

Università degli Studi di Padova
Dipartimento di Scienze Statistiche
Corso di Laurea Triennale in
Statistica per l'Economia e l'Impresa



RELAZIONE FINALE

**L'EFFETTO DEGLI SCONTI NEL RETAIL:
EVIDENZE DA DATI PANEL**

Relatore Prof. Adriano Paggiaro
Dipartimento di Scienze Statistiche

Laureando: Christian Capitanio
Matricola N2033253

Anno Accademico 2025/2026

Indice:

1. Introduzione	7
1.1 Contesto di ricerca	7
1.2 Obiettivo della tesi	8
2. Base dati e variabili di analisi	11
2.1 Struttura del dataset	11
2.2 Qualità dei dati e struttura delle osservazioni mancanti	12
2.2.1 Entità del fenomeno	13
2.2.2 Chiusure per inventario	13
2.2.3 Chiusure strutturali domenicali	13
2.2.4 Caso di store con pattern misti di missing	15
2.2.5 Implicazioni econometriche	16
2.2.6 Conclusione sulla qualità dei dati	16
2.3 Descrizione delle variabili	17
2.3.1 KPI principali	17
2.3.2 KPI derivati	17
2.3.3 Variabili strutturali	18
2.3.4 Variabili temporali	19
2.3.5 Variabili promozionali e attività commerciali	19
3. Analisi descrittive	21
3.1 Struttura del dataset e KPI principali	21
3.1.1 Log-trasformazione e outlier	23
3.2 Struttura temporale della politica promozionale	25
3.3 Eterogeneità nelle osservazioni	26
3.3.1 Tipologia di store	26
3.3.2 Area geografica	28
3.3.3 Stagionalità: feriale vs weekend	28
3.4 Relazioni bivariate tra KPI	29
3.5 Sintesi e implicazioni per l'analisi econometrica	30
4. Metodologia	33
4.1 Approccio generale	33
4.2 Modello base e interpretazione dell'elasticità	33
4.3 Modelli con effetti fissi	34
4.4 Analisi dell'eterogeneità	35
4.5 Specifica degli errori standard	37

4.6 Specificazione 2WFE con trattamento binario	37
4.7 Regression Discontinuity Design.....	38
4.8 Limiti dell'analisi	39
5. Elasticità allo sconto	41
5.1 Elasticità media, OLS.....	41
5.1.1 Risultati	41
5.1.2 Interpretazione	42
5.1.3 Limiti dell'analisi	43
5.2 Elasticità media, 2WFE.....	43
5.2.1 Risultati	44
5.2.2 Interpretazione	45
5.3 Eterogeneità dell'elasticità per contesto strutturale	47
5.3.1 Struttura del campione.....	47
5.3.2 Risultati	48
5.3.3 Discussione di un risultato anomalo.....	53
5.3.4 Interpretazione	54
5.4 Eterogeneità dell'elasticità per contesto promozionale	56
5.4.1 Impostazione metodologica	57
5.4.2 Variazione identificante e qualità dell'identificazione.....	59
5.4.3 Statistiche descrittive per regime	60
5.4.4 Risultati elasticità within-regime	60
5.4.5 Scala dell'effetto: un esercizio illustrativo	63
5.4.6 Interpretazione	64
6. Analisi causale di una campagna promozionale	67
6.1 Contesto e strategia di identificazione	68
6.2 Analisi preliminare: modello OLS.....	71
6.2.1 Risultati	71
6.2.2 Interpretazione	72
6.2.3 Limiti interpretativi	72
6.3 Modello con effetti fissi	73
6.3.1 Risultati	73
6.3.2 Interpretazione	74
6.3.3 Confronto OLS vs 2WFE	75
6.3.4 Limiti dell'approccio con effetti fissi	76
6.4 Regression Discontinuity Design.....	77
6.4.1 Specificazione del modello	78

6.4.2 Risultati	82
6.4.3 Verifica di robustezza	83
6.4.4 Interpretazione	86
6.5 Cosa dicono i tre modelli insieme.....	87
7. Conclusioni	89
7.1 Sintesi dei risultati principali.....	89
7.2 Implicazioni operative.....	90
7.3 Limiti e direzioni di ricerca futura	92
7.4 Riflessione conclusiva	93
Bibliografia.....	95
Appendice.....	97

1. Introduzione

1.1 Contesto di ricerca

La presente tesi si inserisce nell'esperienza di stage svolta presso un operatore internazionale del settore retail fashion, con più di un secolo di storia e una rete di oltre 5.000 punti vendita a livello globale. L'azienda si colloca nel segmento mass market, con un posizionamento orientato a un equilibrio tra accessibilità del prezzo e qualità degli articoli venduti. Il modello distributivo è caratterizzato da una rete di punti vendita fisici oltre che online, integrata da attività promozionali e commerciali coordinate centralmente a livello europeo. La rete commerciale comprende differenti tipologie di store, tra cui punti vendita situati in centri commerciali, negozi in centri urbani strategici e outlet orientati alla leva promozionale. Il contesto analizzato è contraddistinto da un'elevata intensità promozionale: lo sconto rappresenta una delle principali leve utilizzate per influenzare la domanda, stimolare il traffico nei punti vendita e gestire la rotazione dei prodotti lungo il calendario commerciale. Tuttavia, proprio per la sua diffusione, lo sconto introduce un trade-off rilevante: se da un lato può incrementare le vendite, dall'altro può comportare una riduzione del valore netto del fatturato. In questo scenario, comprendere quando e dove le promozioni risultino realmente efficaci diventa un problema centrale per il business e una sfida empirica non banale, poiché le promozioni non vengono attivate in modo casuale ma rispondono a logiche commerciali che rendono difficile separare l'effetto causale dalla selezione.

All'interno di questo contesto, il team CRM (Customer Relationship Management) della divisione Marketing Europa svolge un ruolo chiave nella progettazione e nel monitoraggio delle attività promozionali. Le decisioni relative all'intensità e al timing degli sconti vengono prese su base settimanale e si basano su analisi quantitative delle performance osservate nei diversi mercati.

Durante il periodo di stage, il ruolo ricoperto è stato quello di Data Analyst all'interno del team CRM, con responsabilità legate alla progettazione, testing e monitoraggio delle campagne promozionali a livello europeo oltre alla produzione di analisi statistiche a supporto del processo decisionale settimanale. In particolare, le analisi condotte avevano l'obiettivo di tradurre i dati operativi in indicazioni utili per la pianificazione commerciale, la valutazione di campagne e iniziative aziendali. Le dinamiche promozionali si riflettono infatti direttamente nei principali indicatori osservabili nei dati, come il numero di transazioni e il valore medio dello

scontrino, rendendo possibile una valutazione quantitativa dell'efficacia delle politiche di sconto. Sebbene il contesto operativo del team fosse europeo, l'analisi empirica della tesi si concentra esclusivamente sul mercato italiano. Questa scelta è motivata dalla necessità di lavorare su un contesto commerciale relativamente omogeneo per quadro normativo, offerta di prodotti e servizi commerciali, riducendo l'eterogeneità derivante da differenze strutturali tra paesi. Inoltre, il mercato italiano rappresentava il contesto con la maggiore disponibilità di dati operativi, oltre a garantire una maggiore uniformità nella definizione delle variabili osservabili e una qualità complessivamente più elevata e affidabile del dato, rendendo possibile un'analisi panel dettagliata a livello giornaliero e di singolo punto vendita.

1.2 Obiettivo della tesi

L'obiettivo della tesi è quantificare l'impatto delle politiche promozionali sulle performance di vendita, con particolare attenzione al ruolo dello sconto come leva di stimolo della domanda. In contesti ad alta intensità promozionale emerge una questione centrale: lo sconto genera realmente vendite incremental? E soprattutto, l'efficacia delle promozioni è uniforme o varia sistematicamente in funzione del contesto operativo: area geografica, tipologia di store, periodo della settimana?

Per rispondere a queste domande, l'analisi si sviluppa lungo tre direttrici principali. La prima è la stima dell'elasticità della domanda rispetto allo sconto attraverso un modello panel con effetti fissi a due vie, che consente di controllare per l'eterogeneità non osservata tra negozi e per gli shock temporali comuni, isolando la componente di variazione dello sconto informativa sull'effetto causale. La seconda è l'analisi dell'eterogeneità di tale elasticità in funzione di variabili operative rilevanti (area geografica, tipologia di store e giorno della settimana) attraverso un sistema di interazioni che permette di caratterizzare per quali combinazioni di contesto lo sconto risulta più o meno efficace. La terza è la valutazione causale dell'impatto di specifiche campagne promozionali attraverso una progressione metodologica che va dall'OLS al modello a effetti fissi fino al Regression Discontinuity Design, sfruttando la discontinuità temporale nell'avvio delle promozioni per identificare un effetto causale locale confrontando osservazioni immediatamente adiacenti al cutoff.

Dal punto di vista metodologico, l'approccio adottato riflette le indicazioni della letteratura econometrica sull'identificazione causale in contesti osservativi. I modelli panel con effetti fissi

controllano per le fonti di eterogeneità più rilevanti, mentre il disegno quasi-sperimentale consente di avvicinarsi a un'identificazione causale senza richiedere la randomizzazione del trattamento. È importante sottolineare che un aumento delle vendite non implica necessariamente un miglioramento della redditività: gli sconti comportano una riduzione del margine unitario, e i risultati devono essere interpretati come effetti sulla domanda, non direttamente come misure di profittabilità.

Il contributo della tesi è quindi duplice: da un lato quantificare in modo rigoroso l'efficacia delle promozioni in un contesto retail, dall'altro identificare i contesti operativi in cui lo sconto risulta uno strumento efficace o inefficiente, fornendo indicazioni utili per una gestione più mirata delle politiche promozionali.

2. Base dati e variabili di analisi

2.1 Struttura del dataset

Il dataset utilizzato per l'analisi è costruito a partire dai dati operativi aziendali e si riferisce esclusivamente ai punti vendita fisici operanti sul territorio italiano. Sono esclusi sia il canale online sia le vendite relative ad articoli successivamente oggetto di reso. La struttura è di tipo panel a livello store-giorno, con un orizzonte temporale che copre l'intero anno solare 2025. L'unità di osservazione è la combinazione tra punto vendita e giorno, che consente di seguire l'evoluzione delle performance nel tempo per ciascuno store e di sfruttare la variazione longitudinale per l'identificazione econometrica.

Il dataset copre complessivamente 359 giorni dell'anno solare, poiché durante alcune delle principali festività nazionali, in particolare: Natale, Capodanno, Pasqua, Pasquetta, festa dei Lavoratori e Ferragosto, tutti gli store risultano sistematicamente chiusi. Tale riduzione non costituisce una perdita informativa preoccupante, in quanto riflette una chiusura uniforme e generalizzata della rete commerciale, che non introduce distorsioni nella struttura del campione né problemi di selezione nelle osservazioni disponibili.

Il dataset completo include 64.013 osservazioni distribuite su 216 store e 359 giorni, con una struttura di panel non bilanciato dovuta alla presenza di aperture e chiusure di punti vendita nel corso dell'anno. La distribuzione del numero di osservazioni per store riflette questa eterogeneità: il minimo è di 11 osservazioni, il primo quartile si colloca a 304, la mediana a 358 e la media a 296, con un massimo di 359. La distanza tra mediana e media segnala la presenza di un gruppo di store attivi solo per una frazione dell'anno, che sposta verso il basso la media senza essere rappresentativa della situazione tipica del campione.

Per le principali analisi econometriche è stato considerato un sottoinsieme di punti vendita Like-for-Like (L4L), definito come l'insieme degli store attivi in modo continuativo durante l'intero periodo di osservazione. La scelta del campione L4L risponde a un'esigenza metodologica precisa: includere store con storie osservative incomplete introdurrebbe distorsioni legate all'ingresso o all'uscita dal mercato, che potrebbero essere sistematicamente correlate con le variabili di interesse. Il campione L4L comprende 165 store per un totale di 58.181 osservazioni. La distribuzione del numero di osservazioni per store è in questo caso significativamente più regolare: il minimo è di 282 osservazioni, il primo quartile di 357, la mediana di 358 e la media di 352, con un massimo di 359. L'elevata concentrazione attorno alla

mediana conferma che gli store L4L sono sostanzialmente attivi per l'intero anno, con scarti minimi rispetto al massimo teorico di 359 giorni osservati.

È opportuno segnalare che la selezione del campione L4L introduce una forma di survivorship bias: gli store attivi per l'intero anno sono, per definizione, quelli che non hanno subito chiusure nel periodo di osservazione, possono quindi essere sistematicamente più performanti o strutturalmente più stabili rispetto ai punti vendita marginali esclusi dall'analisi. Le stime di elasticità ottenute su questo campione devono pertanto essere interpretate come riferite agli store a regime, e non necessariamente generalizzabili all'universo completo dei punti vendita, inclusi quelli di nuova apertura o prossimi alla chiusura. Il campione L4L costituisce il riferimento per tutte le analisi econometriche presentate nei capitoli successivi: la sua adozione garantisce che la variazione osservata nei KPI rifletta dinamiche operative genuine e non artefatti legati alla composizione mutevole del campione.

2.2 Qualità dei dati e struttura delle osservazioni mancanti

Dal punto di vista della qualità dei dati, non sono presenti duplicazioni a livello store-giorno e non si riscontrano valori mancanti rilevanti nelle variabili principali, tuttavia si nota che, come mostrato nella Tabella 2.1, è presente un numero di osservazioni store-giorno mancanti rispetto al caso teorico di panel bilanciato ($359 \times 165 = 59235$ osservazioni teoriche contro le 58181 osservazioni attuali), pertanto 1054 osservazioni risultano mancanti (1,77% del campione teorico).

Tabella 2.1 - Struttura delle osservazioni nel campione L4L

Indicatore	Valore
Osservazioni teoriche	59.235
Osservazioni presenti nel dataset	58.181
Osservazioni mancanti	1.054
Osservazioni mancanti su totale teorico (%)	1,77%

Nota: il totale teorico è calcolato come 165 store L4L $\times$ 359 giorni di calendario (escluse le principali festività nazionali in cui tutti gli store risultano chiusi).

Prima di procedere con l'analisi econometrica è pertanto necessario distinguere con precisione due fenomeni strutturalmente differenti che possono generare l'assenza di una combinazione store-giorno nel dataset: da un lato le chiusure operative pianificate, in cui il punto vendita non era fisicamente aperto al pubblico, e dall'altro i casi in cui lo store fosse tecnicamente aperto ma non abbia registrato alcuna vendita, situazione che comporterebbe la mancata osservazione di un esito economicamente rilevante.

2.2.1 Entità del fenomeno

L'analisi sistematica delle combinazioni store-giorno mancanti nel campione L4L mostra che il fenomeno è quantitativamente contenuto. Il 35,8% degli store (59 su 165) non presenta alcun giorno mancante nell'intero anno. Il 66,7% degli store (110 su 165) presenta al massimo un giorno mancante, mentre il 76,4% (126 su 165) presenta al massimo due osservazioni mancanti.

2.2.2 Chiusure per inventario

Gli store con al massimo due giorni mancanti, pari al 76,4% del campione, presentano un profilo coerente con le normali chiusure operative per inventario annuale. Infatti, secondo la prassi operativa descritta dal management, nel retail fashion ogni punto vendita è soggetto almeno una volta all'anno a operazioni di conteggio fisico delle scorte che richiedono generalmente una giornata completa di inattività commerciale; gli store di maggiori dimensioni possono richiedere due giornate consecutive. In questi casi l'assenza dal dataset appare coerente con una sospensione programmata dell'attività commerciale: lo store era chiuso al pubblico e nessuna transazione era tecnicamente possibile. Dal punto di vista econometrico, tali osservazioni non sembrano quindi riflettere una mancata registrazione di vendite effettivamente realizzate, bensì l'assenza stessa dell'operatività commerciale. Sul piano statistico, non è possibile verificare direttamente la natura di ciascuna osservazione mancante, poiché il dataset non include un calendario ufficiale delle chiusure operative.

2.2.3 Chiusure strutturali domenicali

La distribuzione aggregata delle osservazioni mancanti mostra un pattern fortemente concentrato in giornate compatibili con regimi di chiusura programmata. Su 1.054

combinazioni store-giorno mancanti nell'intero campione L4L, 748 (71,0%) ricadono di domenica. Ulteriori 148 osservazioni (14,0%) si concentrano di venerdì; tale valore è quasi interamente spiegato dalla festività di Santo Stefano, che nell'anno analizzato cadeva di venerdì e durante la quale 99 store risultavano chiusi. Le restanti osservazioni mancanti si distribuiscono invece in modo sostanzialmente uniforme tra gli altri giorni della settimana. Complessivamente, circa l'85% delle osservazioni mancanti è concentrato tra domeniche e festività nazionali, suggerendo che il fenomeno sia prevalentemente riconducibile a chiusure operative programmate.

La Tabella 2.2 riporta tutti i punti vendita con la più elevata numerosità di missing nel campione, ad eccezione di un unico caso che verrà discusso separatamente nella sezione successiva. Nonostante il numero elevato di osservazioni mancanti, tutti gli store riportati mostrano una concentrazione estremamente elevata dei missing nella giornata di domenica: in molti casi oltre il 90%, con valori che raggiungono quasi il 98%.

Tabella 2.2 – dati mancanti per giorno della settimana

Store	Osservazioni mancanti	Giorno prevalente	Incidenza
store153	57	Dom	87,7%
store156	54	Dom	88,9%
store123	53	Dom	88,7%
store131	53	Dom	88,7%
store140	51	Dom	92,2%
store166	49	Dom	98,0%
store130	47	Dom	89,4%
store141	47	Dom	93,6%
store146	47	Dom	93,6%
store147	47	Dom	97,9%
store152	47	Dom	95,7%
store132	40	Dom	95,0%

Nota: la tabella riporta i dodici store con il maggior numero di osservazioni mancanti nel campione L4L, escludendo store151 che presenta un profilo di missing strutturalmente diverso che verrà discusso separatamente nella sezione 2.2.4.

La colonna "Incidenza" indica la percentuale di osservazioni mancanti che ricade nella giornata di domenica.

Tale regolarità suggerisce che il fenomeno non sia riconducibile a problemi di qualità del dato o a mancate registrazioni delle vendite, bensì a regimi di chiusura domenicale strutturale, probabilmente associati a specifiche caratteristiche commerciali o territoriali dei punti vendita. La presenza di rare osservazioni domenicali nel dataset appare inoltre coerente con aperture straordinarie effettuate in particolari periodi dell'anno.

La Tabella 2.2 mostra quindi che, anche tra gli store con il maggior numero di osservazioni mancanti dell'intero campione, il pattern dominante rimane quello della chiusura domenicale sistematica. L'unica eccezione rilevante è rappresentata dallo store151.

2.2.4 Caso di store con pattern misti di missing

Tra tutti gli store del campione, un solo punto vendita presenta contemporaneamente un numero elevato di osservazioni mancanti e una distribuzione non riconducibile quasi esclusivamente alle chiusure domenicali. Si tratta dello store151 (Mall|Sud), che con 77 osservazioni mancanti rappresenta il caso con la maggiore numerosità di missing dell'intero dataset. Questo punto vendita mostra una struttura più articolata dei missing, con una componente prevalente di chiusure domenicali ma anche ulteriori giornate di inattività distribuite lungo l'anno. Nel dettaglio, 51 osservazioni mancanti (66,2%) ricadono di domenica, suggerendo la presenza di una chiusura domenicale strutturale. Una quota aggiuntiva di osservazioni mancanti coincide inoltre con festività nazionali (Epifania, Festa della Liberazione, Festa della Repubblica, Ognissanti, Immacolata e Santo Stefano) e giorni limitrofi ad essi.

Dal punto di vista operativo, tale configurazione appare più coerente con una combinazione di chiusure programmate e possibili sospensioni temporanee dell'attività commerciale, piuttosto che con fenomeni sistematici di perdita o corruzione dei dati. In particolare, l'assenza di blocchi prolungati di osservazioni mancanti riduce la plausibilità di problemi strutturali nel processo di raccolta dati o di interruzioni permanenti dell'operatività dello store.

Anche nel caso più estremo osservato nel campione, la struttura empirica dei missing appare quindi compatibile con specifiche modalità operative del punto vendita più che con un problema sistematico di dati mancanti non osservati.

2.2.5 Implicazioni econometriche

L'assenza nel dataset di osservazioni con vendite pari a zero implica che le analisi econometriche successive non consentono di identificare l'effetto delle promozioni sulla probabilità estensiva di generare almeno una vendita, ma esclusivamente sull'intensità delle vendite nei giorni in cui l'attività commerciale è osservata. Formalmente, i coefficienti stimati descrivono quindi il comportamento condizionato a giornate con attività commerciale osservata, ovvero: $E[Y|Y>0]$ e non la distribuzione incondizionata delle vendite. Tale limitazione deve essere esplicitamente riconosciuta nell'interpretazione dei risultati. Tuttavia, alla luce dell'analisi strutturale dei missing, appare ragionevole ritenere che la grande maggioranza delle osservazioni mancanti derivi da chiusure operative programmate piuttosto che da giornate di apertura senza domanda.

2.2.6 Conclusione sulla qualità dei dati

Nel complesso, il quadro che emerge dall'analisi è rassicurante sul piano della qualità e dell'utilizzabilità del dataset. La quota di osservazioni mancanti è contenuta in termini assoluti e, soprattutto, presenta pattern altamente coerenti con regimi di chiusura operativa programmata, sia per inventario sia per chiusure domenicali strutturali. Non è possibile escludere in via assoluta l'esistenza di giornate in cui uno store fosse aperto senza registrare alcuna transazione e quindi senza comparire nel dataset; tuttavia, le evidenze empiriche disponibili suggeriscono che tali casi siano quantitativamente limitati e non sufficientemente sistematici da compromettere in misura sostanziale l'interpretazione econometrica delle stime. Anche i casi con maggiore numerosità di osservazioni mancanti mostrano pattern temporalmente strutturati e operativamente interpretabili, elemento che rafforza ulteriormente la plausibilità di una struttura dei missing prevalentemente riconducibile a regimi di chiusura commerciale.

Le analisi presentate nei capitoli successivi possono pertanto essere condotte sul dataset così strutturato, con la consapevolezza che i coefficienti stimati descrivono il comportamento degli store nei giorni di effettiva operatività commerciale osservata.

2.3 Descrizione delle variabili

2.3.1 KPI principali

Il dataset include un insieme di indicatori chiave di performance (KPI) che misurano le diverse dimensioni di rendimento commerciale degli store.

- **Ticket Count:** Numero di transazioni effettuate. Rappresenta una misura diretta della domanda realizzata ed è uno dei principali driver delle vendite complessive.
- **Sales Net Value:** Valore netto delle vendite, espresso in euro. Costituisce la principale misura di performance economica, integrando sia la quantità venduta sia il livello di prezzo effettivo.
- **Sales Discount Value:** Valore monetario complessivo degli sconti applicati, espresso in euro. Consente di quantificare l'intensità promozionale in termini assoluti e l'impatto diretto delle politiche di pricing sui ricavi.
- **Sales Quantity:** Quantità totale di prodotti venduti. Misura il volume fisico delle vendite, indipendentemente dal valore monetario.

2.3.2 KPI derivati

Le variabili *Sales Net Value*, *Sales Discount Value* e *Sales Quantity* descrivono dimensioni fondamentali dell'attività commerciale, ma risultano fortemente correlate con il *Ticket Count*, poiché un numero più elevato di transazioni giornaliere genera meccanicamente valori aggregati più alti per tali grandezze. Queste variabili rappresentano quindi misure prevalentemente estensive.

Per isolare in modo più chiaro il comportamento medio della singola transazione e ridurre la dipendenza dalla scala operativa dello store, l'analisi introduce alcune variabili derivate normalizzate rispetto al numero di scontrini. Tali indicatori possono essere interpretati come misure intensive, in quanto descrivono la struttura media dell'acquisto piuttosto che il semplice volume aggregato di attività. Questo approccio consente di estrarre informazione aggiuntiva dal dataset e di distinguere più chiaramente gli effetti legati al traffico da quelli legati al comportamento d'acquisto.

- **ATV (Average Ticket Value):** valore medio per scontrino, espresso in euro e definito come $Sales\ Net\ Value / Ticket\ Count$. Rappresenta lo scontrino medio e sintetizza congiuntamente effetti di prezzo e composizione del paniere di acquisto.
- **Discount:** sconto medio applicato alle transazioni, espresso in termini percentuali e definito come $Sales\ Discount\ Value / (Sales\ Net\ Value + Sales\ Discount\ Value)$. Costituisce la principale variabile esplicativa della tesi ed è interpretato come misura sintetica dell'intensità promozionale applicata sul prezzo pieno. La variabile Discount incorpora sia componenti strutturali di pricing (tipiche di format come Outlet) sia componenti transitorie legate a campagne promozionali.
- **UPT (Units per Ticket):** numero medio di articoli acquistati per scontrino, calcolato come $Sales\ Quantity / Ticket\ Count$. Misura la profondità del carrello e, insieme al prezzo medio per unità, contribuisce alla determinazione dello scontrino medio.

2.3.3 Variabili strutturali

Il dataset include un insieme di variabili strutturali che descrivono le caratteristiche dei punti vendita e consentono di controllare per eterogeneità non osservata tra le diverse unità analizzate. Tali variabili risulteranno particolarmente rilevanti nell'ambito dei modelli panel con effetti fissi.

- **Store:** Identificativo univoco del punto vendita. Costituisce l'unità cross-sectional del dataset ed è utilizzato per la costruzione della struttura panel. La variabile permette inoltre l'introduzione di effetti fissi a livello di store, fondamentali per controllare caratteristiche non osservabili ma persistenti nel tempo.
- **Tipologia:** Classificazione del punto vendita (diviso in Outlet, Mall, Urban). Questa variabile cattura differenze strutturali nel modello di business e nel comportamento della clientela, che possono influenzare sia il livello della domanda sia la sensibilità alle politiche promozionali.
- **L4L:** Indicatore di stabilità del punto vendita nel periodo di analisi. Un valore pari a 1 identifica store attivi per l'intero orizzonte temporale considerato, senza aperture o chiusure straordinarie.

- **Region:** Area geografica di appartenenza dello store. Permette di controllare per differenze territoriali nei comportamenti di consumo, nelle condizioni competitive, nel potere d'acquisto e nelle dinamiche commerciali locali.

2.3.4 Variabili temporali

Il dataset include un insieme di variabili temporali che permettono di controllare per dinamiche di calendario, componenti stagionali e variazioni sistematiche nel tempo, elementi fondamentali per isolare correttamente l'effetto delle politiche promozionali sui KPI osservati.

- **Date:** Data giornaliera dell'osservazione. Rappresenta l'unità temporale di base del dataset e consente di analizzare le dinamiche operative ad alta frequenza delle vendite, del traffico e dell'attività promozionale.
- **Week:** Identificativo settimanale dell'anno. È utilizzato per catturare shock temporali comuni a tutti gli store.
- **Day Of Week:** Giorno della settimana (1–7). Permette di distinguere pattern sistematici tra giorni feriali e weekend, che influenzano in modo significativo traffico, comportamento d'acquisto, composizione della clientela e risposta alle promozioni.
- **Month:** Mese dell'anno. Cattura componenti di ciclicità di medio periodo legate al calendario commerciale.

2.3.5 Variabili promozionali e attività commerciali

Il dataset include numerose variabili binarie che identificano l'attivazione di specifiche iniziative promozionali e commerciali adottate dall'azienda nel corso dell'anno. Tali variabili non rappresentano semplicemente la presenza generica di una promozione, ma descrivono in modo più granulare le diverse modalità con cui l'azienda interviene sulla politica di prezzo e sugli incentivi alla domanda. In particolare, le variabili catturano:

- **Campagne promozionali strutturate nel calendario commerciale, come:**
 - saldi stagionali
 - mid season sale
 - periodi di pre-saldo

- eventi commerciali specifici ad alta intensità promozionale (es. Black Friday)
- **Iniziative promozionali strategiche o mirate**, che includono:
 - sconti percentuali su categorie o combinazioni di prodotti
 - promozioni “bundle” o multi-acquisto
 - campagne a soglia di spesa (es. sconto su minimo importo)
 - promozioni outlet o di liquidazione stock
- **Programmi di incentivazione e loyalty**, come:
 - cashback
 - coupon e voucher
 - campagne di redemption su acquisti precedenti
 - collaborazioni con brand o iniziative partner
- **Eventi commerciali e campagne stagionali specifiche**, legati a particolari momenti dell'anno o ricorrenze (es. San Valentino, Single's Day, Festa della Mamma).

È importante sottolineare che tali variabili non identificano necessariamente trattamenti esogeni. Le decisioni promozionali dell'azienda sono infatti potenzialmente correlate alla domanda attesa, alle performance recenti degli store o ad altre informazioni operative osservate dal management ma non interamente catturate nel dataset. Di conseguenza, le successive analisi econometriche pongono particolare attenzione ai possibili problemi di endogeneità, attraverso l'utilizzo di modelli panel con effetti fissi e di strategie quasi-sperimentali finalizzate a migliorare l'identificazione causale dell'effetto delle promozioni.

3. Analisi descrittive

Questa sezione presenta un'analisi esplorativa del dataset con un duplice obiettivo. Il primo è documentare le caratteristiche strutturali del panel (statistiche di sintesi dei KPI, la struttura della politica promozionale, l'eterogeneità tra store) per fornire il contesto necessario all'interpretazione dei risultati econometrici. Il secondo è anticipare empiricamente le tensioni analitiche centrali della tesi: lo sconto genera traffico, ma a quale costo sul fatturato? E questo trade-off è uniforme o varia sistematicamente in funzione del contesto operativo? Le evidenze descrittive che seguono non rispondono a queste domande, quello è compito dei modelli, ma ne motivano la formulazione e orientano le scelte metodologiche adottate nel capitolo successivo.

3.1 Struttura del dataset e KPI principali

Il dataset è un panel di osservazioni giornaliere a livello di store, con 58.181 osservazioni complessive distribuite su 165 store e 359 giorni di calendario. Le unità di analisi coprono un orizzonte temporale di un anno e sono suddivise tra tre tipologie di format commerciale: Mall, Urban e Outlet, e tre aree geografiche: Sud, Centro, Nord. I KPI osservati per ogni coppia store-giorno sono il numero di transazioni (Ticket Count), il fatturato netto (Sales Net), lo sconto medio praticato (Discount), le unità per transazione (UPT) e il valore medio per transazione (ATV). La Tabella 3.1 riporta le statistiche descrittive per queste variabili

Tabella 3.1 - Statistiche descrittive dei KPI principali

KPI	Media	Dev. Std.	P25	Mediana	P75	Min	Max
Ticket Count	31	33	18	31	51	1	501
Sales Net (€)	2.285,4€	1920	1.007,3€	1735,4€	2.977,9€	2,8€	26.714,7€
Discount (%)	19,9%	12	9,7%	18,2%	28,7%	0%	83,4%
UPT	1,5	0,26	1,3	1,5	1,6	1	12,0
ATV (€)	58,1€	15	47,2€	56,1€	67,3€	2,8€	418,6€

Nota: l'unità di osservazione è la coppia store-giorno del campione L4L. Dev. Std. = deviazione standard, misura della dispersione attorno alla media. P25 e P75 sono il primo e terzo quartile: il 25% e il 75% delle osservazioni cade rispettivamente al di sotto di questi valori. La mediana (P50) è il valore centrale della distribuzione.

Il Ticket Count medio giornaliero per store è pari a 31 transazioni, il range va da 1 a 501 transazioni, e il rapporto interquartile ($P75/P25 \approx 2,8$) indica che gli store più attivi movimentano quasi tre volte il volume degli store più piccoli. In un contesto retail fashion questa dispersione è attesa: un punto vendita in un grande centro commerciale del Nord in un sabato di saldi opera in condizioni strutturalmente diverse da un negozio di città media del Sud in un martedì ordinario. La sostanziale coincidenza tra media e mediana (31) suggerisce una distribuzione approssimativamente simmetrica attorno al valore centrale.

Il fatturato netto medio è pari a 2.285€ per store-giorno, con una mediana di 1.735€: la differenza tra i due indica una distribuzione asimmetrica con code molto più lunghe verso destra, trainata da giornate o store con volumi eccezionalmente alti. Il massimo osservato è 26.715€, plausibilmente uno store di grande formato in un giorno di punta. Il minimo di 2,8€ è invece compatibile con una giornata di apertura con una singola transazione registrata.

Le unità vendute per ticket (UPT) presentano una media di 1,5 e una distribuzione molto compressa, con un range interquartile che va da 1,3 a 1,6: la grande maggioranza degli scontrini contiene quindi tra uno e due articoli. Questo è un dato rilevante per il retail fashion, dove l'acquisto monoprodotto è la norma. Il massimo di 12 pezzi per scontrino è classificabile come outlier, dal momento che il valore supera il terzo quartile di oltre tre volte lo scarto interquartile.

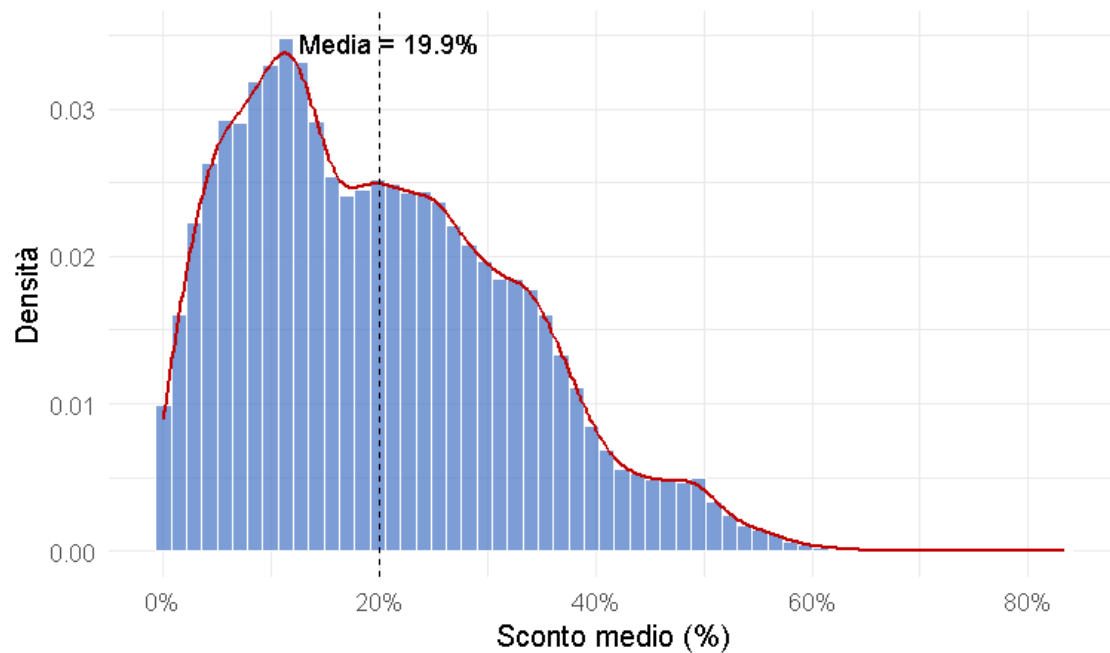
Il valore medio per transazione (ATV) è pari a 58,1€, confermando il posizionamento del brand nella fascia media alta del retail fashion. L'intervallo interquartile, compreso tra 47,2€ e 67,3€, evidenzia una variabilità rilevante del valore dello scontrino, suggerendo una certa eterogeneità nei comportamenti di acquisto tra clienti e contesti operativi. Tale dispersione può riflettere differenze strutturali degli store o l'effetto delle dinamiche promozionali. Anche in questo caso il massimo osservato (418,6€) è classificabile come outlier.

Lo sconto medio giornaliero è pari al 19,9%, con una deviazione standard di 12 punti percentuali. La Figura 3.1 ne mostra la distribuzione che presenta una forma unimodale con un picco intorno al 12-15% e un plateau che si estende fino a circa il 35%.

Questa struttura è compatibile con la coesistenza nel panel di due regimi promozionali distinti: le giornate ordinarie, in cui sono attive promozioni puntuali di entità moderata, e i periodi di saldi, in cui lo sconto si porta strutturalmente sopra il 20-25%. La coda destra è lunga ma non uniforme: la lieve ripresa della densità intorno al 40-45% riflette il contributo degli Outlet, dove sconti in quella fascia costituiscono il posizionamento di prezzo ordinario del canale, non

un'eccezione promozionale. Questa eterogeneità nella distribuzione dello sconto è uno dei motivi per cui l'analisi econometrica non può limitarsi a stimare un effetto medio, ma deve tenere conto del contesto in cui lo sconto viene praticato.

Figura 3.1 – Distribuzione dello sconto medio giornaliero per store



Nota: istogramma della distribuzione della variabile Discount nel campione L4L. La linea rossa rappresenta la densità kernel stimata con kernel gaussiano. La linea tratteggiata indica la media campionaria (19,9%). Sull'asse delle ordinate è riportata la densità, non la frequenza assoluta.

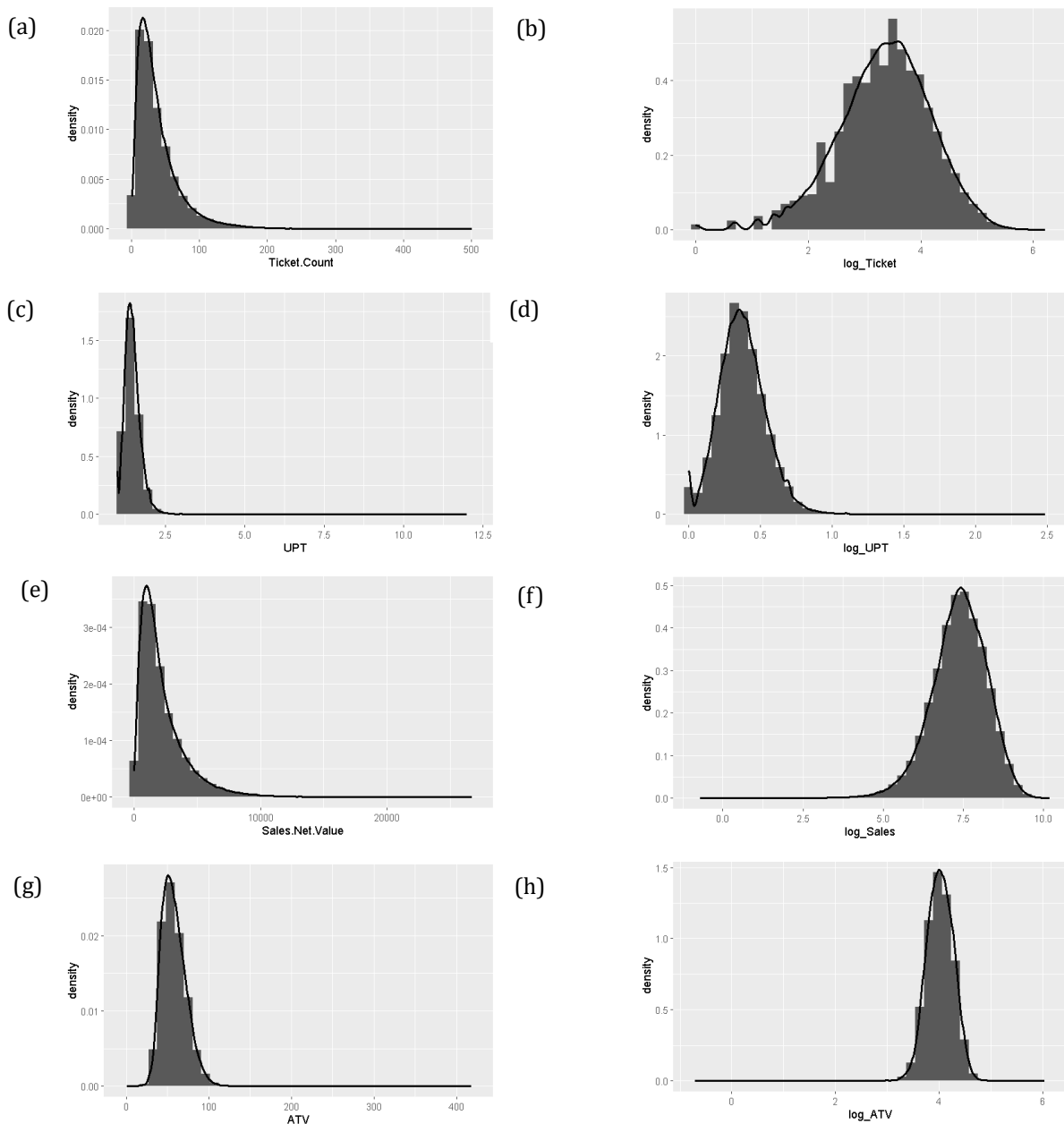
3.1.1 Log-trasformazione e outlier

Le statistiche descrittive evidenziano la presenza di una forte eterogeneità nei KPI osservati. In particolare, alcune variabili presentano code molto estese verso destra, con osservazioni estreme che si discostano in modo significativo dalla massa centrale della distribuzione. Per visualizzare più chiaramente questa struttura, vengono riportate le distribuzioni dei KPI sia in livelli sia in forma logaritmica.

I grafici in livelli (Figura 3.2, pannelli a, c, e, g) mostrano distribuzioni marcatamente asimmetriche, con code lunghe e una concentrazione delle osservazioni su valori relativamente bassi, soprattutto per variabili come Sales Net e Ticket Count. La trasformazione logaritmica (Figura 3.2 pannelli b, d, f, h) attenua sensibilmente questa asimmetria, comprimendo le osservazioni estreme e rendendo le distribuzioni più simmetriche e meno influenzate dai valori

anomali. Questo non modifica l'informazione contenuta nei dati, ma ne migliora la leggibilità e la stabilità statistica, risultando particolarmente utile nelle successive specificazioni econometriche.

Figura 3.2 - Distribuzione dei KPI in livelli e in forma logaritmica



Nota: istogramma con densità kernel sovrapposta. I pannelli (a, c, e, g) mostrano la distribuzione originale dei KPI, i pannelli (b, d, f, h) mostrano la distribuzione dopo trasformazione logaritmica.

Nel contesto retail analizzato, questi valori estremi non rappresentano errori di misurazione o anomalie strutturali, ma riflettono dinamiche operative reali: giornate di picco promozionale, store ad alta affluenza o eventi commerciali straordinari. Per questa ragione, pur essendo statisticamente classificabili come outlier, tali osservazioni vengono mantenute nell'analisi, in quanto informativamente rilevanti e coerenti con la natura del fenomeno studiato.

3.2 Struttura temporale della politica promozionale

Lo sconto non è distribuito uniformemente nel tempo né tra tipologie di campagna: la sua struttura riflette il calendario commerciale, con periodi di intensa attività promozionale che si alternano a fasi di relativa quiete. La Figura 3.3 mostra l'evoluzione settimanale del Ticket Count medio e dello sconto medio aggregato su tutti gli store L4L nel corso dell'anno, evidenziando i periodi di saldi invernali ed estivi e il Black Friday.

Il primo elemento che emerge dalla serie storica è la presenza di un livello base di sconto che persiste lungo tutto l'anno, anche nei periodi privi di campagne promozionali attive. Questo livello strutturale, intorno al 12-15% nelle settimane di quiete promozionale, riflette la componente di pricing ordinaria di ciascuno store, in particolare degli Outlet, per i quali lo sconto non è uno strumento promozionale occasionale ma la condizione operativa permanente. Questa componente strutturale è invariante nel tempo per ciascun store e viene assorbita dall'effetto fisso di store nei modelli panel: quello che i modelli stimano è quindi esclusivamente la risposta alla variazione dello sconto attorno a questo livello base, non l'effetto del livello assoluto.

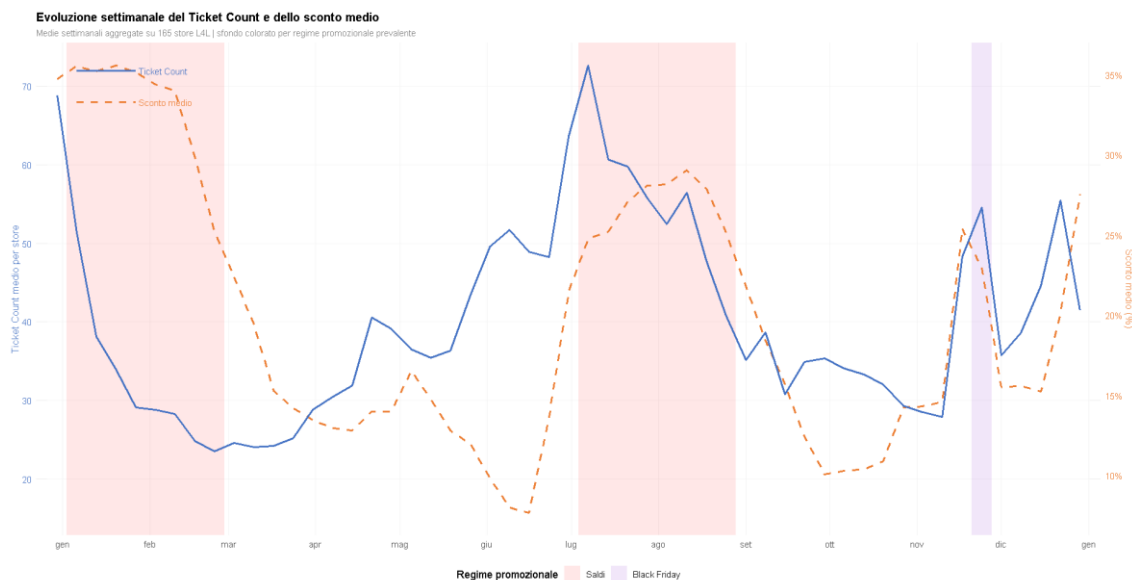
Il secondo elemento rilevante è la co-movimentazione tra sconto e traffico nei periodi di intensa attività promozionale. Durante i saldi invernali (gennaio-febbraio) e quelli estivi (luglio-agosto) si osserva un'elevazione simultanea di entrambe le serie: lo sconto medio sale verso il 28-32% e il Ticket Count mostra i picchi più elevati dell'anno. Questo pattern descrittivo è coerente con l'ipotesi che lo sconto generi traffico, ma non è sufficiente per stabilire una relazione causale: i picchi di traffico nei saldi riflettono anche fattori stagionali indipendenti dallo sconto, come la concentrazione degli acquisti di inizio stagione e le dinamiche di domanda strutturalmente più elevata in quel periodo.

Il Black Friday si distingue per un profilo diverso: la finestra temporale è concentrata in pochi giorni, lo sconto sale rapidamente e il Ticket raggiunge un picco, per poi rientrare altrettanto rapidamente alla fine dell'evento. Questo profilo temporale è strutturalmente diverso da quello

dei saldi, che mostrano invece un'elevazione più graduale e prolungata. I due periodi non sono quindi direttamente confrontabili in termini di meccanismo di risposta della domanda, e l'analisi econometrica li tratta come regimi distinti proprio per questa ragione.

Al di fuori dei saldi e del Black Friday le variazioni dello sconto sono più contenute e meno sistematiche, con oscillazioni intorno al livello base che riflettono l'attivazione di campagne promozionali puntuali: Mid Season Sale, Pre-Saldi, promozioni a soglia, di durata e copertura variabile.

Figura 3.3 – Evoluzione settimanale del Ticket Count e dello sconto medio



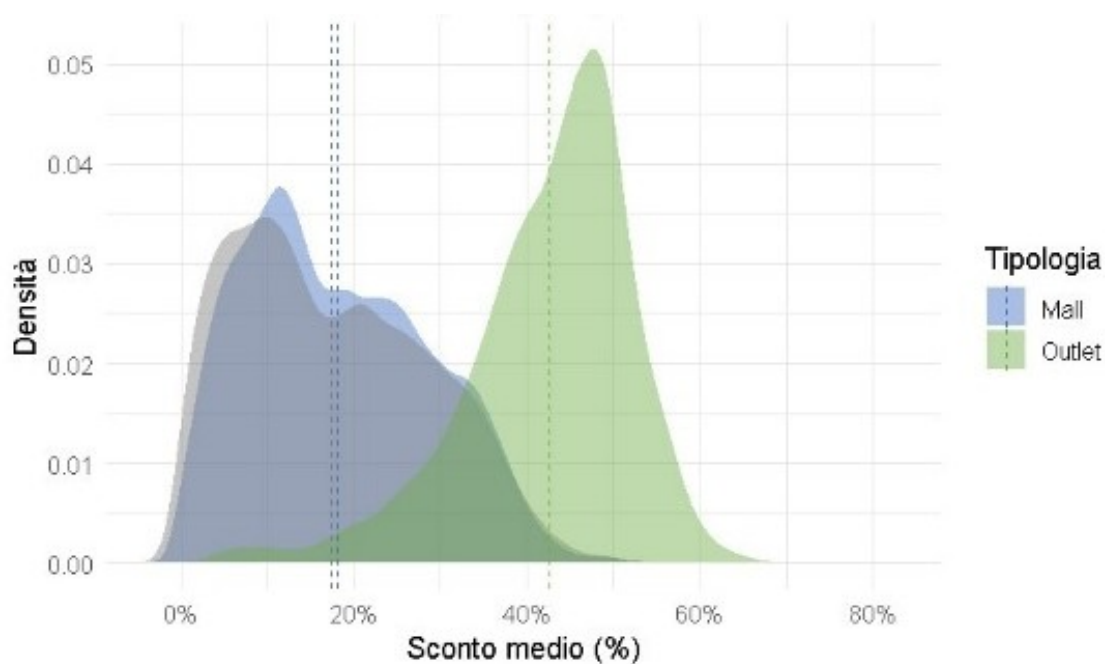
Nota: medie settimanali calcolate sul campione L4L (165 store, 359 giorni). La linea blu rappresenta il Ticket Count medio per store; la linea arancione tratteggiata rappresenta lo sconto medio (Discount), leggibile sull'asse destro. Le bande colorate indicano le settimane in cui il regime promozionale è Saldi (rosa) o Black Friday (lilla).

3.3 Eterogeneità nelle osservazioni

3.3.1 Tipologia di store

Come mostrato in Figura 3.4 la marcata variabilità tra formati segnala un'eterogeneità strutturale che costituisce una delle principali fonti di confondimento nelle stime aggregate. Gli store Outlet, pur rappresentando solo 13 unità su 165, mostrano un profilo operativo marcatamente distinto. L'elemento più caratterizzante è lo sconto medio: 42,6% negli Outlet, più che doppio rispetto a Mall (18,2%) e Urban (17,5%), riflettendo il posizionamento di prezzo proprio di questo canale.

Figura 3.4 - Distribuzione dello sconto medio per tipologia di store



Nota: densità kernel per Mall, Urban e Outlet. Le linee tratteggiate indicano le medie per gruppo. Campione L4L.

La Tabella 3.2 conferma questa differenza strutturale. Il Ticket medio giornaliero degli Outlet è pari a 58,8 transazioni, contro le 39,1 dei Mall e le 36,1 degli Urban, segnale di un traffico strutturalmente più elevato; coerentemente, anche il fatturato medio è superiore (2.685€ contro circa 2.250€ per gli altri due format). Questa eterogeneità rende gli store Outlet una categoria analiticamente separata, e giustifica l'inclusione della variabile di format come dimensione di controllo nei modelli econometrici.

Tabella 3.2 - KPI medi per tipologia di store

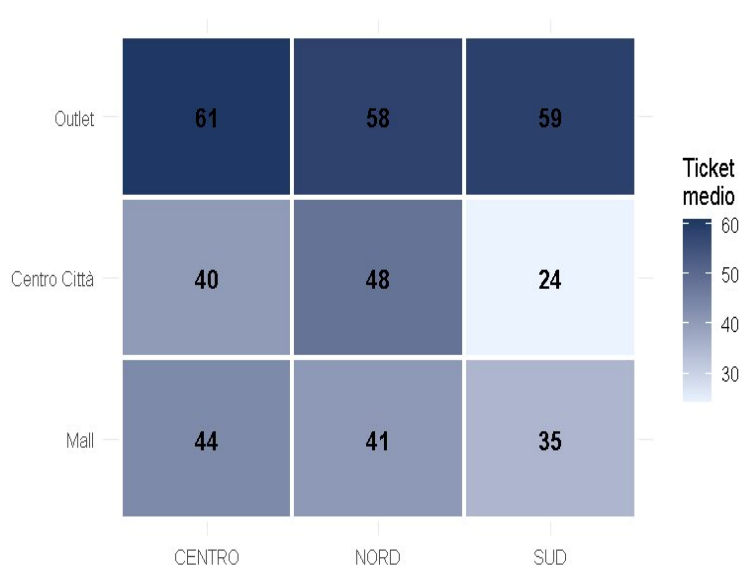
Tipologia	N store	N osservazioni	AVG Ticket	AVG Sales	AVG Discount
Mall	110	39077	39,1	2240€	18,2%
Urban	42	14453	36,1	2279€	17,5%
Outlet	13	4651	58,8	2685€	42,6%

Nota: medie calcolate sul campione L4L (N = 58.181)

3.3.2 Area geografica

Lo sconto medio varia anche in funzione dell'area geografica, seppur in misura più contenuta rispetto alla tipologia. Il Nord presenta il valore più alto (21,8%), seguito da Sud (19,5%) e Centro (17,9%). La heatmap in Figura 3.5 mostra il numero di Ticket medio per la combinazione Tipologia × Regione e rivela una struttura interessante: gli Outlet generano un numero di Ticket elevati e sostanzialmente omogenei tra Nord (58), Centro (61) e Sud (59), mentre per i Mall e i Centri Città la variabilità geografica è più pronunciata, in particolare i Centri Città del Sud mostrano un numero di Ticket medi molto inferiori (24) rispetto al Centro (40) e al Nord (48).

Figura 3.5 - Ticket medio giornaliero per tipologia di store e area geografica



Nota: valori medi del Ticket Count per cella tipologia × area geografica. Campione L4L, N = 58.181. L'intensità del colore è proporzionale al valore medio, celle più scure indicano traffico più elevato.

3.3.3 Stagionalità: feriale vs weekend

Una delle dimensioni di eterogeneità più rilevanti per l'analisi econometrica è la distinzione tra giorni feriali e weekend. La Tabella 3.3 mostra che il Ticket medio nel weekend è 56,6, contro 32,2 nei giorni feriali, pari ad un incremento del 75,9%. L'effetto sulle Sales Net è di dimensione analoga (79,6%). Lo sconto medio, al contrario, è pressoché identico nei due periodi (20,4% vs 20,8%) indicando che il maggiore traffico del weekend non è associato a uno sconto più elevato:

la differenza di volume riflette dunque una diversa composizione della domanda e del comportamento d'acquisto del cliente, non una diversa intensità promozionale.

Tabella 3.3 - KPI medi per fascia temporale

	AVG Ticket	AVG Sales	AVG Discount
Feriale	32,2	1.817€	20,4%
Weekend	56,6	3.263€	20,8%

Nota: medie calcolate sul campione L4L (N = 58.181).

3.4 Relazioni bivariate tra KPI

La matrice di correlazione in Tabella 3.4 riassume le relazioni tra i cinque KPI principali. Nonostante le correlazioni descrittive non abbiano interpretazione causale, queste possono essere utili per evidenziare pattern strutturali, in questo senso due risultati meritano attenzione particolare.

Tabella 3.4 - Matrice di correlazione tra i KPI principali

	Ticket Count	Sales Net	Sconto	UPT	ATV
Ticket Count	1,000	0,928	0,129	0,055	-0,070
Sales Net	0,928	1,000	-0,028	0,131	0,213
Sconto	0,129	-0,028	1,000	0,114	-0,459
UPT	0,055	0,131	0,114	1,000	0,425
ATV	-0,070	0,213	-0,459	0,425	1,000

Nota: valori in verde = correlazione positiva; in rosso = correlazione negativa. Intensità del colore proporzionale al valore assoluto. Soglie: chiaro $\geq 0,10$, medio $\geq 0,40$, scuro $\geq 0,80$.

Il primo è la correlazione molto elevata tra Ticket Count e Sales Net ($r = 0,928$): il volume di traffico è il principale driver del fatturato, e le due variabili sono sostanzialmente intercambiabili come misura dell'attività di uno store.

Il secondo risultato rilevante è la correlazione negativa tra sconto e ATV ($r = -0,459$), la più forte dell'intera matrice al netto della coppia Ticket/Sales. Questo conferma il meccanismo di compressione del valore unitario: all'aumentare dello sconto, il prezzo medio pagato per transazione si riduce in modo sistematico.

La correlazione positiva tra UPT e ATV ($r = 0,425$) indica invece che le transazioni di maggiore valore tendono anche ad avere più articoli, mentre la correlazione quasi nulla tra sconto e Sales Net ($r = -0,028$) anticipa uno dei risultati econometrici centrali: l'associazione quasi nulla tra sconto e Sales Net è coerente con un possibile trade-off tra aumento dei volumi e riduzione del prezzo medio, ipotesi che verrà approfondita nelle analisi econometriche successive.

Un'ultima osservazione aiuta a inquadrare il trade-off centrale dell'analisi. Il fatturato netto è per costruzione il prodotto tra numero di transazioni e valore medio per transazione, il che implica che in forma logaritmica:

$$\log(\text{Sales Net}) = \log(\text{Ticket Count}) + \log(\text{ATV})$$

Lo sconto agisce simultaneamente su entrambe le componenti, ma in direzioni opposte: tende ad aumentare il traffico e il numero di transazioni, e al tempo stesso a comprimere il valore medio dello scontrino. Il saldo netto sul fatturato dipende quindi da quale dei due effetti prevale. Questa identità non è solo una nota contabile: guida direttamente la lettura dei risultati econometrici, dove stimare separatamente l'elasticità su Ticket Count e su ATV permette di scomporre l'effetto totale dello sconto nelle sue due componenti, capendo in quali contesti il volume compensa la perdita di valore unitario e in quali no.

3.5 Sintesi e implicazioni per l'analisi econometrica

L'analisi descrittiva delinea un quadro coerente che orienta le scelte metodologiche delle sezioni successive. Quattro elementi meritano di essere evidenziati.

La forte eterogeneità cross-section tra store, sia per tipologia che per area geografica, rende necessario il controllo per effetti fissi a livello di store. La differenza di oltre 20 punti percentuali nello sconto medio tra Outlet e altri format, e la variabilità nel Ticket che va da 1 a

501, non possono essere trattate come rumore casuale: riflettono differenze strutturali stabili che, se non controllate, confonderebbero la stima dell'effetto dello sconto.

La presenza di variabilità temporale sistematica (stagionalità settimanale, campagne promozionali concentrate) richiede il controllo per effetti fissi temporali. I fixed effects per giorno assorbono tutti i trend comuni a tutti gli store in un dato giorno, isolando la variazione idiosincratICA dello sconto all'interno di ciascuno store.

La distribuzione dello sconto, con una componente strutturale intorno al 10-20% e una componente promozionale che può superare il 40%, suggerisce che il dataset contiene due regimi di prezzo distinti. Questo non è un problema per la specificazione con fixed effects (che sfrutta la variazione within-store nel tempo), ma è rilevante per l'analisi dell'eterogeneità di impatto: le elasticità stimate potrebbero essere diverse a seconda che si tratti di variazione ordinaria dello sconto o di variazione da campagna.

La correlazione quasi nulla tra sconto e Sales Net ($r = -0,028$) e quella negativa con l'ATV ($r = -0,459$) anticipano la tensione centrale dell'analisi, già formalizzata attraverso l'identità log-additiva nella sezione precedente: lo sconto genera traffico ma comprime il prezzo medio, e il saldo netto dipende da quale dei due effetti prevale nel contesto specifico.

In sintesi, le analisi descrittive delineano un contesto in cui lo sconto è strutturalmente sistemico eterogeneo e potenzialmente endogeno nelle sue dinamiche temporali. Queste tre caratteristiche: sistematicità, eterogeneità e endogeneità sono esattamente le ragioni per cui una semplice regressione OLS non è sufficiente e per cui la strategia empirica adottata nei capitoli successivi richiede strumenti più sofisticati.

4. Metodologia

4.1 Approccio generale

L'analisi empirica si basa su un approccio econometrico volto a stimare la relazione tra le politiche promozionali e i principali indicatori di performance commerciale osservati a livello store-giorno. La natura longitudinale del dataset offre una fonte di variazione particolarmente utile per l'identificazione causale: invece di confrontare punti vendita differenti tra loro, è possibile confrontare lo stesso store in momenti diversi del tempo, controllando per le sue caratteristiche strutturali invarianti. Questa logica within-store costituisce il principio identificativo comune ai modelli adottati e rappresenta il principale strumento per ridurre il rischio di attribuire agli sconti variazioni nei KPI che riflettono invece eterogeneità persistente tra punti vendita o shock temporali comuni alle diverse giornate di osservazione.

4.2 Modello base e interpretazione dell'elasticità

Per stimare l'effetto dello sconto sui KPI di vendita, si utilizza una specificazione log-lineare:

$$\log(Y_{it}) = \beta \cdot Discount_{it} + \varepsilon_{it}$$

dove:

- Y_{it} rappresenta il KPI di interesse per lo store i nel giorno t ,
- $Discount_{it}$ è lo sconto medio applicato, espresso in punti percentuali,
- β rappresenta la semi-elasticità dello sconto. Per semplicità espositiva, nel resto della tesi il coefficiente verrà comunque indicato come elasticità promozionale.

La scelta della forma funzionale log-lineare è motivata da considerazioni sia teoriche che pratiche (Wooldridge, 2013). Dal punto di vista teorico, la relazione tra sconto e domanda è tipicamente moltiplicativa piuttosto che additiva; inoltre la trasformazione logaritmica risponde a esigenze di natura statistica: i principali KPI commerciali presentano distribuzioni fortemente asimmetriche a destra e caratterizzate dalla presenza di valori estremi, una configurazione poco compatibile con le ipotesi gaussiane sottostanti ai modelli lineari stimati tramite OLS. Come mostrato nell'analisi descrittiva nella sezione 3.1, la trasformazione logaritmica produce distribuzioni significativamente più simmetriche e maggiormente

approssimabili a una distribuzione normale, migliorando la stabilità delle stime e l'affidabilità dell'inferenza statistica.

La variabile di sconto non viene trasformata logaritmicamente poiché è espressa come percentuale e include valori nulli, per i quali la trasformazione logaritmica sarebbe non definita. In termini di interpretazione statistica, la variazione percentuale del KPI associata a un incremento unitario dello sconto è data da $(e^\beta - 1) \times 100$. Per valori contenuti del coefficiente, tale espressione è approssimabile linearmente con $\beta \times 100$, rendendo l'interpretazione direttamente in termini di variazioni percentuali marginali.

Dal punto di vista dell'interpretazione business, la specificazione log-lineare produce un coefficiente β interpretabile come elasticità: un incremento di un punto percentuale dello sconto è associato a una variazione di $\beta \times 100$ del KPI. Discount è espresso su scala 0–100, pertanto il coefficiente β misura la variazione percentuale del KPI associata a un incremento di un punto percentuale dello sconto medio, a parità di tutte le altre condizioni. Questa interpretazione è particolarmente utile nel contesto retail, dove le decisioni promozionali vengono tipicamente espresse in termini percentuali e dove l'effetto di interesse è la risposta relativa della domanda, non quella assoluta.

Questa specificazione fornisce una prima misura dell'associazione tra sconto e vendite ma non consente di isolare un effetto causale, in quanto non controlla per le differenze strutturali tra store né per i trend temporali comuni.

4.3 Modelli con effetti fissi

L'introduzione degli effetti fissi di store e di tempo trasforma la specificazione base in un modello panel che sfrutta esclusivamente la variazione within-store nel tempo:

$$\log(Y_{it}) = \beta \cdot \text{Discount}_{it} + \alpha_i + \gamma_t + \varepsilon_{it}$$

- α_i è l'effetto fisso di store, che assorbe tutte le caratteristiche invarianti nel tempo del punto vendita.
- γ_t è l'effetto fisso giornaliero, che cattura ogni shock comune a tutti gli store in una determinata unità di tempo.

- β rappresenta l'elasticità dello sconto, identificata dalla variazione within-store nel tempo.

Dal punto di vista dell'identificazione, il modello a effetti fissi a due vie riduce il rischio di distorsione derivante da due classi di problemi simultaneamente. La prima riguarda l'eterogeneità cross-sectional: store diversi differiscono strutturalmente per dimensione, localizzazione, format commerciale e composizione della clientela, e queste differenze sono correlate sia con il livello delle vendite che con l'intensità promozionale tipicamente applicata. Senza controllarle, il confronto tra store ad alta e bassa intensità di sconto rifletterebbe in parte queste differenze strutturali piuttosto che l'effetto causale dello sconto. L'effetto fisso α_i elimina questo problema assorbendo il livello medio di ciascuno store. La seconda classe di problemi riguarda gli shock temporali comuni: eventi macroeconomici, stagionalità, festività e campagne di comunicazione centralizzate influenzano simultaneamente le vendite di tutti i negozi e le decisioni promozionali aziendali. L'effetto fisso γ_t cattura questi shock nella loro componente aggregata, impedendo che vengano erroneamente attribuiti alla variazione dello sconto.

Il coefficiente β stimato in questo modello ha quindi un'interpretazione più pulita rispetto alla specificazione base: misura la variazione percentuale del KPI associata a un incremento di un punto percentuale dello sconto, confrontando lo stesso store con sé stesso in giorni diversi, al netto degli andamenti comuni a tutti i negozi (Wooldridge, 2013). In altre parole ci serviamo esclusivamente della varianza within-store e within-day per la stima del coefficiente di elasticità.

4.4 Analisi dell'eterogeneità

L'elasticità media stimata nel modello a effetti fissi potrebbe mascherare una variazione sistematica dell'effetto dello sconto in funzione del contesto operativo. Per catturare questa eterogeneità, il modello viene esteso introducendo termini di interazione tra lo sconto e le variabili di contesto rilevanti: tipologia di store, area geografica e giorno della settimana. La specificazione assume la forma generale:

$$\log(Y_{it}) = \beta_0 \cdot \text{Discount}_{it} + \sum_k \beta_k (\text{Discount}_{it} \times Z_{k,it}) + \alpha_i + \gamma_t + \varepsilon_{it}$$

- $Z_{k,it}$ rappresentano le variabili di contesto, incluse unicamente in interazione con lo sconto.

- Il coefficiente β_0 rappresenta l'elasticità di riferimento per la categoria baseline.
- Il coefficiente β_k rappresenta la variazione dell'elasticità associata alla specifica modalità k rispetto alla baseline.

La struttura delle interazioni adottata nell'analisi empirica va oltre le interazioni semplici tra sconto e ciascuna variabile di contesto separatamente. Il modello include interazioni fino al terzo ordine, coinvolgendo simultaneamente tipologia di store, area geografica e giorno della settimana. Questa scelta riflette l'ipotesi che le tre dimensioni di eterogeneità non agiscano in modo additivo e indipendente, ma che l'effetto dello sconto in un determinato contesto dipenda dalla combinazione specifica di tutte e tre le caratteristiche. Ad esempio, l'effetto del weekend potrebbe essere diverso non solo tra tipologie di store, ma diverso in modo eterogeneo tra tipologie e regioni simultaneamente. Una specificazione limitata alle sole interazioni di secondo ordine imporrebbe restrizioni additive sulla struttura dell'eterogeneità, assumendo implicitamente che l'effetto di una dimensione di contesto sia indipendente dalle altre, vincolando la struttura dell'eterogeneità in modo potenzialmente arbitrario.

Data la presenza di interazioni multiple, le elasticità rilevanti per l'interpretazione economica non coincidono direttamente con un singolo coefficiente stimato, ma con combinazioni lineari di più parametri del modello. In termini generali, l'effetto marginale dello sconto in un determinato contesto può essere espresso come:

$$\theta = a' \beta^{\wedge}$$

dove $\beta^{\wedge}$ è il vettore dei coefficienti stimati e a è un vettore di pesi che seleziona i coefficienti rilevanti per la combinazione considerata.

Poiché tali elasticità sono funzioni dei parametri stimati, anche la loro incertezza deve essere derivata congiuntamente dai coefficienti coinvolti. Gli errori standard delle elasticità marginali vengono quindi calcolati applicando il delta method, utilizzando la matrice di varianza-covarianza stimata del modello. Nel caso di combinazioni lineari, la varianza dell'effetto marginale è:

$$\text{Var}(\theta) = a' \text{VCOV}(\beta^{\wedge}) a$$

Questo approccio consente di ottenere inferenza corretta sulle elasticità specifiche di ciascun contesto operativo, tenendo conto non solo della varianza dei singoli coefficienti ma anche della loro covarianza.

4.5 Specifica degli errori standard

La matrice di varianza-covarianza utilizzata sia per il calcolo degli errori standard dei coefficienti stimati sia per la propagazione dell'incertezza nelle elasticità marginali è costruita utilizzando errori standard clusterizzati a livello di store. Questa scelta riflette una proprietà strutturale dei dati panel store-giorno: le osservazioni relative allo stesso punto vendita nel tempo non sono indipendenti, poiché condividono caratteristiche non osservate persistenti (la qualità del management locale, la composizione stabile della clientela, la posizione geografica) che generano correlazione seriale nei residui. Ignorare questa dipendenza intra-cluster porterebbe a una sottostima sistematica degli errori standard, con inferenze eccessivamente ottimistiche e test di ipotesi distorti verso il rifiuto dell'ipotesi nulla (Cameron & Miller, 2015).

Il clustering è effettuato a livello di store, che rappresenta l'unità di trattamento naturale del dataset e consente inferenza robusta in presenza di eteroschedasticità tra store e correlazione seriale all'interno di ciascun negozio nel tempo.

4.6 Specificazione 2WFE con trattamento binario

Per valutare l'impatto di specifiche campagne promozionali, il modello a effetti fissi a due vie viene applicato con una variabile di trattamento binaria che indica la presenza della promozione:

$$Y_{it} = \delta \cdot Promo_{it} + \alpha_i + \gamma_t + \varepsilon_{it}$$

- $Promo_{it}$ è una variabile dummy che assume valore 1 se lo store i è esposto alla promozione al tempo t .
- δ misura l'effetto medio della campagna sul KPI di interesse.
- α_i e γ_t sono gli effetti fissi di store e di data

L'identificazione di δ si basa sul confronto tra store trattati e non trattati nello stesso periodo, e tra periodi trattati e non trattati per lo stesso store, al netto degli effetti fissi. L'assunzione fondamentale è quella di parallel trends: in assenza della promozione, le traiettorie dei KPI degli store trattati e non trattati avrebbero seguito un andamento parallelo nel tempo. La sua plausibilità dipende dalla natura del processo di selezione al trattamento: se gli store vengono selezionati per la promozione in base a caratteristiche time-varying correlate con i KPI (ad esempio perché la promozione viene attivata in determinati store in risposta a un loro calo di

traffico) l'assunzione è violata e δ non ha un'interpretazione causale anche in presenza degli effetti fissi. Nel contesto analizzato, la plausibilità del parallel trends è supportata anche dal fatto che le decisioni promozionali vengono definite centralmente a livello aziendale e non in risposta alle performance locali dei singoli store, riducendo il rischio che il trattamento sia assegnato selettivamente agli store con traiettorie di KPI già divergenti.

4.7 Regression Discontinuity Design

Per le promozioni che presentano una data di avvio netta e uniforme si può adottare un approccio di Regression Discontinuity Design che sfrutta la discontinuità temporale nell'attivazione della campagna. L'idea identificativa è che, in un intorno sufficientemente stretto del giorno di avvio, le osservazioni immediatamente precedenti e immediatamente successive al cutoff siano sostanzialmente comparabili, differendo sistematicamente solo per l'esposizione alla promozione. In questa finestra ristretta, la variazione nel trattamento è plausibilmente indipendente dalle caratteristiche non osservate degli store, consentendo un'interpretazione causale del coefficiente stimato senza richiedere che il timing promozionale sia casuale sull'intera finestra temporale.

La running variable è definita come la distanza in giorni dal giorno di avvio della campagna

$$running_{it} = t - t_0$$

dove t_0 è il giorno di inizio promozione.

La variabile di trattamento è definita come:

$$D_{it} = \mathbb{1}(running_{it} \geq 0)$$

Il modello locale è:

$$Y_{it} = \tau D_{it} + \beta_1 \cdot running_{it} + \beta_2 \cdot D_{it} \times running_{it} + \alpha_i + \varepsilon_{it}$$

- τ è il salto discontinuo al cutoff, interpretabile come effetto causale locale dell'avvio della promozione.
- β_1 è la pendenza nel periodo pre-cutoff.
- β_2 è la variazione di pendenza nel periodo post-cutoff, che consente alle due rette di avere inclinazioni diverse ai due lati della discontinuità.

- α_i è l'effetto fisso di store, che raffina l'identificazione eliminando le differenze strutturali tra store.

I pesi kernel triangolari $w_{it}=1-|running_{it}|/h$ attribuiscono peso massimo alle osservazioni nel giorno del cutoff e peso decrescente alle osservazioni più lontane, privilegiando il confronto locale dove l'assunzione di comparabilità è più credibile. La scelta della bandwidth h determina il trade-off tra varianza (ridotta ampliando la finestra) e distorsione (aumentata includendo osservazioni più lontane dal cutoff dove le due traiettorie potrebbero divergere per ragioni indipendenti dalla promozione) (Lee & Lemieux, 2010).

4.8 Limiti dell'analisi

Nonostante l'utilizzo di tecniche avanzate, l'analisi presenta alcune limitazioni:

- possibile endogeneità dello sconto. Sebbene le politiche promozionali analizzate siano definite centralmente a livello aziendale e non dai singoli punti vendita, riducendo quindi il rischio di endogeneità derivante da decisioni discrezionali degli store manager in risposta alle performance locali, non è possibile escludere completamente la presenza di una componente residua di endogeneità. In particolare, l'intensità dello sconto potrebbe essere correlata a condizioni commerciali non osservate e variabili nel tempo, quali livelli di stock residuo, qualità dell'assortimento disponibile o aspettative sulla domanda in specifici segmenti commerciali. In tali casi, una parte della variazione dello sconto potrebbe riflettere strategie aziendali di risposta a dinamiche commerciali attese, piuttosto che variazioni pienamente esogene della leva promozionale;
- utilizzo di una misura aggregata dello sconto (Discount), che non cattura la distribuzione degli sconti a livello di singola transazione,
- limitata variabilità dello sconto in alcune campagne, che può ridurre la precisione delle stime,
- assenza di variabili di controllo relative allo stock disponibile e alla profondità di assortimento. Tale limite è particolarmente rilevante nel contesto outlet, dove gli incrementi dello sconto possono coincidere con fasi avanzate di svuotamento del magazzino: in queste situazioni lo sconto aumenta mentre la disponibilità di prodotto diminuisce, generando potenzialmente una correlazione negativa tra sconto e numero di transazioni osservate;

- assenza di informazioni su fattori esterni (es. condizioni meteo, concorrenza locale).

Tali aspetti vengono considerati nell'interpretazione dei risultati. Inoltre, un limite trasversale a tutti i modelli riguarda il livello di aggregazione dell'analisi. Lavorando a livello store-giorno, i KPI osservati (numero di transazioni, vendite nette, valore medio dello scontrino) sono misure aggregate che sintetizzano in un unico numero il comportamento di clienti potenzialmente molto eterogenei. Questo implica che i coefficienti stimati catturano effetti medi sul punto vendita nel suo complesso, senza distinguere se l'incremento del traffico sia attribuibile all'acquisizione di nuovi clienti, a una maggiore frequenza di visita dei clienti abituali, o a un effetto di sostituzione tra giorni. Analogamente, una variazione dell'UPT o dell'ATV potrebbe riflettere cambiamenti nella composizione del paniere di acquisto, nella dimensione media dei gruppi di acquisto o nella selezione dei clienti che scelgono di visitare il negozio durante la promozione, meccanismi distinti che i dati aggregati non consentono di separare. Questa limitazione circonda il livello interpretativo dei risultati: le stime devono essere lette come effetti sulla performance aggregata del punto vendita, non come misure di risposta comportamentale individuale.

5. Elasticità allo sconto

L'analisi empirica si basa su una strategia econometrica finalizzata a stimare la relazione tra intensità promozionale e performance di vendita a livello store-giorno, sfruttando la natura panel del dataset. Il percorso seguito è progressivo: si parte da una specificazione OLS elementare che fornisce una prima misura dell'associazione grezza tra sconto e KPI, per poi raffinare l'identificazione attraverso l'introduzione di effetti fissi a due vie e infine estendere il modello con termini di interazione per studiare l'eterogeneità dell'effetto.

5.1 Elasticità media, OLS

Come primo passo viene stimata una specificazione log-lineare:

$$\log(Y_{it}) = \beta \cdot Discount_{it} + \varepsilon_{it}$$

dove Y_{it} è il KPI di interesse per lo store i nel giorno t (alternativamente Ticket Count, UPT, Sales Net e ATV) e $Discount_{it}$ è lo sconto medio applicato a livello store-giorno, espresso in punti percentuali. Il coefficiente β misura la variazione percentuale del KPI associata a un incremento di un punto percentuale dello sconto medio, ed è interpretabile direttamente come elasticità della domanda rispetto all'intensità promozionale. La trasformazione logaritmica della variabile dipendente è motivata dalla natura moltiplicativa della relazione tra sconto e domanda, inoltre consente di interpretare i coefficienti come variazioni percentuali del KPI e riduce l'influenza delle osservazioni estreme, particolarmente rilevanti nelle variabili di vendita. Questa specificazione fornisce una prima misura dell'associazione tra promozioni e performance, ma non controlla per fattori non osservati che influenzano simultaneamente sconto e vendite, e le stime devono pertanto essere interpretate come associazioni statistiche e non come effetti causali.

5.1.1 Risultati

Come emerge dalla Tabella 5.1, lo sconto è associato a un aumento del numero di transazioni pari a circa +0,67% per ogni punto percentuale di sconto aggiuntivo, un effetto positivo e fortemente significativo. Un segnale nella stessa direzione emerge anche per l'UPT, che cresce di circa +0,15% (un effetto statisticamente robusto ma economicamente contenuto), che suggerisce come le promozioni incidano solo marginalmente sulla quantità di articoli acquistati per singola transazione. Sul lato del valore, il quadro si inverte: le Sales Net mostrano una

riduzione di circa $-0,29\%$ per ogni punto percentuale di sconto, mentre l'ATV subisce una contrazione più marcata, pari a circa $-0,96\%$. Questi ultimi due coefficienti riflettono l'effetto meccanico dell'abbassamento del prezzo effettivo pagato dal consumatore, che non viene compensato dall'incremento dei volumi.

Tabella 5.1 – Elasticità media dello sconto sui KPI: stime OLS

KPI	Elasticità (%)	Std. Error	t-stat	p-value
log(Ticket Count)	+0,674	0,026	25,58	<0,001***
log(UPT)	+0,149	0,005	27,99	<0,001***
log(Sales Net)	-0,289	0,028	-10,49	<0,001***
log(ATV)	-0,963	0,007	-132,0	<0,001***

*Nota: stime OLS su panel giornaliero store-giorno, N = 58.181 osservazioni, campione L4L. La colonna Elasticità riporta il coefficiente β . I coefficienti e gli errori standard sono riportati in termini percentuali (stima e errore standard moltiplicati per 100). Gli errori standard non sono clusterizzati a livello store. *** $p < 0,001$.*

5.1.2 Interpretazione

Il profilo d'effetto che emerge dall'OLS delinea un trade-off strutturale tra prezzo e quantità. Nelle associazioni osservate, livelli più elevati di sconto risultano associati a un aumento della domanda sul margine estensivo (più transazioni, più articoli per transazione) ma una compressione del valore di ciascuna unità venduta in misura superiore all'incremento dei volumi, producendo un effetto netto negativo sul fatturato. Questo risultato è coerente con la natura delle promozioni di prezzo nel retail: abbassare il prezzo aumenta il traffico ma riduce il margine per unità, e l'equilibrio tra questi due effetti determina se la promozione genera o meno valore incrementale netto. L'ATV è la variabile che più direttamente cattura questo meccanismo: una riduzione di quasi un punto percentuale per ogni punto di sconto applicato è coerente con l'ipotesi che i consumatori non aumentino la spesa in risposta alla promozione, ma semplicemente acquistano agli stessi prezzi scontati.

5.1.3 Limiti dell'analisi

Nonostante la significatività statistica delle stime, i risultati devono essere interpretati con cautela. Il modello adottato non controlla per una serie di fattori rilevanti, tra cui la stagionalità del calendario commerciale, le differenze strutturali tra store e le strategie aziendali di pricing. In particolare, la variabile $Discount_{it}$ è verosimilmente endogena: le politiche promozionali non sono assegnate in modo casuale, ma vengono implementate in specifici contesti operativi, spesso caratterizzati da variazioni attese della domanda, introducendo una correlazione tra intensità promozionale e shock non osservati.

Formalmente, la condizione di esogeneità $E[\varepsilon_{it} | Discount_{it}] = 0$ è plausibilmente violata: le politiche promozionali non sono assegnate in modo casuale ma rispondono a condizioni operative che generano correlazione tra lo sconto e il termine di errore.

Di conseguenza i coefficienti stimati non possono essere interpretati come effetti causali dello sconto, ma riflettono una combinazione tra l'effetto diretto della promozione e le dinamiche di selezione delle politiche commerciali.

5.2 Elasticità media, 2WFE

Per superare i limiti della specificazione precedente si introduce un modello con effetti fissi a due vie:

$$\log(Y_{it}) = \beta \cdot Discount_{it} + \alpha_i + \gamma_t + \varepsilon_{it}$$

L'effetto fisso di store α_i assorbe tutte le caratteristiche strutturali invarianti nel tempo del punto vendita (dimensione, localizzazione, format commerciale, composizione della clientela) eliminando il bias dovuto al fatto che store con diversa intensità promozionale tipica differiscono sistematicamente anche nelle loro performance di base. L'effetto fisso di data γ_t cattura ogni shock comune a tutti i negozi in un determinato giorno, inclusi stagionalità, festività e dinamiche macroeconomiche aggregate, impedendo che questi vengano erroneamente attribuiti alla variazione dello sconto. Il coefficiente β identifica quindi l'elasticità dello sconto confrontando lo stesso store con sé stesso in giorni diversi, al netto degli andamenti comuni (una stima within-store nel tempo che è metodologicamente più credibile di quella OLS). Gli errori standard sono clusterizzati a livello di store per tenere conto della correlazione seriale degli errori all'interno di ciascun punto vendita, che violerebbe l'assunzione di

indipendenza delle osservazioni e porterebbe a una sottostima sistematica dell'incertezza se ignorata.

5.2.1 Risultati

Tabella 5.2 - Elasticità media dello sconto sui KPI: stime 2WFE

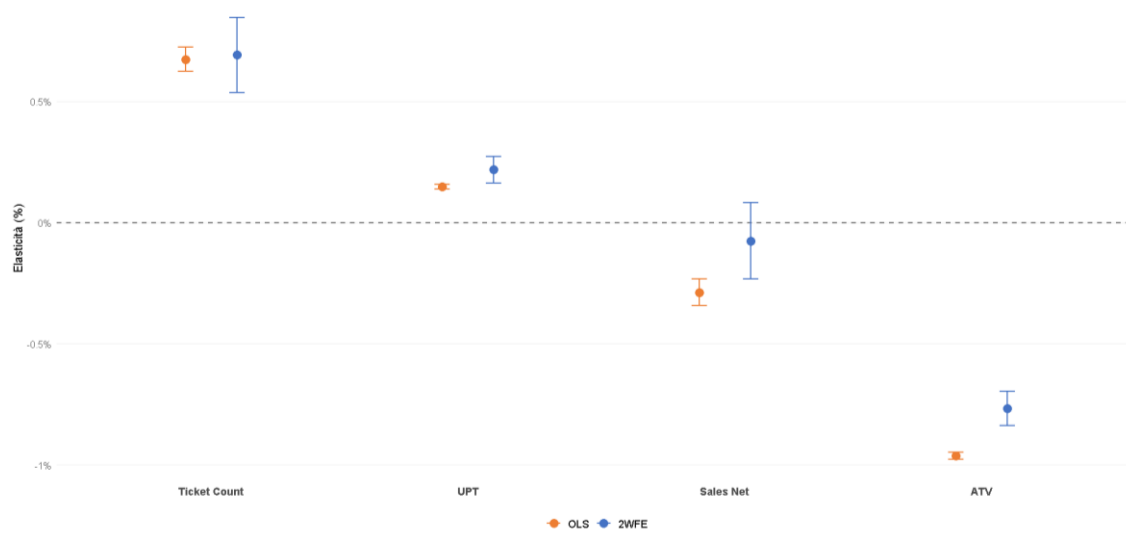
KPI	Elasticità (%)	Std. Error	t-stat	p-value
log(Ticket Count)	+0,692	0,079	8,754	<0,001***
log(UPT)	+0,218	0,028	7,773	<0,001***
log(Sales Net)	-0,076	0,080	-0,944	0,347 ns
log(ATV)	-0,768	0,036	-21,276	<0,001***

*Nota: effetti fissi store (165) e giornalieri (359). Errori standard clusterizzati a livello store. N = 58.181. La colonna Elasticità riporta il coefficiente β . I coefficienti e gli errori standard sono riportati in termini percentuali (stima e errore standard moltiplicati per 100). ns = non significativo; *** $p < 0,001$.*

Il confronto con il modello OLS, mostrato nella Figura 5.1, è il risultato più informativo di questa fase dell'analisi. Per il Ticket, l'effetto dello sconto rimane positivo e stabile (circa +0,69% per ogni punto percentuale di sconto, sostanzialmente invariato rispetto all'OLS) confermando che il segnale sul traffico è robusto e non è un artefatto dell'eterogeneità tra store o dei trend temporali. Anche per l'UPT l'effetto rimane positivo e lievemente più ampio rispetto all'OLS (+0,22%), suggerendo che la stima precedente era se mai leggermente attenuata dalla presenza di eterogeneità non controllata.

Il cambiamento più rilevante riguarda le Sales Net: il coefficiente passa da -0,29% nell'OLS a -0,08% nel 2WFE, mantenendo il segno ma perdendo totalmente intensità nell'effetto e la sua significatività statistica ($p = 0,34$, non significativo). Questo significa che la relazione negativa tra sconto e fatturato osservata nell'OLS era in larga misura un artefatto, rifletteva il fatto che le promozioni vengono attivate in periodi o in store strutturalmente associati a performance di fatturato più deboli, non un effetto causale dello sconto sul valore delle vendite. L'ATV rimane invece negativo e significativo (-0,77%), coerentemente con l'effetto meccanico di riduzione del prezzo medio pagato.

Figura 5.1 - Elasticità media dello sconto sui KPI: confronto OLS e 2WFE



Nota: ogni punto rappresenta il coefficiente β stimato dalla specificazione log-lineare corrispondente, interpretabile come variazione percentuale del KPI per un incremento di +1 punto percentuale dello sconto medio. Le barre verticali rappresentano gli intervalli di confidenza al 95%. Campione L4L, $N = 58.181$.

5.2.2 Interpretazione

L'introduzione degli effetti fissi produce due messaggi distinti a seconda del KPI considerato. Per il Ticket e l'UPT, la stabilità dei coefficienti tra OLS e 2WFE è un segnale di robustezza: l'effetto positivo dello sconto sul volume di transazioni e sulla profondità del carrello non è un artefatto del confondimento, ma riflette una relazione genuina tra intensità promozionale e risposta della domanda. Le promozioni funzionano come strumento di attivazione: aumentano il traffico in negozio e, in misura minore, la quantità acquistata per visita.

Per le Sales Net, la totale perdita di significatività statistica e intensità in modulo del coefficiente racconta invece una storia diversa. La relazione negativa tra sconto e fatturato osservata nell'OLS era in larga misura un artefatto generato dalla selezione temporale endogena delle promozioni. Le politiche di sconto non vengono attivate in modo casuale nel calendario commerciale, ma tendono a essere concentrate nei momenti in cui la domanda attesa è strutturalmente più debole: fine stagione, periodi inter-festivi, fasi di calo fisiologico successivo ai picchi. Questo produce una correlazione negativa spuria tra intensità promozionale e fatturato, non perché lo sconto comprime le vendite, ma perché le vendite erano già depresse prima dell'attivazione della promozione, e la promozione stessa è stata scelta come risposta a quella debolezza. L'effetto fisso temporale γ_t rimuove questa fonte di distorsione assorbendo il

livello medio di tutti gli store in ciascun giorno, elimina la componente di variazione del fatturato che è comune a tutti i negozi e riconducibile alla fase del calendario commerciale, impedendo che venga erroneamente attribuita alla variazione dello sconto. L'effetto fisso di store α_i opera in parallelo su una dimensione diversa: assorbe la selezione cross-sectional, il fatto che store strutturalmente diversi abbiano intensità promozionali tipiche diverse e al tempo stesso performance di base diverse. I due effetti fissi risolvono quindi due problemi di identificazione distinti e complementari. Una volta rimosse queste componenti, il coefficiente collassa verso zero, perdendo sia intensità che significatività statistica confermando che nell'OLS si stava stimando una combinazione tra effetto causale e selezione, non l'effetto causale isolato. Questo risultato acquista una lettura particolarmente netta se si richiama l'identità contabile introdotta nella sezione 3.4: poiché $\log(\text{Sales Net}) = \log(\text{Ticket Count}) + \log(\text{ATV})$, l'effetto netto dello sconto sul fatturato è la somma algebrica dei due effetti stimati separatamente. Nel modello 2WFE, lo sconto aumenta il Ticket del +0,69% e comprime l'ATV del -0,77% per ogni punto percentuale di sconto aggiuntivo: i due effetti si compensano quasi perfettamente, producendo un impatto sulle Sales Net prossimo a zero e statisticamente non distinguibile da esso. Il trade-off tra volume e valore unitario è quindi sostanzialmente in pareggio in media, ma questa media nasconde una forte eterogeneità di contesto che la sezione successiva porta alla luce. L'ATV rimane l'unica variabile monetaria con un effetto robusto, ma questo è attribuibile al meccanismo diretto di abbassamento del prezzo piuttosto che a un cambiamento nel comportamento d'acquisto.

Il quadro complessivo che emerge è quello di promozioni efficaci sul margine estensivo della domanda, capaci di aumentare il numero di transazioni ma non di tradurre sistematicamente questo incremento in maggiore fatturato netto. Questo profilo ha implicazioni dirette per la valutazione della profittabilità delle campagne: un aumento del traffico del +0,69% per ogni punto di sconto applicato deve essere confrontato con la compressione del margine unitario derivante dallo sconto stesso, e il saldo non è necessariamente positivo.

Per dare scala concreta a questi coefficienti, si considera uno store rappresentativo del campione: 31 transazioni medie al giorno, fatturato netto medio di 1.735€, sconto medio del 18%. Una campagna che sposta lo sconto di 5 punti percentuali, dal 18% al 23%, è associata a un incremento atteso del Ticket di circa il 3,5% ($5 \times 0,69\%$), pari a poco più di una transazione aggiuntiva al giorno. Parallelamente, l'ATV si riduce di circa il 3,8% ($5 \times 0,77\%$), passando da 56€ a circa 54€ per scontrino. Il fatturato netto risultante è sostanzialmente invariato: il volume aggiuntivo e la compressione del valore si annullano a vicenda. Questo calcolo illustra in modo

concreto perché l'effetto medio sulle Sales Net sia non significativo, secondo le associazioni stimate dal modello 2WFE: non perché lo sconto sia inerte, ma perché i due meccanismi che lo compongono, più traffico e meno valore per transazione, operano in direzioni opposte con intensità quasi identica.

Vale la pena notare che questa neutralità sul fatturato netto non è neutralità economica: lo sconto aggiuntivo riduce il margine lordo per unità venduta, e l'incremento di traffico non genera ricavi sufficienti a compensare questa riduzione. In media, quindi, una campagna promozionale che agisce esclusivamente sulla leva dello sconto tende a essere marginalmente distruttiva di valore, a meno che non operi in contesti in cui l'elasticità del Ticket sia sufficientemente superiore alla media. Identificare questi contesti è precisamente l'obiettivo della sezione successiva.

5.3 Eterogeneità dell'elasticità per contesto strutturale

L'elasticità media stimata nei modelli precedenti è una misura aggregata che, per costruzione, ignora la possibilità che la risposta della domanda allo sconto vari sistematicamente in funzione del contesto operativo. Un effetto medio di +0,69% sul Ticket potrebbe nascondere combinazioni di store, area geografica e periodo in cui l'effetto è quasi nullo e altre in cui è molto più ampio (informazione che è direttamente rilevante per la pianificazione delle campagne promozionali). Per catturare questa eterogeneità, il modello viene esteso introducendo interazioni tra lo sconto e le principali dimensioni di contesto disponibili nel dataset:

$$\log(Y_{it}) = \beta_0 \cdot Discount_{it} + \sum_k \beta_k (Discount_{it} \times Z_{k,it}) + \alpha_i + \gamma_t + \varepsilon_{it}$$

dove $Z_{k,it}$ include tipologia di store, area geografica e giorno della settimana. Come discusso nella sezione metodologica, la struttura delle interazioni adottata è di terzo ordine (coinvolge simultaneamente tutte e tre le dimensioni) poiché l'ipotesi che le tre fonti di eterogeneità agiscano in modo additivo e indipendente sarebbe eccessivamente restrittiva.

5.3.1 Struttura del campione

Prima di presentare le stime, è opportuno segnalare che le celle del disegno sperimentale non sono bilanciate. La Tabella 5.3 riporta la distribuzione delle osservazioni per regione, tipologia e giorno della settimana.

Tabella 5.3 - Frequenza assoluta delle osservazioni per cella

	Mall	Urban	Outlet	Totale
Centro Feriale	6.862	3.055	512	10.429
Nord Feriale	9.692	3.071	2.547	15.310
Sud Feriale	11.416	4.548	255	16.219
Centro Weekend	2.712	1.199	206	4.117
Nord Weekend	3.862	1.236	1.028	6.126
Sud Weekend	4.533	1.344	103	5.980
Totale	39.077	14.453	4.651	58.181

Nota: campione L4L, N = 58.181. Le celle Outlet | Sud presentano numerosità sensibilmente inferiori rispetto alle altre (255 osservazioni nel feriale, 103 nel weekend), il che limita la generalizzabilità delle stime per quelle combinazioni.

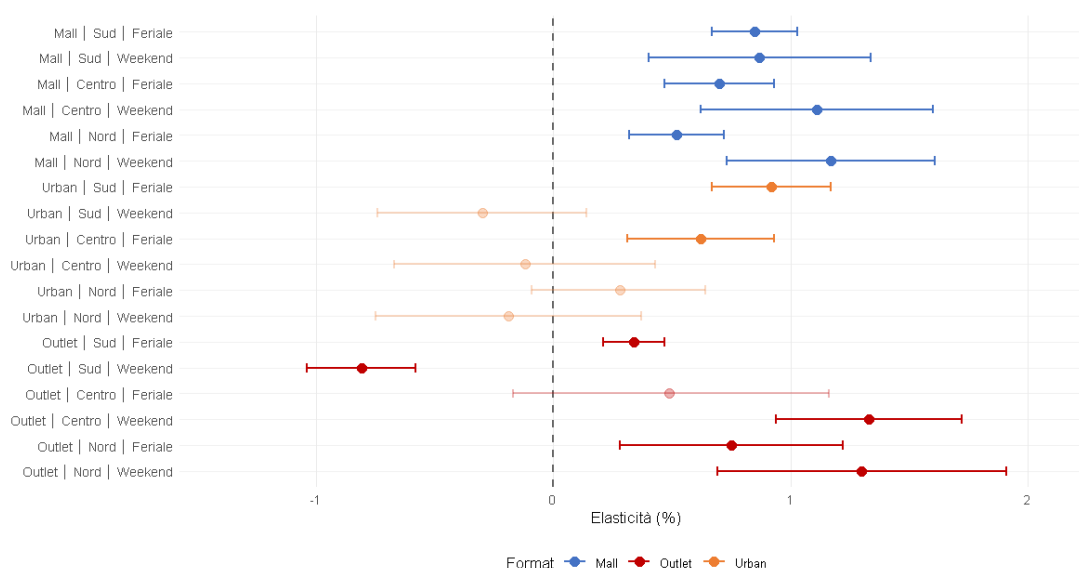
Lo sbilanciamento più rilevante riguarda le celle Outlet|Sud|Weekend (N = 103) che presentano una numerosità sensibilmente inferiore rispetto alle altre. Le stime per queste combinazioni sono statisticamente precise in senso convenzionale, ma potenzialmente sensibili alla composizione specifica degli store inclusi nel campione. I risultati relativi agli Outlet del Sud e del Centro vanno pertanto interpretati con la cautela dovuta a una base empirica ristretta.

5.3.2 Risultati

I risultati del modello con interazioni di terzo ordine sono presentati in termini di elasticità calcolate per ciascuna delle 18 combinazioni di contesto, tre tipologie di store (Mall, Urban, Outlet), tre aree geografiche (Sud, Centro, Nord) e due fasce temporali (feriale, weekend), ottenute sommando il coefficiente base e tutti i coefficienti di interazione rilevanti, con errori standard corretti per propagazione della varianza.

Per ciascun KPI la discussione si appoggia alla rappresentazione grafica delle elasticità con i relativi intervalli di confidenza al 95%; i valori puntuali completi, con errori standard composti, intervalli di confidenza al 95% e livelli di significatività, sono riportati nelle Tabelle A.1-A.4 in Appendice.

Figura 5.2 - Elasticità del Ticket Count per contesto operativo



Nota: stime del modello 2WFE con interazioni. Ogni punto rappresenta l'elasticità marginale del Ticket Count rispetto allo sconto per una specifica combinazione di contesto, calcolata come somma del coefficiente base e dei coefficienti di interazione rilevanti.

Le barre orizzontali rappresentano gli intervalli di confidenza al 95%, costruiti via delta method sulla matrice di varianza-covarianza clusterizzata a livello store. I punti trasparenti indicano stime non significative al 5% ($p > 0,05$), da interpretare come non distinguibili da zero. Campione L4L, $N = 58.181$.

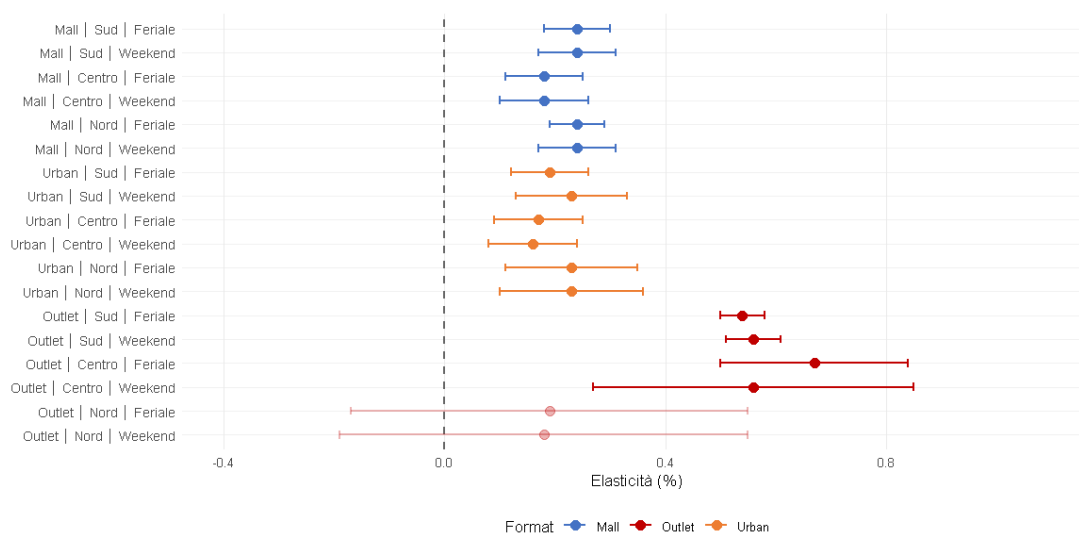
Ticket Count. Per i Mall lo sconto genera traffico positivo e significativo in tutte le combinazioni, sia nel feriale sia nel weekend, in tutte e tre le aree geografiche. L'elemento più rilevante è l'effetto del giorno della settimana. Al Nord gli IC di feriale [+0,32%; +0,72%] e weekend [+0,73%; +1,61%] non si sovrappongono: il weekend produce un'elasticità significativamente più alta e la differenza è statisticamente dimostrabile. Al Centro gli IC si sovrappongono nella fascia +0,62%/+0,93%, rendendo la differenza indicativa ma non conclusiva. Al Sud le stime di feriale (+0,85%) e weekend (+0,87%) sono quasi identiche con IC ampiamente sovrapposti: non c'è effetto weekend nei Mall del Sud. Il messaggio è quindi che il weekend amplifica l'elasticità nei Mall in modo robusto al Nord, possibile al Centro, assente al Sud.

Per gli Urban il quadro si divide nettamente lungo due assi. Nel feriale, l'elasticità è positiva e significativa al Sud (+0,92%) e al Centro (+0,62%), mentre al Nord la stima è non significativa e va trattata come non distinguibile da zero: i dati non consentono di concludere che lo sconto generi traffico negli store Urban del Nord nel feriale. Nel weekend, le stime sono non significative in tutte e tre le aree, al Sud (-0,30% ns), al Centro (-0,12% ns) e al Nord (-0,19% ns), e vanno quindi trattate come non distinguibili da zero indipendentemente dal segno. Lo

sconto non genera traffico misurabile negli store Urban nel weekend in nessuna area geografica. Vale la pena notare che per Urban | Sud il contrasto tra feriale significativamente positivo (+0,92%) e weekend non distinguibile da zero è il più ampio e netto dell'intera tipologia, con IC che non si sovrappongono: la differenza feriale/weekend al Sud è statisticamente dimostrabile.

Per gli Outlet la struttura è la più articolata. Nel feriale, Outlet | Nord è positivo e significativo (+0,75%), Outlet | Sud è positivo e significativo ma di entità più contenuta (+0,34%), mentre Outlet | Centro non raggiunge la significatività (ns) e va trattato come non distinguibile da zero, l'SE elevato (0,338) segnala scarsa potenza statistica per quella cella, probabilmente legata alla limitata numerosità. Nel weekend il quadro si polarizza: Outlet | Centro (+1,33%) e Outlet | Nord (+1,30%) mostrano le elasticità più alte dell'intera Tabella A.1 con IC che escludono chiaramente lo zero, mentre Outlet | Sud | Weekend (-0,81%**) è l'unico caso con elasticità negativa e significativa, risultato che verrà interpretato con la cautela necessaria nella sezione 5.3.3.

Figura 5.3 - Elasticità dell'UPT per contesto operativo



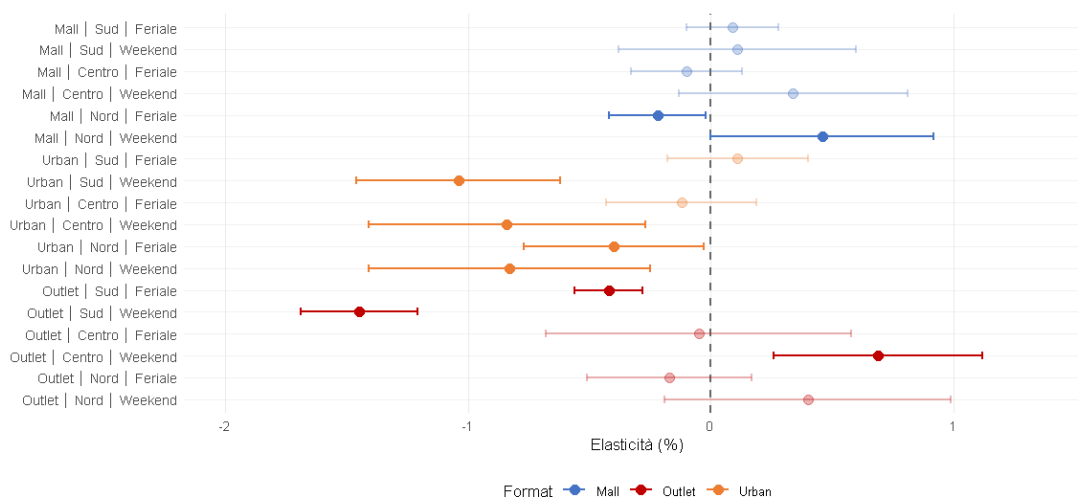
Nota: stime del modello 2WFE con interazioni. Ogni punto rappresenta l'elasticità marginale dell'UPT rispetto allo sconto per una specifica combinazione di contesto, calcolata come somma del coefficiente base e dei coefficienti di interazione rilevanti. Le barre orizzontali rappresentano gli intervalli di confidenza al 95%, costruiti via delta method sulla matrice di varianza-covarianza clusterizzata a livello store. I punti trasparenti indicano stime non significative al 5% ($p > 0,05$), da interpretare come non distinguibili da zero. Campione L4L, $N = 58.181$.

UPT. L'UPT è il KPI con la struttura più ordinata. Per Mall e Urban l'elasticità è positiva, significativa e sostanzialmente uniforme in tutte le combinazioni, con valori compresi tra +0,16% e +0,24% e IC stretti. Non si osserva alcun effetto del giorno della settimana né dell'area

geografica: gli IC di feriale e weekend si sovrappongono ovunque. Per Mall e Urban lo sconto produce un effetto sull'UPT stabile, prevedibile e indipendente dal contesto.

Per gli Outlet la struttura si differenzia. Al Sud e al Centro le elasticità sono significativamente più alte (tra +0,54% e +0,67%), con IC che non si sovrappongono a quelli di Mall e Urban: la differenza tra tipologie è robusta e dimostrabile. Al Nord invece le stime (intorno a +0,18%/+0,19%) sono non significative con standard error molto elevati (0,183 e 0,188) che segnalano scarsa potenza statistica per quella cella, probabilmente legata alla composizione specifica degli store Outlet del Nord nel campione.

Figura 5.4 - Elasticità delle Sales Net per contesto operativo



Nota: stime del modello 2WFE con interazioni. Ogni punto rappresenta l'elasticità marginale delle Sales Net rispetto allo sconto per una specifica combinazione di contesto, calcolata come somma del coefficiente base e dei coefficienti di interazione rilevanti.

Le barre orizzontali rappresentano gli intervalli di confidenza al 95%, costruiti via delta method sulla matrice di varianza-covarianza clusterizzata a livello store. I punti trasparenti indicano stime non significative al 5% ($p > 0,05$), da interpretare come non distinguibili da zero. Campione L4L, $N = 58.181$.

Sales Net. Le Sales Net sono il KPI più rilevante per valutare l'effetto economico netto dello sconto e mostrano la maggiore eterogeneità tra combinazioni.

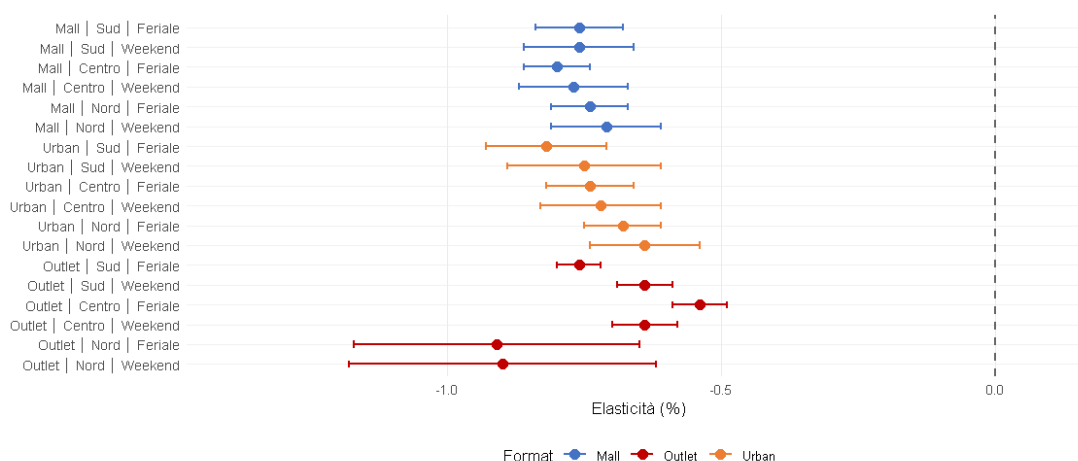
Per i Mall, tutte le combinazioni tranne due non sono significative e vanno trattate come non distinguibili da zero. Le eccezioni sono Mall | Nord | Feriale (-0,22%*) e Mall | Nord | Weekend (+0,46%·). Il primo è significativo al 5% con segno negativo; il secondo è significativo solo al 10% con segno positivo, e gli IC dei due si sovrappongono ampiamente: non è possibile affermare che le due stime siano diverse tra loro. Nel complesso, per i Mall l'effetto dello sconto

sul fatturato netto è nullo o non rilevabile nella grande maggioranza dei contesti: il traffico generato e la compressione del prezzo si compensano.

Per gli Urban il quadro è il più coerente e interpretabile dell'intera sezione. Nel feriale, le Sales Net sono non significative al Sud (+0,11% ns) e al Centro (-0,12% ns), mentre sono significativamente negative al Nord (-0,40%). Nel weekend, le Sales Net sono significativamente negative in tutte e tre le aree: -1,04%** al Sud, -0,84%** al Centro, -0,83%** al Nord. Il pattern è chiaro: nel weekend lo sconto erode il fatturato netto degli store Urban in modo sistematico e robusto in tutta la penisola, mentre nel feriale l'erosione è presente solo al Nord. Gli IC del weekend non si sovrappongono a quelli del feriale al Sud, dove la differenza tra +0,11% ns e -1,04%** è la più netta di tutta la sezione, mentre si sovrappongono parzialmente al Centro e al Nord.

Per gli Outlet il contrasto geografico nel weekend è il risultato più estremo di tutta la sezione. Outlet | Sud | Weekend (-1,45%) e Outlet | Centro | Weekend (+0,69%) hanno IC che non si sovrappongono e segni opposti: questa è la differenza più robusta e più ampia osservabile tra due celle comparabili, tuttavia è facilmente riconducibile ai problemi precedentemente anticipati riguardo la base campionaria degli Outlet | Sud. Outlet | Nord | Weekend (+0,40% ns) non è significativo. Nel feriale, Outlet | Sud è significativamente negativo (-0,42%**), mentre Centro e Nord non sono significativi.

Figura 5.5 - Elasticità dell'ATV per contesto operativo



Nota: stime del modello 2WFE con interazioni. Ogni punto rappresenta l'elasticità marginale dell'ATV rispetto allo sconto per una specifica combinazione di contesto, calcolata come somma del coefficiente base e dei coefficienti di interazione rilevanti. Le barre orizzontali rappresentano gli intervalli di confidenza al 95%, costruiti via delta method sulla matrice di varianza-covarianza clusterizzata a livello store. I punti trasparenti indicano stime non significative al 5% ($p > 0,05$), da interpretare come non distinguibili da zero. Campione L4L, $N = 58.181$.

ATV. L'ATV è il KPI con la struttura più uniforme: l'effetto è negativo e significativo in tutte le 18 combinazioni senza eccezioni. La compressione del prezzo medio è un effetto meccanico diretto dello sconto che non dipende dal contesto operativo. Le variazioni tra combinazioni, da $-0,54\%$ a $-0,91\%$, sono statisticamente distinguibili grazie alla precisione delle stime, ma economicamente contenute e non cambiano la direzione del messaggio.

Non si osserva alcun effetto sistematico del giorno della settimana: gli IC di feriali e weekend si sovrappongono in tutte le tipologie. L'unica anomalia rilevante riguarda gli Outlet del Nord ($-0,91\%$ e $-0,90\%$), con IC che non si sovrappongono a quelli degli Outlet del Sud e del Centro: l'ATV negli Outlet del Nord è compresso dallo sconto in misura significativamente superiore rispetto agli altri Outlet.

L'uniformità dell'ATV offre un criterio di lettura trasversale per le altre tabelle: poiché la compressione del prezzo medio è stabile intorno a $-0,70/-0,80\%$ in quasi tutti i contesti, l'effetto netto sulle Sales Net dipende interamente da quanto l'elasticità del Ticket supera questa soglia. Le combinazioni in cui il Ticket è significativamente positivo e le Sales Net sono non significative sono quelle in cui i due effetti si compensano. Le combinazioni in cui il Ticket non è significativo e le Sales Net sono negative sono quelle in cui la compressione del prezzo non trova alcuna compensazione volumetrica, il caso operativamente più critico.

5.3.3 Discussione di un risultato anomalo

Un risultato che richiede particolare cautela interpretativa riguarda la combinazione Outlet | Sud | Weekend, per la quale l'elasticità del Ticket è $-0,81\%^{***}$ con IC $[-1,04\%; -0,58\%]$ che esclude chiaramente lo zero. Numericamente è la stima più estrema della sezione, ma la sua interpretazione economica è fragile per ragioni strutturali. La stima è identificata da sole 103 osservazioni concentrate su un unico store, la cella più piccola di tutto il disegno sperimentale, all'interno di una specificazione con interazioni di terzo ordine. Un SE di 0,119 su una stima di $-0,81\%$ non è inusuale in contesti con poca variazione identificante, ma il fatto che l'intera inferenza poggia su pochissimi punti vendita rende il coefficiente potenzialmente specifico di quelle osservazioni piuttosto che generalizzabile.

Il meccanismo più plausibile rimane l'endogeneità legata alla gestione delle scorte: gli Outlet tendono ad aumentare l'intensità dello sconto nelle fasi avanzate di liquidazione, quando la disponibilità assortimentale è già ridotta. Lo sconto cresce mentre il numero di prodotti

disponibili diminuisce, producendo una correlazione negativa tra sconto e transazioni che riflette il ciclo di vita del magazzino più che la risposta della domanda al prezzo. Questa lettura è coerente con il fatto che l'UPT nella stessa combinazione sia positivo e significativo (+0,56%***): chi entra acquista più articoli per transazione, il che suggerisce presenza di domanda ma scarsità dell'offerta.

5.3.4 Interpretazione

Il profilo che emerge dalle quattro tabelle si organizza in tre gruppi di contesti con implicazioni operative distinte.

Contesti in cui lo sconto non genera traffico e comprime il fatturato: ridurre l'intensità promozionale è razionale. Il caso più chiaro e robusto è quello degli store Urban nel weekend. Il Ticket non è significativo in nessuna area geografica, lo sconto non genera traffico misurabile nel weekend per questo formato, mentre le Sales Net sono significativamente negative in tutte e tre le aree (-1,04%*** al Sud, -0,84%** al Centro, -0,83%** al Nord). La combinazione di traffico assente e fatturato in erosione indica che nei weekend Urban lo sconto trasferisce valore ai clienti che sarebbero entrati comunque, senza espandere la base di visitatori. In questi contesti ridurre l'intensità promozionale nel weekend avrebbe il duplice effetto di contenere il costo della promozione e di attenuare l'erosione del fatturato netto, senza sacrificare traffico incrementale che i dati non rivelano. Un test pilota di riduzione controllata in un sottoinsieme di store Urban nel weekend fornirebbe evidenza diretta e consentirebbe di validare questa indicazione prima di un'implementazione estesa.

Un profilo simile, seppur meno estremo, caratterizza Urban | Nord | Feriale: il Ticket non è significativo e le Sales Net sono negative (-0,40%*). Anche nel feriale al Nord gli store Urban mostrano uno sconto che non genera traffico ma erode il fatturato.

Contesti in cui lo sconto genera traffico ma non fatturato: mantenere i livelli attuali senza intensificare, valutare leve complementari. I Mall nel feriale mostrano elasticità del Ticket positive e significative in tutte le aree (tra +0,52%*** e +0,85%***), ma Sales Net non significative ovunque tranne che per il debole segnale negativo al Nord. Lo sconto funziona come strumento di attivazione del traffico, ma la compressione del prezzo medio assorbe integralmente il volume aggiuntivo senza produrre crescita di fatturato. In questi contesti intensificare ulteriormente lo sconto aumenterebbe il costo promozionale senza generare ricavi

aggiuntivi. La logica di intervento più appropriata non è modificare il livello di sconto ma affiancare strumenti che aumentino il valore della transazione a parità di traffico (cross-selling, bundle di prodotto, soglie di spesa minima) senza ulteriore compressione del prezzo unitario.

Urban | Sud | Feriale (+0,92%*** Ticket, Sales Net ns) e Urban | Centro | Feriale (+0,62%*** Ticket, Sales Net ns) condividono lo stesso profilo: lo sconto genera traffico reale ma il saldo sul fatturato è neutro. Anche qui la direzione non è intensificare lo sconto ma sfruttare il traffico generato attraverso leve di valore che non comprimano il prezzo medio.

Contesti ad alta elasticità del Ticket con potenziale per campagne promozionali mirate. I Mall nel weekend al Nord (+1,17%) e al Centro (+1,11%) mostrano le elasticità sul Ticket più alte tra tutti i formati non-Outlet, con IC che al Nord non si sovrappongono a quelli del feriale, la differenza è statisticamente dimostrabile. Le Sales Net non sono significative, quindi il saldo tra volume e compressione del prezzo è neutro. Questi sono i contesti in cui campagne promozionali specificamente concentrate nel weekend potrebbero generare il massimo traffico incrementale. Non si tratta necessariamente di aumentare lo sconto medio annuale, ma di concentrare comunicazioni promozionali, offerte temporanee o eventi commerciali nel weekend in queste aree, sfruttando una clientela che il sabato e la domenica risponde più intensamente all'incentivo di prezzo rispetto ai giorni feriali. Mall | Sud | Feriale (+0,85%***) è un contesto ad alta elasticità stabile anche nel feriale, con Sales Net neutre, che si presta a iniziative promozionali infrasettimanali mirate.

Urban | Sud | Feriale (+0,92%) e Urban | Centro | Feriale (+0,62%) presentano elasticità feriali elevate con Sales Net neutre: contesti in cui promozioni infrasettimanali, pensate per i giorni di minore affluenza strutturale, potrebbero generare traffico incrementale significativo senza erodere il fatturato.

Per gli Outlet una nota di cautela è necessaria prima di qualsiasi raccomandazione. Outlet | Centro | Weekend (+1,33%*** Ticket, +0,69%** Sales Net) è l'unica combinazione in tutta la sezione in cui lo sconto produce un incremento significativo sia del traffico sia del fatturato netto. Outlet | Nord | Weekend (+1,30%*** Ticket) mostra la stessa elasticità sul traffico ma con Sales Net non significative. Queste sono le elasticità assolute più alte dell'intera analisi, e documentano che la clientela degli Outlet risponde intensamente allo sconto nel weekend. Tuttavia, il livello di sconto strutturale degli Outlet è già nell'ordine del 37-42% medio, ben al di sopra degli altri formati, e un ulteriore incremento su base già così elevata comporta rischi di margine che i dati aggregati non consentono di quantificare. L'evidenza disponibile descrive una

clientela altamente price-sensitive in questi contesti, ma non implica che aumentare ulteriormente sconti già molto profondi sia commercialmente conveniente. Eventuali interventi sugli Outlet dovrebbero essere valutati in termini di composizione dell'offerta e gestione dell'assortimento piuttosto che di pura intensità dello sconto.

5.4 Eterogeneità dell'elasticità per contesto promozionale

L'analisi condotta nella sezione precedente ha documentato come l'elasticità allo sconto vari sistematicamente in funzione del contesto operativo: format, area geografica e giorno della settimana. Una seconda dimensione di eterogeneità, distinta dalla prima e ad essa complementare, riguarda il tipo di promozione attiva.

Per analizzare questa dimensione si introduce una classificazione delle osservazioni in undici categorie promozionali, costruite aggregando le dummy elementari del dataset. Le categorie sono:

- **Saldi stagionali.** Saldi invernali ed estivi, disaggregati nelle tre fasi previste dal calendario commerciale (Fase 1, Fase 2, Fase 3) con intensità di sconto crescente al progredire della stagione. Rappresentano il regime promozionale a maggiore copertura temporale e di rete.
- **Mid Season Sale (MSS).** campagne promozionali di medio periodo collocate tra una stagione commerciale e la successiva, con l'obiettivo di sostenere il traffico nei periodi di domanda strutturalmente più debole.
- **Pre-Saldi.** iniziative che anticipano formalmente l'avvio dei saldi, rivolte a stimolare la domanda nelle settimane precedenti al periodo promozionale principale, spesso con accesso anticipato riservato a segmenti specifici di clientela.
- **Black Friday.** evento promozionale concentrato in un numero limitato di giorni, caratterizzato da alta intensità di sconto e forte pressione comunicativa a livello di mercato.
- **Evento Calendario.** campagne legate a specifiche ricorrenze commerciali o occasioni di consumo, come San Valentino, Festa della Mamma o il Single's Day, con sconto strutturalmente più contenuto rispetto alle altre categorie.

- **Cashback.** programmi di rimborso parziale sul valore d'acquisto, condizionati al raggiungimento di una soglia minima di spesa. A differenza degli sconti diretti, il beneficio è differito rispetto al momento dell'acquisto.
- **Liquidazione.** attività finalizzate allo smaltimento di stock residuo o articoli di fine serie, generalmente attivate in modo selettivo su specifici store o categorie di prodotto e scollegate dal calendario promozionale ordinario.
- **Promozioni a soglia e multi-acquisto.** iniziative che subordinano il beneficio promozionale al raggiungimento di un importo minimo di spesa o all'acquisto di più articoli, con l'obiettivo di aumentare il valore medio della transazione oltre che il traffico.
- **Partnership e coupon.** campagne realizzate in collaborazione con partner commerciali esterni o attraverso la distribuzione di voucher, attive su un sottoinsieme della rete e con logiche di attivazione eterogenee.

I giorni privi di qualsiasi promozione attiva costituiscono la categoria di riferimento (baseline).

5.4.1 Impostazione metodologica

Il modello stimato è un 2WFE con interazioni tra lo sconto medio giornaliero e i flag di categoria:

$$\log(Y_{it}) = \alpha_i + \gamma_t + \beta_0 \cdot \text{Discount}_{it} + \sum_k \beta_k (\text{Discount}_{it} \times \text{flag}_{k,it}) + \varepsilon_{it}$$

dove α_i e γ_t sono gli effetti fissi di store e data, β_0 cattura l'elasticità marginale dello sconto nella baseline, e i coefficienti β_k misurano il differenziale rispetto alla baseline per ciascuna categoria. Gli errori standard sono clusterizzati a livello di store. Le elasticità totali per ciascuna categoria, $\beta_0 + \beta_k$, sono calcolate con propagazione dell'errore via delta method. Il modello non include le dummy di categoria come regressori separati, solo come interazioni con lo sconto: gli effetti fissi di data γ_t assorbono già tutto ciò che è comune a tutti gli store in un determinato giorno, incluso il fatto che quel giorno sia Black Friday, saldi o no-promo. Aggiungere le dummy di categoria separatamente introdurrebbe multicollinearità con γ_t , poiché le categorie promozionali variano nel tempo ma non tra store nello stesso giorno. Il livello medio di ciascun regime è quindi già assorbito dagli effetti fissi, e quello che il modello stima è esclusivamente l'elasticità marginale dello sconto all'interno di ciascun regime, sfruttando la variazione cross-store di Discount nello stesso giorno e nella stessa categoria.

Questa distinzione ha una conseguenza interpretativa precisa: le elasticità stimate in questa sezione non rispondono alla domanda quale campagna è più efficace? La domanda a cui il modello risponde è più specifica e più utile dal punto di vista operativo: dato che un determinato regime promozionale è già attivo, quanto vale un punto percentuale di sconto aggiuntivo rispetto al livello corrente? Si tratta di una misura del potenziale di crescita marginale intra-regime, ovvero di quanto margine residuo esiste per intensificare ulteriormente la leva dello sconto all'interno di ciascuna campagna.

Non tutti i regimi promozionali presenti nel dataset sono ugualmente interpretabili attraverso questa logica. Affinché l'elasticità marginale possa essere letta come risposta allo sconto all'interno del regime promozionale, è necessario che il regime sia attivo su una quota di store prossima al 100% nei giorni in cui è presente. In questa situazione, l'effetto fisso temporale γ_t assorbe il livello medio del sistema. L'identificazione del coefficiente non deriva quindi dal confronto tra giorni appartenenti a regimi diversi, ma esclusivamente dalla variabilità residua dello sconto tra store che operano nello stesso giorno e sotto il medesimo regime promozionale.

Questa condizione è soddisfatta dai Saldi, dal Mid Season Sale, dai Pre-Saldi, dalle Evento Calendario e dal Black Friday. L'interpretazione economica è particolarmente naturale: si tratta infatti di campagne che l'azienda realizza comunque in momenti prestabiliti del calendario commerciale. La decisione rilevante non riguarda quindi l'attivazione della promozione, ma la calibrazione della sua intensità. La domanda analitica diventa: dato che il regime promozionale è già in corso, esiste ancora margine per aumentare lo sconto ottenendo traffico incrementale sufficiente a giustificare il costo?

Le restanti categorie presentano invece coperture significativamente inferiori: Cashback (84,3% medio, con circa un terzo dei giorni sotto l'80%), Promo Soglia (64,3%), Coupon/Partnership (43,3%) e Liquidazione (5,8%). Questa differenza genera una logica identificativa diversa, infatti in questi casi una quota rilevante della variazione utilizzata dal modello deriva dal confronto tra store esposti e non esposti alla promozione nello stesso giorno, confondendo l'effetto dell'intensità dello sconto ed effetto dell'attivazione stessa della campagna. Per questa ragione tali categorie vengono mantenute nel modello come controlli, al fine di evitare distorsioni da omissione di contesti promozionali rilevanti, ma i relativi coefficienti non vengono interpretati sostanzialmente nelle analisi successive.

5.4.2 Variazione identificante e qualità dell'identificazione

Prima di presentare le stime è opportuno documentare la variazione dello sconto all'interno di ciascun regime, poiché questa determina direttamente la qualità dell'identificazione econometrica. Un regime in cui tutti gli store applicano esattamente lo stesso sconto non offre variazione cross-store su cui identificare il coefficiente, e la stima risultante avrà errori standard molto ampi e scarso potere statistico.

Tabella 5.4 - Variazione dello sconto dentro ogni regime

Regime	N oss.	AVG Discount (%)	SD (pp)	IQR (pp)	P10 (%)	P90 (%)
Mid Season Sale	9.037	13,1	11,5	9,7	3,7	27,6
Evento Calendario	8.570	12,0	9,9	10,4	2,5	27,0
Saldi Fase 3	8.268	28,9	9,3	11,0	18,3	41,0
Saldi Fase 1	6.171	30,2	8,6	12,2	19,7	40,1
Pre-Saldi	7.599	21,8	8,6	9,2	9,4	31,0
Saldi Fase 2	4.546	31,5	7,7	10,0	22,4	40,9
Black Friday	815	22,6	7,4	6,4	16,2	31,2

Nota: SD = deviazione standard cross-store dello sconto nei giorni in cui il regime è attivo. IQR = intervallo interquartile. Valori più alti indicano maggiore variazione identificante. Il regime Evento Calendario presenta una numerosità così elevata grazie al prolungamento delle stesse attività promozionali.

La Tabella 5.4 rivela un elemento importante per l'interpretazione dei risultati. Il Black Friday presenta la deviazione standard più bassa tra tutti i regimi (7,4 pp) e con 815 osservazioni complessive. Questa combinazione di bassa variazione intra-regime e numerosità limitata ci porta ad aspettarci IC più grandi per le stime riguardanti il Black Friday rispetto alle altre categorie prese in considerazione. Il Mid Season Sale si colloca all'estremo opposto: 11,5 pp di deviazione standard su 56 giorni e oltre 9.000 osservazioni, offrendo la base identificativa più solida tra tutti i regimi.

5.4.3 Statistiche descrittive per regime

Prima di presentare le elasticità marginali è necessario documentare le caratteristiche strutturali di ciascun regime, poiché il livello di partenza dei KPI determina il significato economico assoluto di qualsiasi variazione percentuale.

Tabella 5.5 - Statistiche descrittive per regime promozionale

Regime	AVG Ticket	AVG Sales Net (€)	AVG Discount (%)	ATV (€)	UPT
Black Friday	62,1	4.029	22,6	62,3	1,43
Saldi Fase 1	58,1	2.797	30,2	48,6	1,42
Evento Calendario	47,2	2.507	12,0	52,4	1,42
Pre-Saldi	45,6	2.408	21,8	52,0	1,42
Saldi Fase 2	43,9	2.165	31,5	51,1	1,51
Saldi Fase 3	37,2	1.915	28,9	52,3	1,48
Mid Season Sale	34,1	2.240	13,1	65,6	1,51

Nota: i regimi sono ordinati per AVG Ticket decrescente.

La Tabella 5.5 mostra come i regimi con più alto traffico assoluto siano Black Friday (62,1 ticket/giorno) e Saldi Fase 1 (58,1). È inoltre evidente una progressiva riduzione del traffico lungo il ciclo dei saldi. Il Mid Season Sale e il Pre-Saldi presentano invece sconto medio contenuto (13,1% e 21,8% rispettivamente) con traffico relativamente più basso. L'Evento Calendario mostra il profilo più anomalo: traffico alto (47,2 ticket) con sconto strutturalmente basso (12,0%), a suggerire che in questi giorni il traffico sia già elevato per ragioni legate alla ricorrenza commerciale indipendentemente dall'intensità dello sconto.

5.4.4 Risultati elasticità within-regime

Ticket Count. La dispersione delle elasticità tra i regimi interpretabili è sostanziale: il rapporto tra l'elasticità massima (Mid Season Sale, +0,95%) e quella minima tra le significative (Saldi Fase 1, +0,40%) è di circa 2,4 volte. Questo range indica che il potenziale di crescita marginale varia in misura economicamente rilevante in funzione del contesto promozionale in cui lo sconto viene applicato.

Tabella 5.6 - Elasticità marginali within-regime con IC

Regime	Ticket Count	UPT	Sales Net	ATV
Mid Season Sale	+0,95%*** [+0,72; +1,18]	+0,16%*** [+0,08; +0,24]	+0,08% ns [-0,16; +0,32]	-0,87%*** [-0,96; -0,78]
Pre-Saldi	+0,78%*** [+0,51; +1,04]	+0,19%*** [+0,11; +0,26]	+0,00% ns [-0,24; +0,24]	-0,78%*** [-0,88; -0,67]
Saldi Fase 3	+0,67%*** [+0,46; +0,88]	+0,30%*** [+0,20; +0,39]	-0,21%· [-0,42; +0,01]	-0,88%*** [-0,99; -0,77]
Saldi Fase 2	+0,59%*** [+0,30; +0,87]	+0,07% ns [-0,05; +0,19]	-0,53%*** [-0,81; -0,25]	-1,12%*** [-1,25; -0,99]
Saldi Fase 1	+0,40%** [+0,11; +0,69]	+0,09% ns [-0,04; +0,22]	-0,56%*** [-0,83; -0,28]	-0,96%*** [-1,07; -0,84]
Evento Calendario	+0,24%· [-0,01; +0,50]	+0,24%*** [+0,16; +0,31]	-0,22%· [-0,48; +0,04]	-0,46%*** [-0,56; -0,36]
Black Friday	+0,19% ns [-0,34; +0,72]	+0,06% ns [-0,09; +0,22]	-1,03%*** [-1,61; -0,44]	-1,22%*** [-1,42; -1,01]

*Nota: *** $p < 0,001$; ** $p < 0,01$; * $p < 0,05$; · $p < 0,10$; ns = non significativo. Errori standard clusterizzati a livello store, propagati via delta method.*

Un elemento di cautela è necessario prima di confrontare le stime puntuali: gli intervalli di confidenza di regimi adiacenti nella classifica si sovrappongono in misura significativa. In particolare, gli IC di Mid Season Sale [+0,72; +1,18] e Pre-Saldi [+0,51; +1,04] si sovrappongono nella fascia +0,72/+1,04, rendendo non formalmente distinguibili le due elasticità sulla base dei soli dati disponibili. Lo stesso vale per Pre-Saldi e Saldi Fase 3, i cui IC si sovrappongono quasi interamente. Questo non invalida le stime puntuali come indicazioni ordinali, ma impone cautela nel trarre conclusioni nette su quale regime specifico presenti il potenziale maggiore all'interno di fasce adiacenti. Le differenze sono invece più nette agli estremi della distribuzione: l'IC del Mid Season Sale non si sovrappone a quello di Saldi Fase 1, e nessuno degli IC dei regimi significativi si sovrappone a quello del Black Friday, che include lo zero.

UPT. L'effetto dello sconto sull'UPT è positivo e significativo nella maggior parte dei regimi, con valori compresi tra +0,07% (Saldi Fase 2, ns) e +0,30% (Saldi Fase 3). Il Black Friday e i Saldi

Fase 1 e 2 mostrano UPT non distinguibile da zero: lo sconto aggiuntivo in questi regimi non modifica la profondità del carrello. L'Evento Calendario costituisce un caso a parte: nonostante il segnale sul Ticket sia debole (+0,24%), l'UPT è positivo e significativo (+0,24%***), il che suggerisce che chi entra in negozio in questi giorni risponda allo sconto acquistando più articoli anche quando il traffico totale non cresce in modo apprezzabile.

Sales Net. Il quadro sulle vendite nette è il più informativo dal punto di vista operativo e si divide nettamente in tre gruppi. Mid Season Sale (+0,08% ns) e Pre-Saldi (+0,00% ns) mostrano un effetto sulle Sales Net statisticamente non distinguibile da zero: intensificare lo sconto in questi regimi genera traffico senza erodere il fatturato netto, poiché l'incremento di volume compensa la riduzione del prezzo medio. Saldi Fase 3 mostra un segnale negativo (-0,21%) con IC che include marginalmente lo zero [-0,42; +0,01]. Saldi Fase 1 e Saldi Fase 2 mostrano invece effetti negativi forti e altamente significativi (rispettivamente -0,56%*** e -0,53%) con IC che escludono chiaramente lo zero. Il Black Friday presenta l'effetto più pronunciato di tutti (-1,03%) con IC [-1,61; -0,44], interamente nel territorio negativo.

ATV. L'effetto sull'ATV è negativo, significativo e sostanzialmente uniforme in tutti i regimi, con valori compresi tra -0,46% (Evento Calendario) e -1,22% (Black Friday). Questa uniformità riflette il meccanismo diretto dello sconto sul prezzo medio pagato per transazione ed è attesa per costruzione: abbassare il prezzo riduce il valore medio dello scontrino. La variazione residua tra regimi è interpretabile come effetto composizione. In particolare, nei contesti in cui l'elasticità dell'ATV supera in valore assoluto l'unità (Black Friday e Saldi Fase 2), un aumento dell'1% dello sconto è associato a una riduzione dell'ATV superiore all'1%. Questo risultato non è facilmente riconducibile al solo effetto meccanico di riduzione del prezzo e suggerisce la presenza di ulteriori meccanismi economici. Una prima interpretazione riguarda la composizione della clientela: lo sconto aggiuntivo potrebbe attrarre, al margine, consumatori più sensibili al prezzo e caratterizzati da livelli di spesa mediamente inferiori, amplificando la compressione del valore dello scontrino oltre quanto atteso dal solo effetto di prezzo. Una seconda interpretazione, non alternativa ma complementare alla precedente, riguarda la composizione del paniere acquistato. Sconti particolarmente elevati possono infatti rendere relativamente più convenienti articoli collocati nelle fasce di prezzo inferiori, inducendo parte dei consumatori a sostituire prodotti di fascia media con prodotti meno costosi ma fortemente promozionati. In questo caso la riduzione dell'ATV rifletterebbe non soltanto un cambiamento nel profilo dei clienti che entrano in negozio, ma anche uno spostamento della domanda verso

categorie di prodotto a minor valore unitario. I dati disponibili non consentono di distinguere in modo netto tra questi due canali, che possono plausibilmente operare simultaneamente.

5.4.5 Scala dell'effetto: un esercizio illustrativo

Disponendo delle elasticità marginali e dei livelli medi di ciascun regime, è possibile costruire un ordine di grandezza illustrativo dell'effetto atteso di un incremento di 5 punti percentuali dello sconto.

Tabella 5.7 - Effetto illustrativo di +5 pp di sconto per regime

Regime	AVG Discount (%)	AVG Ticket	El. Ticket (%)	Δ Ticket†	El. Sales (%)	Δ Sales† (€)
Pre-Saldi	21,8	45,6	+0,78***	+1,78	+0,00 ns	0
Mid Season Sale	13,1	34,1	+0,95***	+1,62	+0,08 ns	+9
Saldi Fase 2	31,5	43,9	+0,59***	+1,29	-0,53***	-58
Saldi Fase 3	28,9	37,2	+0,67***	+1,25	-0,21·	-20
Saldi Fase 1	30,2	58,1	+0,40**	+1,16	-0,56***	-78
Evento Calendario	12,0	47,2	+0,24·	+0,57	-0,22·	-28
Black Friday	22,6	62,1	+0,19 ns	+0,60	-1,03***	-207

*Nota: conversione illustrativa: $\Delta Y = \text{livello medio regime} \times (5 \times \text{elasticità}\%) / 100$. Non costituisce stima econometrica dell'effetto in livelli. *** $p < 0,001$; ** $p < 0,01$; * $p < 0,05$; · $p < 0,10$; ns = non significativo.*

È necessario essere espliciti sulla natura di questo esercizio: i valori che seguono non sono stime econometriche dell'effetto causale in livelli. Sono conversioni illustrative che moltiplicano l'elasticità 2WFE per il livello medio osservato nel regime, assumendo che quest'ultimo sia una base rappresentativa. Vanno letti come ordini di grandezza per comunicare la scala dell'effetto, non come previsioni puntuali.

5.4.6 Interpretazione

I risultati si organizzano in modo coerente attorno a tre profili distinti, che rispondono direttamente alla domanda operativa di partenza.

Regimi con potenziale residuo - intensificare ha senso. Il Mid Season Sale e il Pre-Saldi mostrano il profilo più favorevole: elasticità sul Ticket elevate (+0,95% e +0,78% rispettivamente), entrambe altamente significative, e Sales Net con effetto non distinguibile da zero. In termini illustrativi, un incremento di 5 pp di sconto nel Mid Season Sale è associato a circa +1,6 transazioni aggiuntive per store per giorno senza erosione del fatturato netto. Il Pre-Saldi mostra un profilo analogo con +1,8 transazioni illustrative. Questi sono i due regimi in cui l'evidenza empirica disponibile è più coerente con l'ipotesi che esista margine residuo per intensificare lo sconto senza compromettere il fatturato. Va però ribadito che gli IC dei due regimi si sovrappongono ampiamente: non è possibile affermare con certezza quale dei due presenti il potenziale maggiore, né che la differenza tra i due sia economicamente rilevante.

Regimi in zona di transizione - mantenere con attenzione. La Saldi Fase 3 si colloca in una posizione intermedia: elasticità sul Ticket positiva e significativa (+0,67%) ma Sales Net con un segnale negativo marginale (-0,21%) il cui IC tocca quasi lo zero. L'evidenza suggerisce che ci si trovi al confine tra una zona di potenziale residuo e una zona di saturazione: intensificare ulteriormente lo sconto nella Fase 3 potrebbe generare traffico aggiuntivo ma a scapito del fatturato netto. L'Evento Calendario presenta invece un profilo diverso: il segnale sul Ticket è positivo ma solo marginalmente significativo (+0,24%), con IC che include lo zero. Questo regime è caratterizzato da uno sconto strutturalmente molto basso (12,0% medio) e da un traffico già elevato (47,2 ticket medi) probabilmente guidato dalla ricorrenza commerciale più che dall'incentivo di prezzo. L'indicazione è di cautela: lo sconto potrebbe non essere la leva più efficace in questi contesti.

Regimi con segnali di saturazione - ridurre non sarebbe irrazionale. Saldi Fase 1 e Saldi Fase 2 mostrano un profilo che merita attenzione: l'elasticità sul Ticket è positiva ma più contenuta rispetto alle fasi successive (rispettivamente +0,40% e +0,59%), mentre l'effetto sulle Sales Net è negativo e altamente significativo in entrambi i casi (-0,56%*** e -0,53%***). In termini illustrativi, un incremento di 5 pp di sconto nella Fase 1 è associato a +1,2 transazioni aggiuntive ma a una perdita stimata di circa 78€ di fatturato netto per store per giorno. Questo non implica necessariamente che lo sconto attuale sia sbagliato, i saldi hanno obiettivi che

vanno oltre il fatturato netto, inclusa la rotazione delle scorte, ma suggerisce che si sia già in una zona di rendimenti decrescenti rispetto alla leva dello sconto. Un esperimento controllato di riduzione moderata dello sconto nelle fasi iniziali dei saldi consentirebbe di verificare se il traffico si riduce in modo significativo o se la domanda è sufficientemente anelastica da sostenere una maggiore redditività unitaria.

Caso da attenzione. Il Black Friday rappresenta il caso più estremo tra i regimi analizzati. L'elasticità sul Ticket è positiva ma non significativa (+0,19%, IC [-0,34; +0,72]), mentre l'effetto sulle Sales Net è fortemente negativo (-1,03%). Come discusso nella sezione 5.4.2, la mancata significatività del coefficiente sul Ticket è in parte attribuibile alla limitata variazione identificante disponibile ($\text{disc_sd} = 7,4$ punti percentuali concentrata in soli cinque giorni di osservazione) e non può quindi essere interpretata come evidenza conclusiva di assenza di risposta del traffico. L'effetto sulle Sales Net risulta invece statisticamente robusto e di entità rilevante: in termini illustrativi, un punto percentuale di sconto aggiuntivo durante il Black Friday è associato a una riduzione di circa 41€ di fatturato netto per store e per giorno.

Dal punto di vista manageriale, il risultato suggerisce che ulteriori incrementi dello sconto durante il Black Friday rischiano di generare una distruzione di valore più che un'espansione della domanda. La forte compressione dell'ATV indica infatti che il beneficio promozionale viene assorbito principalmente attraverso una riduzione del valore medio dello scontrino, senza evidenza di un corrispondente incremento del traffico. Questo pattern è coerente sia con l'ingresso di consumatori particolarmente sensibili al prezzo sia con uno spostamento degli acquisti verso articoli di fascia inferiore o maggiormente promozionati. In entrambi i casi, l'implicazione operativa è analoga: la leva dello sconto sembra incentivare un downgrade del mix di acquisto più che una reale espansione della spesa. Una possibile direzione di intervento consiste quindi nel ripensare la composizione dell'offerta promozionale, estendendo la convenienza anche a categorie o prodotti di fascia più elevata, così da stimolare la conversione verso articoli a maggior valore anziché concentrare la domanda sui segmenti più economici. In questa prospettiva, il problema non appare tanto la capacità del Black Friday di generare traffico, quanto la sua difficoltà nel preservare il valore dello scontrino.

La progressione dei saldi. Un pattern degno di nota emerge dall'analisi congiunta delle tre fasi dei saldi. L'elasticità sul Ticket cresce monotonamente dalla Fase 1 (+0,40%) alla Fase 2 (+0,59%) alla Fase 3 (+0,67%), mentre l'effetto sulle Sales Net migliora progressivamente

passando da fortemente negativo nelle Fasi 1 e 2 a marginalmente negativo nella Fase 3. Questo pattern è coerente con almeno due meccanismi non mutuamente esclusivi: nella fase finale dei saldi la clientela residua tende a essere più selettivamente price-sensitive, rispondendo maggiormente all'incentivo di sconto; oppure, con il progredire dei saldi, l'assortimento disponibile si riduce e gli sconti più profondi diventano necessari per stimolare acquisti su prodotti meno appetibili. Entrambi i meccanismi producono la stessa firma empirica: elasticità crescente e impatto sul fatturato meno negativo man mano che si avanza nelle fasi. È opportuno notare che gli IC delle Fasi 2 e 3 si sovrappongono sostanzialmente ([+0,30; +0,87] e [+0,46; +0,88] rispettivamente), rendendo non formalmente distinguibili le due stime: la progressione è un pattern indicativo, non una differenza statisticamente provata.

6. Analisi causale di una campagna promozionale

Le analisi del Capitolo 5 erano orientate a una domanda prevalentemente di ottimizzazione. L'obiettivo era comprendere come varia il rendimento marginale della leva prezzo nei diversi contesti operativi e promozionali. Le promozioni considerate in questo capitolo appartengono invece a una categoria diversa. Si tratta di iniziative discrezionali, attivabili o meno a seconda degli obiettivi commerciali e delle risorse disponibili. In questo caso la domanda economica e manageriale non riguarda più l'ottimizzazione dell'intensità dello sconto all'interno di una campagna già prevista, ma l'efficacia stessa della campagna. Prima di chiedersi quanto sconto applicare, è necessario stabilire se la promozione abbia generato un effetto incrementale reale rispetto allo scenario in cui non fosse stata attivata. Per questo motivo l'analisi richiede un diverso approccio metodologico, orientato all'identificazione causale dell'effetto della promozione piuttosto che alla stima dell'elasticità marginale dello sconto. In particolare le decisioni sull'attivazione, l'intensità e il timing delle promozioni vengono prese su base settimanale in condizioni di incertezza e con risorse limitate da allocare tra iniziative concorrenti, spesso senza poter distinguere se l'aumento delle vendite osservato durante una promozione sia realmente causato da essa o rifletta semplicemente la stagionalità o le caratteristiche degli store coinvolti. In un portafoglio promozionale complesso come quello descritto nella sezione 2 questa distinzione ha conseguenze dirette: significa poter scegliere le iniziative che generano traffico incrementale reale e scartare quelle che consumano margine senza aggiungere valore. I metodi di identificazione causale presentati in questa sezione rispondono precisamente a questa esigenza: fornire stime dell'effetto promozionale che siano interpretabili come effetti reali e non come associazioni spurie, e che possano quindi orientare concretamente le scelte future.

L'interesse di questo esercizio non risiede esclusivamente nella valutazione delle due campagne considerate, ma anche nella possibilità di mostrare come tecniche di identificazione causale possano essere integrate nei processi decisionali aziendali. In presenza di un gruppo di controllo plausibile, tali strumenti consentono infatti di valutare l'efficacia di iniziative pilota prima di un'estensione all'intera rete e di monitorarne successivamente le performance con maggiore affidabilità rispetto a semplici confronti descrittivi.

Tra le numerose iniziative promozionali descritte nel capitolo precedente, le due campagne considerate rappresentano un caso particolarmente interessante dal punto di vista analitico. A differenza delle promozioni nazionali estese simultaneamente all'intera rete, queste iniziative

interessano soltanto una parte degli store e generano quindi una variazione cross-sectional che rende possibile la costruzione di un controfattuale credibile. In altre parole, durante il periodo promozionale coesistono punti vendita esposti e non esposti all'iniziativa, creando le condizioni necessarie per distinguere l'effetto della promozione dalle normali variazioni stagionali della domanda.

6.1 Contesto e strategia di identificazione

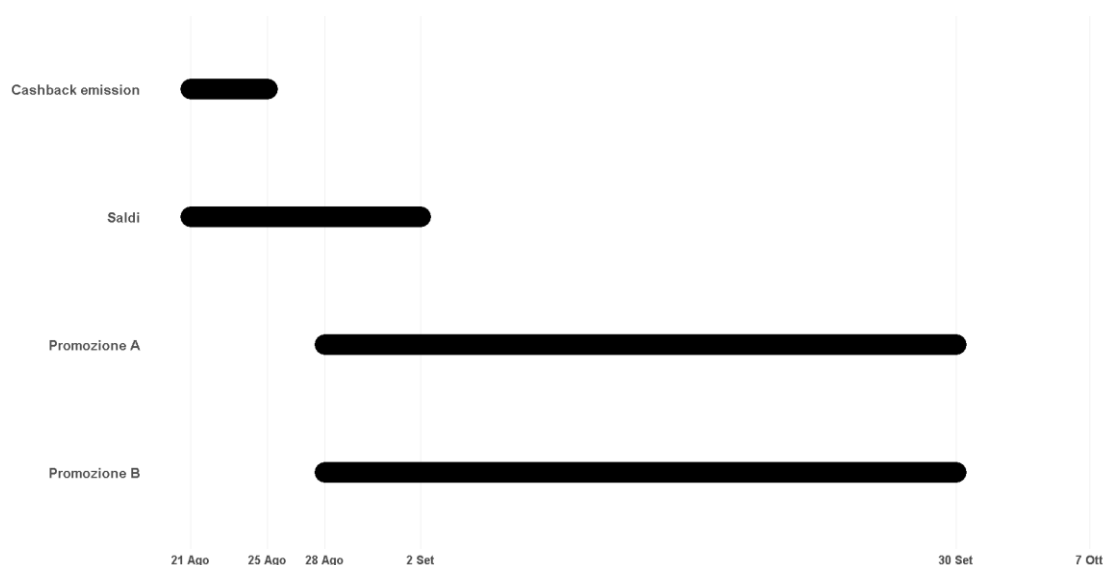
Le due promozioni analizzate nascono da una logica commerciale precisa: intervenire sul traffico in negozio in un momento dell'anno strutturalmente difficile, la transizione tra la fine dei saldi estivi e l'inizio della stagione autunnale, attraverso due leve distinte ma complementari.

La Promozione A è una campagna di co-marketing con un partner commerciale mass market: i clienti che effettuano un acquisto presso i punti vendita del partner ricevono un buono sconto del 20% da utilizzare presso i negozi dell'azienda, subordinato al raggiungimento di una soglia minima di spesa. L'obiettivo dichiarato è l'acquisizione di nuova clientela, attirando in negozio consumatori che non avrebbero altrimenti visitato lo store in quel periodo.

La Promozione B è un cashback del 20%, anch'esso condizionato a una soglia minima di spesa, distribuito ai clienti che avevano effettuato acquisti durante la fase di cashback emission, svoltasi tra il 17 luglio e il 25 agosto 2025. A differenza della Promozione A, questa iniziativa si rivolge a clientela già acquisita, incentivandone il ritorno attraverso un beneficio maturato in precedenza. Entrambe le promozioni sono attive dal 28 agosto al 30 settembre 2025.

La sovrapposizione tra l'avvio delle promozioni e la coda dei saldi estivi, come mostrato nella Figura 6.1, costituisce la principale sfida identificativa dell'analisi: durante quei sei giorni le variazioni osservate nelle performance degli store possono riflettere sia l'effetto delle promozioni sia la fase finale dei saldi estivi. Le strategie empiriche introdotte nelle sezioni successive sono progettate proprio per separare questi due effetti.

Figura 6.1 - Timeline delle iniziative promozionali nel periodo di analisi



Nota: il grafico riporta la durata e la sovrapposizione temporale delle quattro iniziative promozionali attive nel periodo di analisi: Promozione A, Promozione B, cashback emission e saldi estivi.

Dal punto di vista della struttura di copertura delle promozioni, come mostrato nella Tabella 6.1, si nota una marcata asimmetria tra i gruppi. Su 165 store osservati, 37 risultano esposti contemporaneamente a entrambe le iniziative, 94 sono esposti alla sola Promozione B, 2 alla sola Promozione A e 32 non ricevono alcun trattamento. Questa distribuzione ha una conseguenza metodologica diretta: non tutti gli obiettivi di stima godono dello stesso grado di identificabilità empirica.

La Promozione B può essere studiata in modo relativamente autonomo e credibile. I 94 store esposti esclusivamente a questa iniziativa, confrontati con i 32 store non trattati, offrono una base di confronto numericamente adeguata e non contaminata dalla presenza simultanea della Promozione A.

L'identificazione separata dell'effetto della Promozione A è invece empiricamente fragile. Soltanto 2 store risultano esposti esclusivamente a questa iniziativa, una base insufficiente per trarre conclusioni robuste. Per questo motivo, l'effetto della Promozione A viene analizzato esclusivamente nella sua dimensione combinata con la Promozione B, quindi analizzeremo la variazione attesa del KPI nei negozi esposti simultaneamente a entrambe le promozioni.

Tabella 6.1 - Frequenza assoluta di assegnazione al trattamento

Trattamento	Store trattati
Promo A & Promo B	37
Promo A	2
Promo B	94
Nessun trattamento	32

Nota: N = 165 store L4L. Ciascuna cella riporta il numero di store esposti alla combinazione di trattamento corrispondente.

L'analisi si articola pertanto attorno a due oggetti di stima distinti: l'effetto della sola Promozione B, identificato nel confronto tra i 94 store trattati esclusivamente da questa iniziativa e i 32 store di controllo, e l'effetto combinato delle due promozioni, identificato nei 37 store con doppia esposizione. Questa distinzione, che emerge direttamente dalla struttura dei dati, guida l'intera interpretazione dei risultati.

L'identificazione è tuttavia complicata da due fattori contestuali comuni a entrambi gli oggetti di stima. Il primo è la sovrapposizione con il periodo dei saldi estivi, uno shock comune a tutti gli store che agisce direttamente sui KPI di vendita, rendendo difficile separarne l'effetto da quello promozionale. Il secondo è la presenza della precedente iniziativa di cashback emission, conclusasi poco prima dell'avvio delle promozioni di interesse, che potrebbe aver modificato il comportamento d'acquisto dei clienti nel periodo immediatamente precedente all'avvio delle promozioni, rendendo difficile separare l'effetto baseline da quello promozionale per gli store esposti alla Promozione B.

Di fronte a queste criticità, l'analisi adotta una strategia di identificazione progressiva articolata in tre stadi di crescente rigore. Nel primo si stima un modello OLS che cattura la relazione grezza tra promozioni e KPI senza alcun controllo: questa stima serve a stabilire un punto di riferimento e a rendere visibile l'entità del potenziale bias. Nel secondo si introduce un modello a effetti fissi a due vie, che controlla per l'eterogeneità strutturale tra negozi e per gli shock temporali comuni, producendo stime più pulite basate sulla variazione within-store nel tempo. Nel terzo si adotta un approccio di Regression Discontinuity Design (RD), che sfrutta la discontinuità temporale nell'avvio delle promozioni per identificare un effetto causale locale, confrontando le osservazioni immediatamente adiacenti al cutoff, le più comparabili tra loro per costruzione.

6.2 Analisi preliminare: modello OLS

Come primo stadio della strategia di identificazione, si stima un modello lineare in cui i KPI di interesse sono messi in relazione con le due dummy promozionali e la loro interazione. La specificazione adottata è la seguente:

$$Y_{it} = \delta_A \text{PromoA}_{it} + \delta_B \text{PromoB}_{it} + \delta_{AB} (\text{PromoA} \times \text{PromoB})_{it} + \varepsilon_{it}$$

dove Y_{it} rappresenta il KPI di interesse per lo store i nel giorno t ; PromoA_{it} e PromoB_{it} sono variabili dummy che assumono valore 1 in presenza rispettivamente della Promozione A e della Promozione B; $(\text{PromoA} \times \text{PromoB})_{it}$ è il termine di interazione che cattura l'effetto addizionale derivante dall'attivazione simultanea delle due iniziative. Il coefficiente $\hat{\delta}_B$ misura l'effetto della Promozione B negli store esposti esclusivamente a questa iniziativa, ovvero il confronto di interesse primario discusso nella sezione precedente. Il coefficiente $\hat{\delta}_A$ misura l'effetto della Promozione A nei soli 2 store con esposizione esclusiva, una stima da interpretare con estrema cautela. La somma $\hat{\delta}_A + \hat{\delta}_B + \hat{\delta}_{AB}$ misura invece l'effetto combinato nei 37 store con doppia esposizione.

6.2.1 Risultati

Tabella 6.2 - Effetto delle promozioni sui KPI: stime OLS

	log(Ticket)	log(UPT)	log(Sales Net)	log(ATV)
Promo A	+151,7%*** (21,9)	+14,0%*** (2,2)	+150,9%*** (23,6)	-0,2% (2,6)
Promo B	-24,1%*** (1,3)	-3,0%*** (0,4)	-27,2%*** (1,3)	-4,1%*** (0,5)
Interazione (A:B)	-37,5%*** (5,6)	-9,8%*** (1,8)	-36,7%*** (6,2)	+1,5% (2,7)

*Nota: stime OLS, nell'intervallo 18 agosto–12 ottobre 2025. Errori standard non clusterizzati in parentesi, trasformati via delta method. Coefficienti espressi come variazione percentuale tramite $(\exp(\beta)-1) \times 100$; l'intercetta è mantenuta in scala logaritmica. L'effetto combinato nei 37 store con doppia esposizione si ottiene sommando i coefficienti in scala log prima della conversione: Ticket +19,4%, Sales Net +15,7%, UPT -0,3%, ATV -2,9%. *** $p < 0,001$.*

6.2.2 Interpretazione

I coefficienti OLS raccontano una storia apparentemente controintuitiva. Il segno negativo e fortemente significativo di $\hat{\delta}_B$ su tutti i KPI, $-24,1\%$ sul Ticket, $-3,0\%$ sull'UPT, $-27,2\%$ sulle Sales Net e $-4,1\%$ sull'ATV, non deve essere letto come evidenza che la Promozione B deprime le vendite, ma come evidenza che gli store esposti esclusivamente a questa iniziativa sono strutturalmente diversi dagli store di controllo. In assenza di effetti fissi, il modello confronta due gruppi non comparabili: i 94 store trattati dalla sola Promo B con i 32 store di controllo, senza tenere conto del fatto che questi due gruppi differiscono per dimensione, localizzazione, formato e livello di traffico di base. Il coefficiente negativo riflette queste differenze preesistenti più che l'effetto della promozione. In altri termini, il segno negativo è coerente con l'ipotesi che gli store esposti alla Promozione B avessero, mediamente, performance di base inferiori rispetto ai controlli.

L'effetto combinato $\hat{\delta}_A + \hat{\delta}_B + \hat{\delta}_{AB}$ è invece positivo per i KPI di volume: $+19,4\%$ sul Ticket e $+15,7\%$ sulle Sales Net, entrambi significativi. Per l'UPT l'effetto è nullo ($-0,3\%$, non significativo), mentre per l'ATV risulta lievemente negativo ($-2,9\%$, $p < 0,01$). Anche queste stime restano però associazioni aggregate: la coincidenza temporale con i saldi estivi e la selezione non casuale degli store trattati impediscono un'interpretazione causale.

6.2.3 Limiti interpretativi

Le promozioni non sono assegnate casualmente, ma rispondono a logiche operative e commerciali: è plausibile che vengano attivate in periodi caratterizzati da una domanda attesa più debole, generando una relazione tra trattamento e outcome influenzata da timing endogeno. L'intervallo temporale analizzato si sovrappone parzialmente con il periodo dei saldi estivi, che rappresenta uno shock comune a tutti gli store e che influisce direttamente sui KPI di vendita, introducendo una rilevante fonte di confondimento. Nel complesso, le stime OLS vanno lette esclusivamente come associazioni aggregate, utili come punto di riferimento per valutare l'entità della distorsione nei modelli successivi.

6.3 Modello con effetti fissi

Per superare i limiti interpretativi del modello OLS, si introduce un modello con effetti fissi di store e di data. La specificazione adottata è la seguente:

$$Y_{it} = \delta_A \text{Promo}A_{it} + \delta_B \text{Promo}B_{it} + \delta_{AB}(\text{Promo}A \times \text{Promo}B)_{it} + \delta_C \text{Promo}C_{it} + \alpha_i + \gamma_t + \varepsilon_{it}$$

L'effetto fisso di negozio α_i assorbe tutte le caratteristiche strutturali e time-invariant del punto vendita (dimensione, localizzazione, format commerciale, composizione della clientela abituale), eliminando il bias dovuto a differenze sistematiche tra store trattati e non trattati. L'effetto fisso di data γ_t cattura ogni shock comune a tutti i negozi in un determinato giorno, inclusi la stagionalità, i giorni della settimana e l'andamento aggregato dei saldi estivi. Il modello include inoltre un indicatore per la promozione C come controllo esplicito, al fine di isolare il suo potenziale effetto confondente sul periodo precedente all'avvio delle promozioni di interesse. L'identificazione avviene quindi confrontando ciascun negozio con sé stesso nei periodi di attivazione promozionale rispetto ai periodi ordinari, al netto degli andamenti comuni a tutti i negozi in ciascun giorno: un confronto intra-store nel tempo che è metodologicamente ben più credibile del confronto cross-sezionale implicito nell'OLS.

6.3.1 Risultati

Tabella 6.3 - Effetto delle promozioni sui KPI: stime 2WFE

	log(Ticket)	log(UPT)	Log(Sales Net)	log(ATV)
Promo A	+12,9% (7,4)	+0,8% (2,0)	+14,2% (9,2)	+1,3% (2,0)
Promo B	+5,3%* (2,3)	+0,5% (0,9)	+3,4% (3,1)	-2,0% (1,4)
Interazione (A * B)	-10,4% (7,7)	-1,7% (2,1)	-10,9% (9,5)	-0,6% (2,2)

Nota: stime 2WFE nell'intervallo 18 agosto–12 ottobre 2025. Effetti fissi store (165) e giornalieri (56 giorni). Errori standard clusterizzati a livello store in parentesi, moltiplicati per 100. Coefficienti espressi come variazione percentuale approssimata ($\beta \times 100$), valida per valori contenuti. L'effetto combinato nei 37 store con doppia esposizione: Ticket +7,8%, Sales Net +6,7%, UPT -0,4%, ATV -1,3%. $\cdot p < 0,10$; $ p < 0,05$.*

6.3.2 Interpretazione

Il risultato più rilevante del modello 2WFE riguarda la Promozione B. Una volta controllato per l'eterogeneità strutturale tra store e per gli shock temporali comuni, il coefficiente $\hat{\delta}_B$ si inverte di segno rispetto all'OLS: da $-24,1\%$ diventa $+5,3\%$ sul Ticket ($p < 0,05$), con le Sales Net che passano da $-27,2\%$ a $+3,4\%$, pur rimanendo statisticamente non significative. Questo capovolgimento non è anomalo: replica esattamente il meccanismo già osservato in sezione 5.2 per lo sconto medio. L'OLS catturava una correlazione spuria generata dalla selezione non casuale degli store nel gruppo trattato: i negozi esposti alla Promozione B erano strutturalmente diversi dai controlli, e quella differenza preesistente dominava il segnale promozionale. Il 2WFE, confrontando ciascun store con sé stesso nel tempo, elimina questa fonte di confondimento e restituisce la prima stima credibile di un effetto positivo della promozione sul traffico.

L'effetto su UPT e ATV rimane non significativo, suggerendo che la promozione abbia operato esclusivamente sul margine estensivo della domanda: ha aumentato la frequenza di visita senza modificare la struttura della transazione, né la profondità del carrello né il valore medio dello scontrino.

Rileggere questi risultati attraverso l'identità log-additiva $\log(\text{Sales Net}) = \log(\text{Ticket}) + \log(\text{ATV})$ aiuta a chiarire i meccanismi sottostanti. Per la Promozione B, il Ticket cresce di $+5,3\%$ ma l'ATV si comprime di $-2,0\%$, non significativo ma sistematicamente negativo: i due effetti si compensano quasi interamente, producendo un impatto sulle Sales Net positivo ma non distinguibile da zero. Questo profilo suggerisce che i clienti che riscattano il cashback tornino in negozio per acquisti di valore contenuto, probabilmente appena sopra la soglia minima necessaria per utilizzare il beneficio, senza incrementare la spesa media rispetto a una visita ordinaria. La promozione genera traffico reale ma non modifica il comportamento d'acquisto una volta in negozio.

Per l'effetto combinato il quadro cambia: il Ticket cresce di $+7,8\%$ e l'ATV è sostanzialmente invariato ($-1,3\%$, non significativo), producendo un effetto sulle Sales Net positivo ($+6,7\%$ significativo). La differenza rispetto alla sola Promo B non è solo quantitativa ma qualitativa: la Promozione A porta in negozio clienti che non hanno un beneficio da riscattare e si comportano come clienti ordinari, senza la tendenza a minimizzare la spesa che caratterizza il riscatto del cashback. Questo suggerisce che le due iniziative raggiungano segmenti di clientela con propensioni di spesa strutturalmente diverse, e che la loro complementarità non sia

semplicemente additiva sul traffico ma generi una composizione della domanda più favorevole di quanto ciascuna riesca a fare separatamente.

6.3.3 Confronto OLS vs 2WFE

Il confronto tra le stime OLS e quelle 2WFE offre indicazioni metodologiche ed economiche distinte per i due oggetti di stima.

Per la Promozione B, il cambiamento è radicale sia in segno che in significatività: il coefficiente passa da -24,1% a +5,3% sul Ticket e da -27,2% a +3,4% sulle Sales Net. Questo risultato illustra con particolare chiarezza il ruolo degli effetti fissi nella correzione del bias da selezione: gli store esposti esclusivamente alla Promozione B non erano un campione casuale della popolazione di store, ma un gruppo con caratteristiche strutturali sistematicamente diverse dagli store di controllo. Il confronto cross-sezionale dell'OLS confondeva queste differenze preesistenti con l'effetto della promozione, producendo una stima distorta e fuorviante. Il 2WFE, confrontando ciascun store con sé stesso nel tempo, elimina questa fonte di confondimento e restituisce una stima positiva e significativa, che rappresenta la prima evidenza credibile di un effetto reale della Promozione B sui KPI di volume.

Per l'effetto combinato, la dinamica è diversa: la contrazione da +19,4% a +7,8% sul Ticket non cambia il segno della stima ma ne riduce considerevolmente la magnitudine, indicando che parte delle associazioni positive osservate nell'OLS era riconducibile alla selezione non casuale degli store trattati e alla coincidenza temporale con i saldi estivi.

Il quadro complessivo che emerge dal confronto tra OLS e 2WFE conferma un pattern già osservato nella sezione 5: i modelli senza controlli distorcono sistematicamente le stime, e solo il confronto within-store restituisce segnali interpretabili. Per la Promozione B il cambiamento è radicale: da -24,1% a +5,3% sul Ticket, da -27,2% a +3,4% sulle Sales Net. La contrazione da +19,4% a +7,8% sull'effetto combinato non cambia il segno ma ne riduce la magnitudine, indicando che parte dell'associazione positiva osservata nell'OLS era gonfiata dalla coincidenza temporale con i saldi estivi e dalla selezione non casuale degli store. In termini operativi: il cashback da solo genera traffico reale ma non incremento di fatturato, mentre la combinazione con la Promozione A produce un effetto economicamente rilevante e statisticamente robusto su entrambi i KPI, pari a +7,8% sul Ticket e +6,7% sulle Sales Net.

6.3.4 Limiti dell'approccio con effetti fissi

Il modello 2WFE rappresenta un progresso metodologico rilevante rispetto all'OLS, ma non risolve integralmente il problema dell'identificazione causale. Gli effetti fissi di store e di data assorbono rispettivamente l'eterogeneità strutturale tra negozi e gli shock comuni a tutti i punti vendita in ciascun giorno, ma non sono sufficienti a eliminare tutte le possibili fonti di distorsione.

Il limite più rilevante riguarda la possibile presenza di shock non osservati variabili nel tempo a livello di singolo store. Se la decisione di attivare la promozione in un determinato negozio fosse correlata con dinamiche locali della domanda non catturate dall'effetto fisso temporale comune (ad esempio un calo del traffico specifico di quell'area geografica, o un'iniziativa competitiva locale), il coefficiente stimato resterebbe distorto anche dopo l'introduzione dei fixed effects. L'effetto fisso di data controlla per ciò che accade in media a tutti i negozi nello stesso giorno, ma non per ciò che accade in modo eterogeneo tra store nello stesso giorno.

Un ulteriore limite riguarda la dimensione temporale dell'effetto. Il 2WFE stima un effetto medio sull'intero periodo in cui le promozioni sono attive, attribuendo lo stesso peso ai giorni immediatamente successivi al lancio e a quelli più distanti. Se l'effetto promozionale è concentrato nel momento di avvio e si esaurisce rapidamente (come la struttura del cashback potrebbe suggerire, dato il ritardo tra acquisto e maturazione del beneficio), questa media temporale diluisce il segnale fino a renderlo non rilevabile. Questo limite vale in misura simile per entrambi gli oggetti di stima: sia l'effetto della Promozione B che quello combinato potrebbero essere più pronunciati nella fase di lancio e risultare attenuati nella media di periodo stimata dal 2WFE.

Per affrontare queste limitazioni si introduce nella sezione successiva un approccio di Regression Discontinuity Design. L'RD sfrutta la discontinuità temporale nell'avvio delle promozioni confrontando esclusivamente le osservazioni immediatamente adiacenti al cutoff, dove l'assunzione di comparabilità è più credibile e dove un eventuale effetto di lancio sarebbe massimamente visibile. Questo approccio non richiede che il timing promozionale sia casuale sull'intera finestra temporale: richiede soltanto che, in un intorno sufficientemente stretto del giorno di avvio, non vi siano altre discontinuità sistematiche in grado di spiegare il salto osservato nei KPI.

6.4 Regression Discontinuity Design

Il modello a effetti fissi a due vie ha prodotto stime robuste dell'elasticità media dello sconto sull'intero periodo promozionale, controllando per l'eterogeneità strutturale tra store e per gli shock temporali comuni. Rimane però una domanda a cui il 2WFE non è progettato per rispondere: qual è l'effetto di una campagna promozionale nel momento specifico del suo lancio? Per costruzione, il 2WFE stima un effetto medio di periodo, attribuendo lo stesso peso ai giorni immediatamente successivi all'avvio e a quelli più distanti. Se l'effetto promozionale è concentrato nelle prime giornate e si attenua rapidamente, questa media temporale ne diluisce il segnale. Per isolare la risposta immediata al lancio si introduce un approccio quasi-sperimentale basato sul Regression Discontinuity Design (RD).

L'idea fondamentale è che, in un intorno sufficientemente stretto del giorno di lancio, le osservazioni immediatamente precedenti e quelle immediatamente successive al cutoff siano sostanzialmente comparabili, differendo sistematicamente solo per l'esposizione alla promozione. In questa finestra ristretta, la variazione nel trattamento può essere trattata come plausibilmente esogena, consentendo un'interpretazione causale del coefficiente stimato. L'identificazione si basa sull'assunzione di continuità locale: in assenza della promozione, l'evoluzione attesa dei KPI avrebbe proseguito in modo regolare attraverso il cutoff, senza generare alcun salto discontinuo nel momento di avvio della campagna. Eventuali discontinuità osservate possono pertanto essere attribuite all'introduzione del trattamento.

È importante sottolineare come l'RD e il 2WFE rispondano a domande causali parzialmente diverse e si riferiscano a popolazioni di store distinte. Il 2WFE stimava l'effetto medio sull'intero periodo promozionale, articolato su due oggetti di stima: l'effetto della sola Promozione B sui 94 store con esposizione esclusiva, e l'effetto combinato delle due promozioni sui 37 store con doppia esposizione. L'analisi RD è limitata ai 37 punti vendita esposti simultaneamente alla Promozione A e alla Promozione B, poiché solo per questi store il cutoff temporale identifica un cambiamento omogeneo del regime promozionale. Il parametro stimato dall'RD non è dunque l'effetto medio di periodo, bensì l'effetto di lancio del sistema promozionale combinato in quella popolazione specifica, identificato nel confronto tra i giorni immediatamente precedenti e immediatamente successivi al cutoff. La lettura congiunta dei due modelli consente di distinguere due dimensioni dell'effetto promozionale: la risposta

immediata al lancio, catturata dall'RD, e la sua persistenza nel medio periodo, catturata dal 2WFE.

6.4.1 Specificazione del modello

La variabile di running è definita come la distanza in giorni dal giorno di avvio della campagna promozionale:

$$running_{it} = t - t_0$$

dove t_0 rappresenta il giorno di avvio. La dummy di trattamento è:

$$D_{it} = \mathbb{1}(running_{it} \geq 0)$$

Il modello RD locale lineare con effetti fissi di store è:

$$\log(Y_{it}) = \tau D_{it} + \beta_1 \cdot running_{it} + \beta_2 \cdot D_{it} \times running_{it} + \alpha_i + \varepsilon_{it}$$

dove α_i è l'effetto fisso di store e ε_{it} è il termine di errore con varianza stimata clusterizzando a livello store. Il parametro di interesse è τ , che cattura il salto discontinuo nel logaritmo del KPI al momento dell'avvio della promozione, al netto della traiettoria tendenziale nel periodo pre-cutoff e di quella nel periodo post-cutoff. Il modello stima due rette locali distinte ai due lati del cutoff, con pendenze β_1 (pre-cutoff) e $\beta_1 + \beta_2$ (post-cutoff). Il coefficiente τ si interpreta come variazione percentuale approssimata del KPI al cutoff, calcolata come $(\exp(\hat{\tau}) - 1) \times 100$.

L'inclusione degli effetti fissi di store rafforza lo schema RD rispetto al disegno cross-sezionale standard: condizionando per α_i , l'identificazione avviene interamente within-store, eliminando il rischio che differenze strutturali tra store in termini di dimensione, localizzazione o composizione della clientela contaminino la stima dell'effetto di lancio.

Scelta della bandwidth. L'analisi adotta bandwidth $h = 2$, $h = 4$ e $h = 10$ giorni attorno al cutoff, limitatamente ai 37 store con doppia esposizione promozionale. La scelta non è basata su procedure automatiche di ottimizzazione statistica, ma su considerazioni contestuali legate alla struttura commerciale del periodo analizzato. Il terzo giorno antecedente all'avvio delle promozioni coincide con la conclusione dell'iniziativa di cashback emission precedente, mentre il quinto e il sesto giorno successivi al cutoff corrispondono all'ultimo weekend dei saldi estivi,

entrambi eventi che possono generare variazioni nei KPI indipendenti dall'effetto causale della promozione. Le tre bandwidth consentono di valutare la robustezza dei risultati rispetto all'inclusione di osservazioni temporalmente più lontane dal cutoff, bilanciando il trade-off tra varianza (ridotta ampliando la finestra) e distorsione potenziale (aumentata includendo giorni con fattori confondenti).

Il modello è stimato sui log-KPI grezzi, questa scelta è coerente con il principio generale di manipolare il meno possibile la variabile dipendente nell'RD, lasciando che il modello locale lineare approssimi direttamente la funzione di risposta attorno al cutoff. Una conseguenza di questa scelta è che la variazione infrasettimanale della domanda, sistematicamente diversa tra giorni feriali e weekend, costituisce una fonte di rumore all'interno di ciascuna finestra. Finestre più strette, che includono pochi giorni e quindi potenzialmente pochi cicli settimanali completi, sono più esposte a questo rumore rispetto a finestre più ampie. Questo spiega, almeno in parte, il pattern crescente delle stime al crescere di h , discusso nella sezione dei risultati.

Pesi kernel. I pesi kernel triangolari $w_{it} = 1 - |running_{it}|/h$ attribuiscono peso massimo alle osservazioni nel giorno del cutoff e peso decrescente linearmente alle osservazioni più lontane, privilegiando il confronto locale.

Giorno di lancio. Un elemento contestuale merita attenzione prima di presentare i risultati. La promozione è stata avviata il 28 agosto 2025, un giovedì. L'assunzione di continuità locale, fondamento identificativo dell'RD, implica che le condizioni medie del mercoledì precedente fossero sostanzialmente comparabili a quelle del giovedì di lancio in assenza della promozione. Questa assunzione è plausibile per una finestra sufficientemente stretta e in assenza di altri eventi discontinui nel giorno di avvio, condizione che la struttura del calendario commerciale discussa nella scelta della bandwidth rende ragionevolmente credibile per $h=2$ e $h=4$.

Il giorno di avvio introduce tuttavia un secondo elemento di cautela che riguarda specificamente l'analisi della pendenza post-cutoff. Poiché la promozione è partita di giovedì, la composizione feriale/weekend dei due lati della finestra è sbilanciata per tutte le bandwidth superiori a $h=2$. Per $h=4$ il pre-cutoff include un solo giorno di weekend (domenica 24 agosto) contro due giorni nel post-cutoff (sabato 30 e domenica 31 agosto); per $h=10$ il rapporto è di due giorni pre contro tre post. Poiché il traffico e il fatturato sono strutturalmente più alti nel weekend, questo sbilanciamento, combinato col fatto che nel post-cutoff i weekend sono più vicini al giorno di salto (quindi avendo un peso kernel maggiore), tende a produrre una

pendenza post-cutoff artificialmente più ripida rispetto a quella pre-cutoff, indipendentemente dall'effetto causale della promozione. L'unica finestra in cui la composizione weekend è perfettamente bilanciata è $h=2$, che include zero giorni di weekend su entrambi i lati.

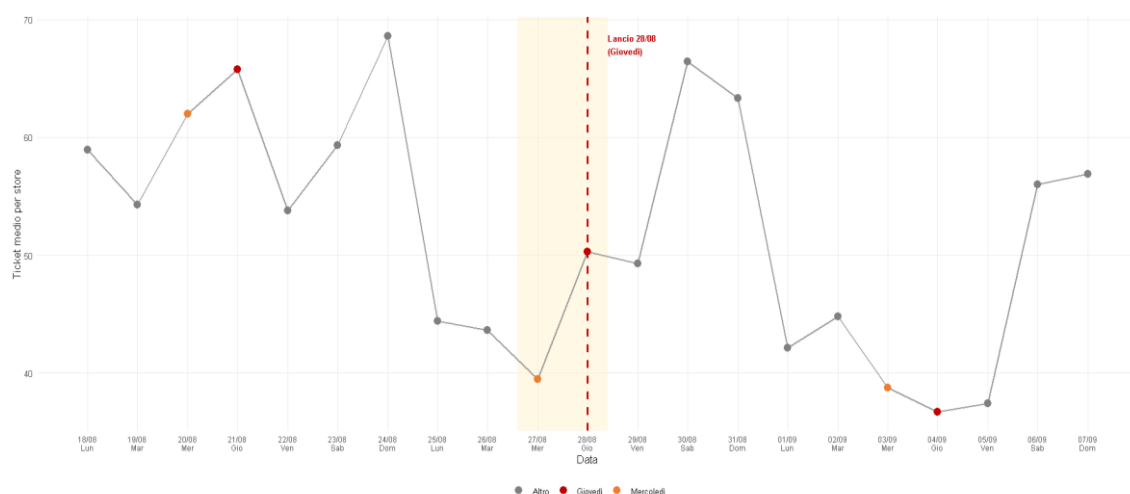
Per queste ragioni, l'analisi principale si concentra sul parametro τ , il salto discontinuo al cutoff, come oggetto di stima primario, trattando il coefficiente β_2 sull'interazione $D \times \text{running}$ con la cautela dovuta al potenziale confondimento infrasettimanale. Una verifica più credibile della variazione di pendenza post-lancio viene condotta nella sezione 6.4.3 su dati destagionalizzati, dove la rimozione della componente infrasettimanale rende il confronto tra le due pendenze più credibile e consente di valutare in modo più affidabile se la promozione abbia modificato la traiettoria dei KPI oltre al livello.

Continuità locale. Un'ulteriore verifica a supporto dell'assunzione di continuità locale consiste nel confrontare l'andamento del numero medio di Ticket attorno al cutoff per il gruppo trattato e per il gruppo di controllo. Come mostrato dalla Figura 6.3 e, soprattutto, dal confronto con la Figura 6.2, il gruppo trattato presenta una discontinuità marcata in corrispondenza dell'avvio della promozione: il valore medio del KPI passa da 38,6 nel periodo immediatamente precedente il cutoff a 46,4 nel periodo successivo. Nel gruppo di controllo, invece, la variazione osservata nello stesso intervallo temporale è molto più contenuta, passando da 44,8 a 47,6.

La presenza di un lieve incremento anche nel gruppo di controllo suggerisce che possano esistere dinamiche temporali comuni o una moderata componente stagionale che interessa l'intero sistema. Tuttavia, l'entità di tale variazione appare nettamente inferiore rispetto a quella osservata nei punti vendita trattati. Questa evidenza indica che la discontinuità rilevata nel gruppo trattato difficilmente può essere attribuita esclusivamente a normali fluttuazioni temporali o stagionali condivise da tutti gli store.

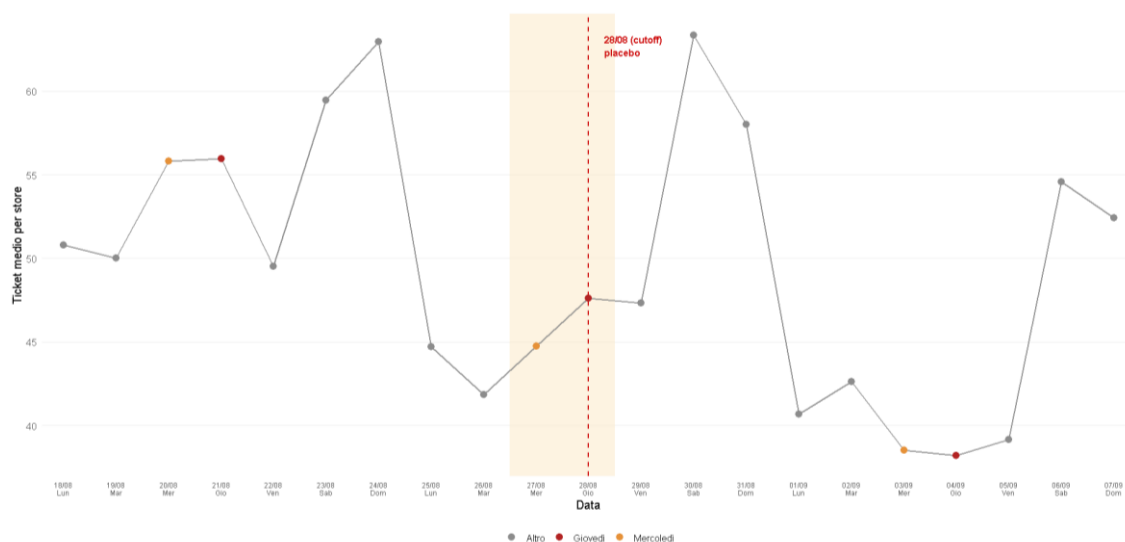
Pur non consentendo una verifica diretta del controfattuale dei punti vendita trattati, il confronto tra i due gruppi fornisce quindi evidenza coerente con l'assunzione di continuità locale richiesta dal disegno RD e non evidenzia elementi tali da mettere in discussione l'interpretazione causale della discontinuità stimata nelle analisi successive.

Figura 6.2 - AVG Ticket nella finestra ± 10 giorni dal cutoff: gruppo di trattati



Nota: medie giornaliere calcolate sui 37 store esposti simultaneamente alla Promozione A e alla Promozione B. La linea tratteggiata rossa indica il giorno di avvio della campagna (28 agosto 2025, giovedì). La banda gialla evidenzia la coppia mercoledì 27 agosto, giovedì 28 agosto. I punti arancioni indicano i mercoledì e quelli rossi i giovedì presenti nella finestra, per rendere visibile la comparabilità strutturale tra i due giorni della settimana.

Figura 6.3 - AVG Ticket nella finestra ± 10 giorni dal cutoff: gruppo di controllo



Nota: medie giornaliere calcolate sui 32 store che non aderivano né alla promozione A né alla promozione B. La linea tratteggiata rossa indica il giorno di avvio della campagna (28 agosto 2025, giovedì). La banda gialla evidenzia la coppia mercoledì 27 agosto, giovedì 28 agosto. I punti arancioni indicano i mercoledì e quelli rossi i giovedì presenti nella finestra, per rendere visibile la comparabilità strutturale tra i due giorni della settimana.

6.4.2 Risultati

Tabella 6.4 - Stime RD: dati grezzi

		log(Ticket)		log(Sales Net)		log(UPT)		log(ATV)	
Termine	h	Stima	p	Stima	p	Stima	p	Stima	p
τ (Effetto %)	2	+20,4% (7,3)	0,002**	+19,3% (7,9)	0,008**	-4,4% (3,6)	ns	-0,9% (4,1)	ns
τ (Effetto %)	4	+31,4% (7,9)	<0,001***	+28,5% (8,3)	<0,001***	-5,4% (3,1)	.	-2,2% (3,5)	ns
τ (Effetto %)	10	+40,7% (5,9)	<0,001***	+41,1% (6,5)	<0,001***	-2,8% (1,7)	.	+0,2% (1,9)	ns
β_1 (running)	2	+0,026 (0,039)	ns	+0,031 (0,041)	ns	+0,002 (0,021)	ns	+0,004 (0,022)	ns
β_1 (running)	4	-0,076 (0,029)	**	-0,061 (0,028)	*	+0,014 (0,012)	ns	+0,016 (0,013)	ns
β_1 (running)	10	-0,069 (0,007)	***	-0,063 (0,007)	***	-0,000 (0,002)	ns	+0,006 (0,002)	*
β_2 (D×running)	2	—	—	—	—	—	—	—	—
β_2 (D×running)	4	+0,175 (0,055)	**	+0,177 (0,058)	**	-0,003 (0,015)	ns	+0,003 (0,015)	ns
β_2 (D×running)	10	+0,031 (0,008)	***	+0,035 (0,009)	***	+0,006 (0,003)	.	+0,004 (0,004)	ns

*Nota: RD locale lineare con pesi kernel triangolari, effetti fissi store (37 store), errori standard clusterizzati a livello store in parentesi. τ riportato come variazione percentuale $(\exp(\hat{\tau})-1) \times 100$ con SE via delta method. β_1 e β_2 in scala logaritmica: misurano la variazione del $\log(KPI)$ per giorno e non si prestano a conversione percentuale diretta. — indica collinearità: con $h=2$ la finestra di 4 giorni non consente di identificare β_2 separatamente. N: $h=2 \rightarrow 111$ oss., $h=4 \rightarrow 258$ oss., $h=10 \rightarrow 700$ oss. *** $p < 0,001$; ** $p < 0,01$; * $p < 0,05$; . $p < 0,10$; ns = non significativo.*

Ticket Count. Il coefficiente $\hat{\tau}$ è positivo e significativo in tutte e tre le specifiche: +20,4%, con $h=2$ ($p = 0,002$), +31,4% con $h=4$ ($p < 0,001$) e +40,7% con $h=10$ ($p < 0,001$). Il pattern crescente con la bandwidth è coerente con il problema di composizione infrasettimanale discusso nella sezione 6.4.1: le finestre più ampie includono più weekend nel post-cutoff rispetto al pre-cutoff, gonfiando artificialmente la stima. La finestra $h=2$ è l'unica con composizione weekend bilanciata e produce la stima più conservativa (+20,4%); $h=4$ offre il miglior compromesso tra prossimità al cutoff e riduzione del rumore. La significatività è robusta in tutte e tre le specifiche, il che costituisce evidenza solida di un effetto positivo al cutoff indipendentemente dalla bandwidth scelta.

log(Sales Net). Il pattern è quasi speculare a quello del Ticket: +19,3% con $h=2$ ($p = 0,008$), +28,5% con $h=4$ ($p < 0,001$) e +41,0% con $h=10$ ($p < 0,001$). La vicinanza tra l'effetto sul Ticket

e quello sulle Sales Net nelle finestre più strette (+20,4% vs +19,3% per $h=2$, +31,4% vs +28,5% per $h=4$) è coerente con l'assenza di effetto sull'ATV: l'incremento del fatturato è quasi interamente spiegato dall'incremento del numero di transazioni.

log(UPT) e log(ATV). Nessuna specifica produce un effetto significativo al livello convenzionale del 5% per questi due KPI. Per l'UPT le stime sono negative e di piccola magnitudine in tutte le bandwidth, con significatività marginale per $h=4$ e $h=10$ ($p \approx 0,09-0,10$): un segnale debolmente suggestivo ma non conclusivo, che scompare nella finestra più stretta dove l'identificazione è più locale. Per l'ATV le stime sono numericamente trascurabili con p-value molto elevati in tutte le specifiche. L'assenza di effetto su UPT e ATV è il risultato più stabile dell'intera analisi RD sui dati grezzi.

Il termine β_2 sull'interazione $D \times running$, che cattura l'eventuale variazione di pendenza nel periodo post-lancio rispetto al pre-lancio, risulta positivo e significativo per Ticket e Sales Net a $h=4$ e $h=10$ nei dati grezzi. Come discusso nella sezione 6.4.1, questo risultato è però contaminato dallo sbilanciamento nella composizione weekend tra i due lati della finestra e non può essere interpretato come evidenza di una reale accelerazione della traiettoria post-lancio. L'analisi di questa dimensione viene condotta in modo più credibile nella sezione 6.4.3 sui dati destagionalizzati, dove il confondente infrasettimanale è rimosso.

6.4.3 Verifica di robustezza

Come verifica di robustezza, il modello RD è stato stimato anche sui log-KPI destagionalizzati. La destagionalizzazione rimuove la componente infrasettimanale stimando per ciascun KPI, separatamente per ogni store, un modello con dummy per giorno della settimana e sottraendo gli effetti stagionali stimati preservando il livello medio della serie:

$$\log(Y)^{dest}_{it} = \log(Y)_{it} - \hat{\gamma}_i(t)$$

dove $\hat{\gamma}_i(t)$ cattura l'effetto medio del giorno della settimana stimato sullo store i . Questa procedura è un'approssimazione: una destagionalizzazione rigorosa richiederebbe dati dello stesso periodo nell'anno precedente per stimare la componente stagionale in modo separato dal segnale di interesse. In assenza di dati storici pluriennali, la stagionalità infrasettimanale viene stimata sugli stessi dati utilizzati per il modello RD, il che introduce una dipendenza tra

la fase di destagionalizzazione e quella di stima. I risultati che seguono vanno quindi letti come una verifica indicativa piuttosto che come un test formalmente indipendente.

Prima di commentare le stime, è utile portare un'ulteriore evidenza empirica rispetto l'assunzione di continuità locale su cui si fonda l'RD. L'assunzione implica che le condizioni attese nel giorno immediatamente precedente al cutoff, mercoledì 27 agosto, fossero sostanzialmente comparabili a quelle del giorno di avvio, giovedì 28 agosto, in assenza della promozione. Una verifica diretta è offerta dai coefficienti di stagionalità infrasettimanale stimati sul campione: il coefficiente per il mercoledì è $-0,022$ log-punti rispetto al lunedì, quello per il giovedì è $-0,023$, con una differenza di appena $+0,001$ log-punti, pari a circa $+0,1\%$ nel Ticket medio. Mercoledì e giovedì hanno quindi mediamente un profilo di domanda strutturalmente quasi identico, il che supporta empiricamente l'assunzione di continuità.

Passando ai risultati della versione destagionalizzata emergono tre principali conclusioni:

Tabella 6.5 - Stime RD: dati destagionalizzati

Termine	h	log(Ticket) dest.		log(Sales Net) dest.		log(UPT) dest.		log(ATV) dest.	
		Stima	p	Stima	p	Stima	p	Stima	p
τ (Effetto %)	2	+30,2% (8,6)	<0,001***	+34,0% (9,3)	<0,001***	-2,3% (3,7)	ns	+2,9% (4,0)	ns
τ (Effetto %)	4	+35,3% (8,1)	<0,001***	+33,8% (8,5)	<0,001***	-4,8% (3,0)	ns	-1,1% (3,3)	ns
τ (Effetto %)	10	+21,0% (4,8)	<0,001***	+20,9% (5,2)	<0,001***	-3,2% (1,5)	*	-0,0% (1,8)	ns
β_1 (running)	2	-0,051 (0,039)	ns	-0,070 (0,041)	.	-0,012 (0,020)	ns	-0,019 (0,022)	ns
β_1 (running)	4	-0,069 (0,027)	*	-0,055 (0,027)	*	+0,012 (0,012)	ns	+0,014 (0,013)	ns
β_1 (running)	10	-0,041 (0,007)	***	-0,034 (0,007)	***	+0,001 (0,002)	ns	+0,007 (0,002)	**
β_2 (D×running)	2	—	—	—	—	—	—	—	—
β_2 (D×running)	4	-0,083 (0,039)	*	-0,090 (0,043)	*	-0,010 (0,015)	ns	-0,007 (0,015)	ns
β_2 (D×running)	10	+0,009 (0,008)	ns	+0,012 (0,008)	ns	+0,005 (0,003)	.	+0,003 (0,004)	ns

*Nota: stessa specificazione del modello principale (RD locale lineare, pesi kernel triangolari, effetti fissi store, errori clusterizzati a livello store in parentesi). τ riportato come variazione percentuale $(\exp(\hat{\tau})-1) \times 100$ con SE via delta method. β_1 e β_2 in scala logaritmica. La destagionalizzazione è condotta stimando separatamente per ciascun KPI e per ciascun store un modello con dummy per giorno della settimana, sottraendo gli effetti stagionali stimati e preservando il livello medio. — indica collinearità a $h=2$. *** $p<0,001$; ** $p<0,01$; * $p<0,05$; . $p<0,10$; ns = non significativo.*

Il salto τ su Ticket e Sales Net è positivo, significativo e robusto in tutte le bandwidth. La rimozione del rumore infrasettimanale conferma che l'effetto al cutoff non è un artefatto della composizione dei giorni nella finestra, è un segnale reale. Con i dati destagionalizzati le stime ($h=2$ +30,2% Ticket, +34,0% Sales Net e $h=4$ +35,3% Ticket, +33,8% Sales Net) sono molto più stabili tra loro rispetto ai dati grezzi (+20,4% vs +31,4%), coerentemente con l'ipotesi che il rumore infrasettimanale amplificasse artificialmente la differenza tra finestre. L'attenuazione a $h=10$ (+21,0% Ticket, +20,9% Sales Net) potrebbe riflettere una progressiva diluizione dell'effetto di lancio all'aumentare della distanza dal cutoff.

L'UPT e l'ATV rimangono non significativi in tutte le specifiche, confermando la robustezza della conclusione sull'assenza di effetto sulla composizione dello scontrino.

Il comportamento di β_2 (Dr) cambia radicalmente rispetto ai dati grezzi, e questo è il risultato metodologicamente più informativo dell'intera sezione. Sui dati grezzi Dr era positivo e significativo per Ticket e Sales Net a $h=4$ e $h=10$, suggerendo una pendenza post-lancio più ripida di quella pre-lancio. Sui dati destagionalizzati il termine β_2 si inverte di segno rispetto ai dati grezzi: a $h=4$ diventa negativo e significativo (-0,083 per il Ticket, -0,090 per le Sales Net), mentre a $h=10$ diventa non significativo. Questo risultato smentisce la lettura dei dati grezzi, dove β_2 positivo sembrava indicare una accelerazione della traiettoria post-lancio, mostrando che quell'apparente accelerazione era interamente un artefatto del calendario. Tuttavia è necessaria una cautela interpretativa: la destagionalizzazione è stata condotta stimando la componente infrasettimanale sulla media dell'intero anno, mentre i weekend immediatamente successivi al lancio (che cadono nella fase finale dei saldi estivi) hanno quasi certamente una stagionalità strutturalmente diversa dalla media annuale. Questo introduce una distorsione residua che impedisce di interpretare le pendenze stimate sui dati destagionalizzati come le vere pendenze della domanda in quel periodo specifico. Il risultato utile non è quindi la stima precisa di β_2 , ma il fatto che l'apparente accelerazione post-lancio nei dati grezzi scompaia una volta rimossa anche solo una parte del rumore infrasettimanale, il che è sufficiente per concludere che il cambiamento di pendenza osservato nei dati grezzi non rifletteva un effetto reale della promozione sulla traiettoria dei KPI.

6.4.4 Interpretazione

I risultati dell'RD disegnano un profilo d'effetto internamente coerente e significativo attraverso tutte e tre le specifiche di bandwidth. Il lancio del sistema promozionale combinato ha generato una discontinuità positiva e significativa nel volume di traffico e nel fatturato al momento dell'avvio, mentre non ha modificato il comportamento d'acquisto in termini di composizione dello scontrino. L'effetto è quindi di natura estensiva, agisce sul margine di partecipazione e non sul margine intensivo della spesa. La convergenza tra l'incremento percentuale del Ticket e quello delle Sales Net nelle finestre più strette conferma che l'incremento del fatturato è quasi interamente attribuibile all'incremento del numero di transazioni, senza che il valore medio dello scontrino contribuisca.

Il pattern crescente delle stime con l'ampiezza della finestra nei dati grezzi è riconducibile al problema di composizione infrasettimanale discusso nella sezione 6.4.1: le finestre più ampie includono più weekend nel post-cutoff rispetto al pre-cutoff, gonfiando artificialmente la pendenza post-lancio e, di riflesso, la stima del salto. La sezione 6.4.3 ha mostrato che sui dati destagionalizzati questo pattern scompare: le stime a $h=2$ (+30,2%) e $h=4$ (+35,3%) sono molto più stabili tra loro, confermando che il rumore infrasettimanale era la causa principale della crescita osservata nei dati grezzi. La finestra $h=4$ sui dati destagionalizzati offre il miglior compromesso identificativo: è sufficientemente prossima al cutoff da rendere credibile l'assunzione di comparabilità, ed è sufficientemente ampia da ridurre il rumore residuo. La finestra $h=10$ incorpora giorni successivi alla fine dei saldi estivi, introducendo un potenziale confondente che rende quella stima meno credibile nonostante la maggiore precisione statistica.

Riguardo alla variazione di pendenza post-lancio, l'analisi sui dati destagionalizzati produce un risultato parzialmente informativo. Sui dati grezzi β_2 era positivo e significativo, suggerendo una accelerazione della traiettoria post-lancio. Sui dati destagionalizzati questo pattern scompare, β_2 si inverte di segno a $h=4$ e diventa non significativo a $h=10$, il che è sufficiente per concludere che l'accelerazione apparente nei dati grezzi era un artefatto del calendario e non un effetto reale della promozione. Non è invece possibile trarre conclusioni forti sulla vera pendenza post-lancio: la destagionalizzazione condotta sulla media annuale non cattura adeguatamente la stagionalità specifica di quei weekend di fine agosto, e altri potenziali

confondenti intorno al cutoff (come la fine dei saldi estivi) rendono le pendenze stimate non pienamente interpretabili come risposta causale alla promozione.

Un elemento di cautela riguarda la validità esterna. L'RD identifica un effetto locale al cutoff, riferito ai 37 store con doppia esposizione in un intorno ristretto dall'avvio. Questo campione non coincide con l'universo degli store analizzati nel 2WFE, e l'effetto stimato descrive la risposta al momento specifico del lancio in quella popolazione, non l'effetto medio sull'intera durata della campagna né l'effetto della sola Promozione B sui 94 store con esposizione esclusiva.

6.5 Cosa dicono i tre modelli insieme

La progressione OLS, 2WFE, RD ha prodotto un quadro coerente e internamente consistente. L'OLS catturava associazioni spurie dominate dalla selezione degli store trattati. Il 2WFE, controllando per eterogeneità strutturale e shock temporali comuni, ha restituito il primo segnale credibile: un effetto positivo e significativo sul Ticket per la Promozione B (+5,3%) e un effetto combinato su Ticket e Sales Net per i 37 store con doppia esposizione (+7,8% e +6,7%). L'RD ha aggiunto la dimensione temporale, stimando il salto al momento del lancio: circa 30% sul Ticket nelle finestre più credibili sui dati destagionalizzati.

Il dato più informativo emerge proprio dalla lettura congiunta di questi due numeri. L'effetto medio di periodo stimato dal 2WFE (+7-8%) è sensibilmente inferiore all'effetto di lancio stimato dall'RD (circa 30%). La spiegazione più coerente con entrambi i risultati è che l'effetto promozionale non sia stato uniforme lungo i 33 giorni di campagna: elevato nelle giornate immediatamente successive al lancio, è andato progressivamente scemando nelle settimane successive, fino a stabilizzarsi su un livello residuo più contenuto. In questa lettura il 2WFE cattura la media di un fenomeno che era molto più intenso all'inizio, mentre l'RD ne isola il picco iniziale.

Questa dinamica ha una implicazione operativa che vale la pena segnalare, pur con la cautela dovuta all'impossibilità di verificarla direttamente con i dati disponibili: se l'effetto incrementale di una campagna di questo tipo si concentra nei primissimi giorni, la struttura ottimale potrebbe differire da quella adottata. Campagne più brevi e più intense al lancio, eventualmente replicate in momenti diversi della stagione, potrebbero in linea di principio

generare un effetto complessivo simile a fronte di un costo promozionale distribuito su un numero inferiore di giorni. Verificare questa ipotesi richiederebbe un disegno sperimentale controllato, ad esempio un A/B test sulla durata della campagna, che i dati osservativi disponibili non consentono di condurre, ma che rappresenta una naturale estensione operativa di questa analisi.

7. Conclusioni

7.1 Sintesi dei risultati principali

La presente tesi si è proposta di quantificare l'impatto delle politiche promozionali sulle performance di vendita di un operatore retail fashion nel mercato italiano, con un duplice obiettivo: stimare in modo rigoroso l'efficacia dello sconto come leva di stimolo della domanda e identificare i contesti operativi in cui tale leva risulta più o meno efficiente.

Il risultato più robusto e trasversale all'intera analisi è che lo sconto è uno strumento efficace sul margine estensivo della domanda: genera traffico e, in misura minore, aumenta la profondità del carrello, ma non produce crescita strutturale del fatturato nella maggior parte dei contesti operativi. Questo profilo emerge con coerenza in tutti i modelli stimati e in tutte le combinazioni di contesto analizzate, e rappresenta la conclusione empiricamente più solida della tesi.

Il secondo risultato rilevante riguarda l'eterogeneità dell'effetto: la risposta della domanda allo sconto non è uniforme, ma varia in modo sistematico e prevedibile in funzione della combinazione di format commerciale, area geografica e giorno della settimana. Gli store Outlet del Centro-Nord nel weekend rappresentano il contesto di massima efficacia promozionale, con elasticità del Ticket tra le più alte osservate e un effetto netto positivo anche sul fatturato. Gli store Urban nel weekend mostrano invece il profilo opposto: elasticità del Ticket non significativa o negativa e fatturato in contrazione, un contesto in cui lo sconto non solo non genera valore incrementale ma potrebbe erodere il margine senza compensazione volumetrica. Questa variabilità non è rumore statistico: riflette differenze strutturali nella composizione della clientela e nella logica commerciale dei diversi format, e ha implicazioni dirette per la pianificazione delle campagne promozionali.

Il terzo risultato riguarda la dimensione temporale dell'effetto promozionale, emersa con particolare chiarezza dall'analisi causale della campagna cashback. L'RD stima un effetto di lancio robusto e significativo, circa +30% sul Ticket nelle finestre più credibili sui dati destagionalizzati, concentrato nei giorni immediatamente successivi all'avvio. Il confronto con il 2WFE, che stima un effetto medio di +7-8% sull'intero periodo di 33 giorni, suggerisce che l'effetto fosse più elevato nelle giornate iniziali e sia andato progressivamente scemando nel corso delle settimane successive. Questa dinamica non è direttamente identificabile dall'RD, che per costruzione osserva solo una finestra ristretta attorno al cutoff, ma è la lettura più coerente con l'insieme dei risultati.

7.2 Implicazioni operative

È opportuno precisare, prima di procedere, che le indicazioni che seguono si basano esclusivamente su effetti stimati sulla domanda, ovvero traffico, volume e fatturato netto. In assenza di dati sul costo del venduto e sui margini unitari, non è possibile trarre conclusioni dirette sulla profittabilità delle singole iniziative promozionali: un aumento del traffico o del fatturato non implica necessariamente un miglioramento della marginalità, poiché il costo dello sconto erogato potrebbe eccedere il beneficio incrementale generato. Le indicazioni operative vanno quindi lette come indicazioni sulla efficacia della domanda, non come osservazioni sulla convenienza economica complessiva.

Differenziazione per contesto. La principale implicazione operativa dell'analisi dell'eterogeneità è che una politica promozionale uniforme, lo stesso sconto applicato a tutti gli store in tutti i giorni, appare subottimale rispetto alle potenzialità del dato disponibile. I risultati mostrano che la risposta del traffico allo sconto varia in modo sistematico e prevedibile in funzione della combinazione di format, area geografica e giorno della settimana. Concentrare l'intensità promozionale nei contesti ad alta elasticità del Ticket, come gli Outlet del Centro-Nord nel weekend e i Mall nel weekend in tutte le aree, e ridurla nei contesti a bassa elasticità, come gli store Urban nel weekend dove l'effetto sul traffico è statisticamente nullo, consentirebbe di ottenere lo stesso o maggiore effetto sul volume di transazioni a parità di intensità promozionale complessiva. Questa rimodulazione avrebbe implicazioni sul costo promozionale sostenuto, ma la valutazione della convenienza netta richiederebbe l'integrazione con i dati di margine, che esulano dall'ambito di questa analisi. La granularità necessaria per implementare questa logica, a livello di format, area geografica e giorno della settimana, è esattamente quella che il dataset analizzato consente di misurare e aggiornare su base continuativa. In termini concreti, la logica di rimodulazione suggerisce due possibili strategie non mutualmente esclusive: concentrare il budget promozionale esistente nei contesti ad alta elasticità per massimizzare il traffico incrementale generato, oppure ridurre l'intensità promozionale media nei contesti a bassa elasticità mantenendo invariato l'effetto sul traffico e liberando risorse da riallocare. La scelta tra le due dipende dagli obiettivi di periodo, ma entrambe richiedono che le decisioni promozionali siano prese a un livello di granularità maggiore di quello attualmente adottato se si applicano politiche uniformi per tutti gli store.

Gestione del timing promozionale. La lettura congiunta di RD e 2WFE suggerisce che l'effetto promozionale fosse più concentrato nelle giornate di lancio di quanto la durata della campagna

lascerebbe supporre: l'effetto stimato al momento dell'avvio è sensibilmente superiore all'effetto medio di periodo, un divario coerente con l'ipotesi che l'intensità della risposta sia andata progressivamente scemando nel corso delle settimane successive. Questo disallineamento tra concentrazione temporale dell'effetto e distribuzione del costo promozionale rappresenta un elemento di inefficienza strutturale nel disegno della campagna: il costo del cashback, distribuito sull'intera durata di 33 giorni, è sostenuto in larga parte per mantenere un effetto che potrebbe essersi già in larga misura esaurito nelle prime giornate. Una struttura alternativa, campagne più brevi e replicate in momenti diversi della stagione, potrebbe in linea di principio generare un effetto complessivo analogo a fronte di un numero inferiore di giorni di costo promozionale. Verificare questa ipotesi richiederebbe un disegno sperimentale controllato che i dati osservativi disponibili non consentono, ma costituisce una direzione operativa concreta che emerge naturalmente dalla lettura congiunta dei modelli.

Valutazione degli strumenti promozionali per regime. L'analisi per categoria promozionale ha evidenziato che regimi diversi mostrano profili di efficacia marginale molto distinti. Il contrasto più netto emerge tra Mid Season Sale e Black Friday: il primo mostra elasticità elevata sul Ticket e fatturato neutro, un contesto favorevole all'intensificazione dello sconto; il secondo mostra traffico non responsivo e forte erosione del fatturato per transazione, suggerendo che in quel contesto la leva dello sconto abbia già raggiunto la saturazione e ulteriori incrementi siano controproducenti sul piano della domanda. I Saldi stagionali mostrano una progressione interna coerente: l'elasticità cresce dalla Fase 1 alla Fase 3, mentre l'impatto sul fatturato migliora progressivamente, indicando che le fasi iniziali dei saldi sono già in zona di rendimenti decrescenti rispetto alla leva dello sconto. Queste differenze suggeriscono che la calibrazione dell'intensità promozionale dovrebbe tenere conto non solo del contesto operativo ma anche del regime promozionale in cui lo sconto viene erogato, distinguendo i momenti in cui esiste ancora potenziale residuo da quelli in cui la domanda è già strutturalmente satura.

Le tre dimensioni, contesto, timing e strumento, non sono indipendenti tra loro e la loro combinazione definisce lo spazio delle decisioni promozionali più rilevanti: quale intensità di sconto applicare, in quale format e area, in quale momento della campagna e con quale meccanismo di erogazione. L'analisi presentata in questa tesi fornisce evidenza empirica su ciascuna delle tre dimensioni separatamente; la loro integrazione in un framework decisionale unitario richiede i dati di margine e un orizzonte di sperimentazione che esulano dall'ambito di questa analisi, ma rappresentano la naturale estensione operativa del lavoro svolto.

7.3 Limiti e direzioni di ricerca futura

Nonostante la solidità metodologica dell'approccio adottato, la tesi presenta alcune limitazioni che circoscrivono il perimetro interpretativo dei risultati e indicano direzioni naturali di approfondimento.

Disponibilità di serie storiche pluriennali. Un limite trasversale a tutti i modelli riguarda l'orizzonte temporale del dataset, limitato a un singolo anno solare. Con un solo anno di osservazioni, stagionalità e componente promozionale non possono essere separate in modo netto, un picco di traffico in un determinato periodo potrebbe riflettere entrambe, e le elasticità stimate riflettono le condizioni specifiche di quell'anno, senza garanzia di stabilità negli anni successivi. Questo limite si manifesta in modo particolare nell'analisi RD, dove la destagionalizzazione è stata condotta sulla media annuale anziché su dati storici dello stesso periodo, introducendo una dipendenza tra le due fasi di stima che limita l'interpretabilità dei coefficienti di pendenza.

Estensione all'analisi del margine. Il limite più rilevante dal punto di vista della rilevanza manageriale è che l'analisi misura esclusivamente gli effetti sulla domanda, ovvero traffico, volume e fatturato netto, senza considerare l'impatto sulle marginalità. Un incremento del Ticket Count del +0,69% per ogni punto percentuale di sconto aggiuntivo ha un valore commerciale molto diverso a seconda del margine associato alle transazioni incrementalmente generate. Un'estensione naturale di questa analisi consisterebbe nell'integrare i dati di margine per stimare l'effetto netto delle promozioni sulla profittabilità, distinguendo i contesti in cui lo sconto genera valore economico da quelli in cui lo erode.

Analisi a livello di singola transazione. L'aggregazione a livello store-giorno comporta una perdita di informazione sulla composizione della domanda. Lavorare su dati di singola transazione consentirebbe di distinguere se l'incremento del traffico generato dalle promozioni è attribuibile all'acquisizione di nuovi clienti, a una maggiore frequenza di visita dei clienti abituali o a un effetto di sostituzione tra giorni, meccanismi con implicazioni molto diverse per la valutazione della profittabilità delle campagne. In particolare, se il picco di traffico osservato al lancio delle promozioni cashback è prevalentemente generato da clienti di ritorno che riscattano un beneficio già maturato, il contributo incrementale netto della campagna all'acquisizione di nuova clientela è molto più limitato di quanto i dati aggregati suggeriscano.

Integrazione con dati CRM. L'integrazione dei dati operativi con il database CRM aziendale consentirebbe di distinguere le transazioni di clienti nuovi da quelle di clienti già acquisiti, di tracciare la risposta individuale alle promozioni nel tempo e di valutare l'effetto delle campagne sulla frequenza di riacquisto e sul valore del ciclo di vita del cliente. Questa integrazione trasformerebbe l'analisi da una valutazione dell'efficacia aggregata delle promozioni a una valutazione della loro efficacia differenziale per segmento di clientela, un'informazione di grande rilevanza per la progettazione di campagne più mirate ed efficienti.

Estensione geografica. L'analisi è stata condotta esclusivamente sul mercato italiano. Un'estensione ad altri mercati europei in cui l'azienda opera consentirebbe di valutare se le eterogeneità osservate sono specifiche del contesto italiano o riflettono pattern più generali, e di identificare best practice trasferibili tra mercati.

7.4 Riflessione conclusiva

Il quadro che emerge da questa tesi è che i dati operativi a livello store-giorno contengono informazione sufficiente per distinguere i contesti in cui lo sconto genera valore da quelli in cui lo trasferisce semplicemente ai clienti che sarebbero entrati comunque, per identificare i regimi promozionali con potenziale residuo da quelli già in saturazione, e per misurare la risposta di una campagna specifica sia nelle sue fasi iniziali sia nell'interezza del suo periodo. Questa tesi ha cercato di estrarre quella informazione in modo rigoroso, con la consapevolezza che le indicazioni operative derivate riguardano la domanda e non la profittabilità, e che il passo successivo (integrare i dati di margine) è necessario prima di qualsiasi decisione di riallocazione del budget promozionale.

Il contributo metodologico è forse il dato più trasferibile: la progressione OLS, 2WFE, RD, applicata a dati operativi standard senza richiedere esperimenti randomizzati, ha prodotto stime causalmente più credibili di quelle che un'analisi descrittiva avrebbe consentito. Incorporare questa logica nella progettazione ordinaria delle campagne trasformerebbe il monitoraggio delle campagne stesse da valutazione delle performance a valutazione dell'effetto incrementale reale, con un cambiamento di prospettiva che va oltre l'utilizzo di strumenti innovativi.

Bibliografia

Cameron, A. C., & Miller, D. L. (2015). A Practitioner's Guide to Cluster-Robust Inference. *Journal of Human Resources*, 50(2), 317–372.

Lee, D. S., & Lemieux, T. (2010). Regression Discontinuity Designs in Economics. *Journal of Economic Literature*, 48(2), 281–355.

Wooldridge, J. M. (2013). *Introductory econometrics: A modern approach* (Fifth edition). Cengage Learning.

Appendice

Tabella A.1 - Elasticità del Ticket Count rispetto allo sconto con IC

Combinazione	Elasticità	SE	IC 95%	Sig.
Mall Sud Feriale	+0,85%	0,091	[+0,67%; +1,03%]	***
Mall Sud Weekend	+0,87%	0,238	[+0,40%; +1,34%]	***
Mall Centro Feriale	+0,70%	0,118	[+0,47%; +0,93%]	***
Mall Centro Weekend	+1,11%	0,251	[+0,62%; +1,60%]	***
Mall Nord Feriale	+0,52%	0,100	[+0,32%; +0,72%]	***
Mall Nord Weekend	+1,17%	0,227	[+0,73%; +1,61%]	***
Urban Sud Feriale	+0,92%	0,129	[+0,67%; +1,17%]	***
Urban Sud Weekend	-0,30%	0,226	[-0,74%; +0,14%]	ns
Urban Centro Feriale	+0,62%	0,161	[+0,31%; +0,93%]	***
Urban Centro Weekend	-0,12%	0,278	[-0,67%; +0,43%]	ns
Urban Nord Feriale	+0,28%	0,187	[-0,09%; +0,64%]	ns
Urban Nord Weekend	-0,19%	0,285	[-0,75%; +0,37%]	ns
Outlet Sud Feriale	+0,34%	0,066	[+0,21%; +0,47%]	***
Outlet Sud Weekend	-0,81%	0,119	[-1,04%; -0,58%]	***
Outlet Centro Feriale	+0,49%	0,338	[-0,17%; +1,16%]	ns
Outlet Centro Weekend	+1,33%	0,200	[+0,94%; +1,72%]	***
Outlet Nord Feriale	+0,75%	0,239	[+0,28%; +1,22%]	**
Outlet Nord Weekend	+1,30%	0,310	[+0,69%; +1,91%]	***

Nota: *** $p < 0,001$; ** $p < 0,01$; * $p < 0,05$; ns = non significativo. IC costruiti come stima $\pm 1,96 \times SE$. Errori standard calcolati via delta method sulla matrice di varianza-covarianza clusterizzata a livello store.

Tabella A.2 - Elasticità dell'UPT rispetto allo sconto con IC

Combinazione	Elasticità	SE	IC 95%	Sig.
Mall Sud Feriale	+0,24%	0,030	[+0,18%; +0,30%]	***
Mall Sud Weekend	+0,24%	0,035	[+0,17%; +0,31%]	***
Mall Centro Feriale	+0,18%	0,035	[+0,11%; +0,25%]	***
Mall Centro Weekend	+0,18%	0,043	[+0,10%; +0,26%]	***
Mall Nord Feriale	+0,24%	0,027	[+0,19%; +0,29%]	***
Mall Nord Weekend	+0,24%	0,034	[+0,17%; +0,31%]	***
Urban Sud Feriale	+0,19%	0,036	[+0,12%; +0,26%]	***
Urban Sud Weekend	+0,23%	0,053	[+0,13%; +0,33%]	***
Urban Centro Feriale	+0,17%	0,043	[+0,09%; +0,25%]	***
Urban Centro Weekend	+0,16%	0,042	[+0,08%; +0,24%]	***
Urban Nord Feriale	+0,23%	0,059	[+0,11%; +0,35%]	***
Urban Nord Weekend	+0,23%	0,064	[+0,10%; +0,36%]	***
Outlet Sud Feriale	+0,54%	0,021	[+0,50%; +0,58%]	***
Outlet Sud Weekend	+0,56%	0,023	[+0,51%; +0,61%]	***
Outlet Centro Feriale	+0,67%	0,088	[+0,50%; +0,84%]	***
Outlet Centro Weekend	+0,56%	0,150	[+0,27%; +0,85%]	***
Outlet Nord Feriale	+0,19%	0,183	[-0,17%; +0,55%]	ns
Outlet Nord Weekend	+0,18%	0,188	[-0,19%; +0,55%]	ns

*Nota: *** $p < 0,001$; ** $p < 0,01$; * $p < 0,05$; ns = non significativo. IC costruiti come stima $\pm 1,96 \times SE$. Errori standard calcolati via delta method sulla matrice di varianza-covarianza clusterizzata a livello store.*

Tabella A.3 - Elasticità delle Sales Net rispetto allo sconto con IC

Combinazione	Elasticità	SE	IC 95%	Sig.
Mall Sud Feriale	+0,09%	0,099	[-0,10%; +0,28%]	ns
Mall Sud Weekend	+0,11%	0,249	[-0,38%; +0,60%]	ns
Mall Centro Feriale	-0,10%	0,119	[-0,33%; +0,13%]	ns
Mall Centro Weekend	+0,34%	0,241	[-0,13%; +0,81%]	ns
Mall Nord Feriale	-0,22%	0,103	[-0,42%; -0,02%]	*
Mall Nord Weekend	+0,46%	0,237	[-0,00%; +0,92%]	.
Urban Sud Feriale	+0,11%	0,147	[-0,18%; +0,40%]	ns
Urban Sud Weekend	-1,04%	0,212	[-1,46%; -0,62%]	***
Urban Centro Feriale	-0,12%	0,156	[-0,43%; +0,19%]	ns
Urban Centro Weekend	-0,84%	0,289	[-1,41%; -0,27%]	**
Urban Nord Feriale	-0,40%	0,187	[-0,77%; -0,03%]	*
Urban Nord Weekend	-0,83%	0,297	[-1,41%; -0,25%]	**
Outlet Sud Feriale	-0,42%	0,071	[-0,56%; -0,28%]	***
Outlet Sud Weekend	-1,45%	0,124	[-1,69%; -1,21%]	***
Outlet Centro Feriale	-0,05%	0,321	[-0,68%; +0,58%]	ns
Outlet Centro Weekend	+0,69%	0,219	[+0,26%; +1,12%]	**
Outlet Nord Feriale	-0,17%	0,173	[-0,51%; +0,17%]	ns
Outlet Nord Weekend	+0,40%	0,302	[-0,19%; +0,99%]	ns

Nota: *** $p < 0,001$; ** $p < 0,01$; * $p < 0,05$; ns = non significativo. IC costruiti come stima $\pm 1,96 \times SE$. Errori standard calcolati via delta method sulla matrice di varianza-covarianza clusterizzata a livello store.

Tabella A.4 - Elasticità dell'ATV rispetto allo sconto con IC

Combinazione	Elasticità	SE	IC 95%	Sig.
Mall Sud Feriale	-0,76%	0,040	[-0,84%; -0,68%]	***
Mall Sud Weekend	-0,76%	0,053	[-0,86%; -0,66%]	***
Mall Centro Feriale	-0,80%	0,031	[-0,86%; -0,74%]	***
Mall Centro Weekend	-0,77%	0,052	[-0,87%; -0,67%]	***
Mall Nord Feriale	-0,74%	0,034	[-0,81%; -0,67%]	***
Mall Nord Weekend	-0,71%	0,051	[-0,81%; -0,61%]	***
Urban Sud Feriale	-0,82%	0,056	[-0,93%; -0,71%]	***
Urban Sud Weekend	-0,75%	0,071	[-0,89%; -0,61%]	***
Urban Centro Feriale	-0,74%	0,039	[-0,82%; -0,66%]	***
Urban Centro Weekend	-0,72%	0,054	[-0,83%; -0,61%]	***
Urban Nord Feriale	-0,68%	0,037	[-0,75%; -0,61%]	***
Urban Nord Weekend	-0,64%	0,052	[-0,74%; -0,54%]	***
Outlet Sud Feriale	-0,76%	0,021	[-0,80%; -0,72%]	***
Outlet Sud Weekend	-0,64%	0,026	[-0,69%; -0,59%]	***
Outlet Centro Feriale	-0,54%	0,024	[-0,59%; -0,49%]	***
Outlet Centro Weekend	-0,64%	0,028	[-0,70%; -0,58%]	***
Outlet Nord Feriale	-0,91%	0,134	[-1,17%; -0,65%]	***
Outlet Nord Weekend	-0,90%	0,142	[-1,18%; -0,62%]	***

Nota: *** $p < 0,001$; ** $p < 0,01$; * $p < 0,05$; ns = non significativo. IC costruiti come stima $\pm 1,96 \times SE$. Errori standard calcolati via delta method sulla matrice di varianza-covarianza clusterizzata a livello store.