalmente mediante confronto con le fonti originali, controllo incrociato con la letteratura scientifica e revisione critica dei testi e dei risultati prodotti. In particolare, ogni informazione inserita nell'elaborato è stata validata sulla base delle fonti primarie consultate e dei risultati delle analisi effettuate, assumendo piena responsabilità del contenuto finale presentato.

**DICHIARA INOLTRE**

che tutto il materiale generato dall'IA è stato attentamente verificato, rivisto e rielaborato dall'autrice, che si assume la piena responsabilità per la qualità, l'accuratezza e l'integrità scientifica del lavoro.