Order Now at Will, Return Later with Remorse: The Dark Side of PayTech*

Dragon Yongjun Tang† Cunyi Yang‡ Yan Zeng§

August 25, 2026

ABSTRACT

Using proprietary data on 12.9 million TikTok E-commerce orders in China, we study how buy-now-pay-later (BNPL) affects post-checkout transaction outcomes. We track each order through final after-sales resolution, allowing us to distinguish transactions subsequently reversed from purchases ultimately retained. BNPL significantly raises product returns, with the estimate accounting for approximately 8% of returns, but market-level estimates show no significant increase in retained orders or value. Consistent with payment decoupling in mental accounting, BNPL accelerates payment, its return effect weakens beginning on the repayment due date, and the effect is stronger where marginal purchases are easier to complete.

Keywords. buy-now-pay-later (BNPL); consumer credit; product returns

*We thank Haoyu Gao, Bing Han, Kai Li, Shen Lin, Fangzhou Lu, Kasper Roszbach, Tianyue Ruan, Wenfeng Wu, Liyan Yang, Xuewei Yang, and Wei Zhang for helpful comments and suggestions. We also thank seminar and conference participants at Sun Yat-sen University, The University of Osaka, the 2026 China Fintech Research Conference, the 16th Annual International Conference on the Chinese Economy, the 3rd Arne Ryde Conference on Crypto, Fintech and Payments, the 8th CSOOPEM Quantitative Finance and Insurance Annual Conference, and the 2026 INFORMS Conference on Service Science. All remaining errors are our own.

†HKU Business School, The University of Hong Kong. Email: yjtang@hku.hk

‡HKU Business School, The University of Hong Kong. Email: cunyi@hku.hk

§Lingnan College, Sun Yat-sen University. Email: zengy36@mail.sysu.edu.cn

1 Introduction

Expanding access to consumer credit is widely viewed as a means of supporting household consumption. Payment technology (PayTech) implements this idea by embedding payment, installment, and credit options directly into digital checkout; BNPL is a prominent form. Recent evidence links BNPL to consumers' financial health and merchant outcomes. deHaan, Kim, Lourie, and Zhu (2024) examine overdraft charges and credit-card costs after BNPL adoption, while Berg, Burg, Keil, and Puri (2025) study its consequences for merchants. Yet easier credit origination need not create retained demand. In returnable e-commerce, a transaction can be recorded as spending, payment volume, and gross merchandise value (GMV) at checkout before being reversed. We ask whether BNPL generates incremental retained demand after product returns or mainly expands reversible checkout activity.

Online platforms have become major retail venues. In China, online sales of physical goods reached RMB 13.09 trillion in 2025, or 26.1% of total retail sales. These checkout measures, however, need not represent demand that ultimately remains. Platforms record orders and GMV at checkout, before the return decision is realized; only after the return window closes can we distinguish retained demand from transactions that are subsequently reversed.[1] The reversal margin is economically substantial. The National Retail Federation and Happy Returns project that U.S. retailers will process $849.9 billion in returns in 2025 and that 19.3% of online sales will be returned; industry and media reports in China estimate return rates in live commerce at 30% to 50%.[2] Returns can reflect learning about product fit, post-purchase reassessment, opportunistic behavior, and the design of return channels. They also impose costs on reverse logistics, after-sales service, and inventory management. Observing returns by payment method therefore connects checkout credit to both

[1] Demand associated with purchases retained after the return decision is termed retained demand.

[2] Sources are the National Bureau of Statistics of China, 2025 retail-sales release; National Retail Federation and Happy Returns, *2025 Retail Returns Landscape*; and Economic Observer (2023). Coresight Research (2023), the European Environment Agency (2024), and People's Daily Overseas Edition (2025) document especially high return rates in online apparel.

the transaction's ultimate outcome and the resources consumed when a purchase is reversed. Testing whether FinTech credit creates retained demand therefore requires following financed transactions beyond checkout and through the return decision.

We obtain a unique order-level dataset from Douyin E-commerce, ByteDance's China-market counterpart to TikTok (hereafter TikTok). The data follow nearly 13 million transactions from checkout through payment, fulfillment, and final after-sales resolution. They cover about 390 online shops and roughly 45,600 products in China and allow us to separate checkout volume, returned activity, and retained demand along the complete transaction path. China provides a useful setting for separating embedded credit from digital-payment adoption because Alipay and WeChat Pay are mature wallet-based methods, whereas TikTok BNPL adds deferred credit within a familiar checkout environment.[3] TikTok's integration of content discovery, livestream shopping, checkout, fulfillment, and after-sales processing also exposes the return margin in the same environment where credit is offered.

We estimate whether otherwise comparable transactions are more likely to be reversed when financed with BNPL and whether market-level BNPL use generates activity that remains after returns. The baseline design compares orders for the same product, sent to the same destination street, and submitted in the same date×hour cell, while controlling for transaction characteristics. Robustness tests address observed platform context, time-varying composition, payment-method overlap, and seller-side fulfillment. Because unobserved impulsivity, liquidity needs, or perceived product fit may still affect both BNPL take-up and returns, we also use a shift–share instrumental-variables design that interacts initial province–product BNPL exposure with lagged platform-wide BNPL conditions. Under the exclusion restriction, the design identifies a local effect for orders whose payment choice responds to these platform conditions through their initial exposure.

[3] Agarwal, Ghosh, Li, and Ruan (2024) show that a demonetization-induced shift from cash to digital payments changes consumer spending behavior. Because consumers in our setting are already familiar with digital payments, we can study embedded credit without conflating it with behavioral changes arising from the initial transition away from cash.

BNPL-financed orders are more likely to be returned. In the preferred fixed-effects specification, BNPL raises realized product returns by 3.3 percentage points, or 8.0% relative to the 41.6% sample mean. The shift–share instrumental-variables (IV) design gives a two-stage least squares (2SLS) estimate of about 3.7 percentage points. Product fit, quality, fulfillment, and post-purchase information all contribute to returns; that a single checkout choice produces an effect of this size shows that payment design materially changes whether transactions ultimately remain.

The market-level decomposition provides the central economic interpretation. Higher BNPL use raises returned order counts and returned transaction value, but we find no statistically significant increase in retained orders or retained transaction value. By separating transaction initiation from transaction retention, the decomposition reveals a margin obscured by checkout-based measures such as GMV and credit originations: both can rise even when the demand that remains for merchants and platforms does not. This margin also matters for who is positioned to benefit. Credit providers can earn financing revenue when retained purchases that would otherwise use immediate payment are shifted to BNPL, and public pricing indicates that this credit carries borrowing costs well above the bank-loan benchmark.[4] The resulting incidence is therefore asymmetric.

The transaction-path evidence is most consistent with the reduced-salience channel implied by mental accounting, although product-fit learning can account for some post-delivery reversals. BNPL orders reach payment completion more quickly, and the return response falls temporarily beginning on the default repayment date. About 44% of the baseline return effect occurs before shipment, when consumers have not yet learned product fit through use. BNPL also fails to create more simultaneous candidate bundles, and existing bundles shift primarily toward complete reversal rather than selective retention. The evidence therefore

[4]Platform-disclosed fees imply annualized BNPL borrowing costs of about 20%, compared with a 3.0%–3.45% one-year loan prime rate (LPR) during our sample. Sources are the platform's checkout and repayment interfaces and National Interbank Funding Center LPR announcements authorized by the People's Bank of China.

points to payment decoupling at checkout, rather than liquidity relief or product trial alone, as the force that admits additional transactions vulnerable to reversal.

These findings make two contributions. First, we contribute to research on consumer credit and FinTech by separating gross purchase activity from demand that remains after returns. Di Maggio, Williams, and Katz (2022) document higher consumer spending after BNPL access, deHaan, Kim, Lourie, and Zhu (2024) examine consumers' downstream financial health, and Berg, Burg, Keil, and Puri (2025) study merchant revenue after BNPL adoption. Campbell and Ramadorai (2025) emphasize that the value of large administrative datasets depends on whether they capture economically relevant adjustment margins over time. By following financed transactions beyond checkout, we capture this otherwise hidden margin and show that BNPL expands checkout activity and returned activity without a commensurate increase in retained orders or retained transaction value. This distinction reveals an economically important margin concealed by gross purchase measures: digital credit can generate more recorded spending, sales, and originations without creating comparable demand that ultimately remains.

Second, we extend mental accounting by showing how consumption–payment decoupling can shape the dynamic path from purchase initiation to transaction retention. Thaler (1985) explains how consumers mentally organize financial activity, Prelec and Loewenstein (1998) show how payment timing and medium couple consumption to payment costs, and Soman (2001) demonstrates that payment mechanisms affect spending. BNPL separates current consumption from future repayment. This decoupling lowers payment salience at checkout and can allow a marginal purchase to cross the threshold; the return option then reveals whether that transaction is reconsidered. Our evidence follows this sequence: BNPL orders complete payment more quickly, reversals occur both before and after shipment, and the return effect temporarily weakens when the recurring repayment due date makes deferred payment salient again. These findings extend the empirical reach of payment decoupling from contemporaneous spending to the subsequent reconsideration, reversal, and final retention

of financed transactions.

Section 2 describes the institutional setting and develops the conceptual framework. Section 3 introduces the data and sample. Section 4 presents the order-level estimates, robustness checks, IV evidence, and market-level retained-demand analysis. Section 5 examines payment decoupling using repayment cues, checkout speed, return timing, candidate-order bundles, and cross-sectional diagnostics. Section 6 concludes.

2 Institutional Setting and Conceptual Framework

2.1 BNPL as Embedded Consumer Credit

Cornelli, Gambacorta, and Pancotto (2023) describe BNPL as point-of-sale credit that allows consumers to obtain a product immediately while deferring or dividing payment, often without interest when payments are made on schedule. In online retail, this credit is embedded directly in the checkout interface. The distinction from credit cards is institutional rather than simply contractual. Credit cards typically provide a general-purpose revolving account that can be used across participating merchants. TikTok BNPL is offered within a specific platform's checkout and combines deferred repayment with a platform credit limit, payment ordering, and platform-specific promotions; users may repay in full in the next billing cycle or choose installments.

Worldpay's *Global Payments Report 2025* estimates that BNPL accounted for 5% of global e-commerce transaction value in 2024, or $342 billion, and 4% in China. In 2024, BNPL accounted for 23% of e-commerce transaction value in Sweden, 20% in Germany, 19% in Norway, 16% in Finland, 15% in Australia, and 13% in New Zealand. In the United States, 16% of adults used BNPL in 2025, up from 10% in 2021; six large providers originated 335.8 million loans totaling $45.2 billion in 2023. The United Kingdom (UK) market grew from £0.06 billion in 2017 to more than £13 billion in 2024, and 20% of UK adults used BNPL

or deferred-payment credit in the 12 months to May 2024.[5]

China is a useful setting for separating learning about digital payments from the effects of embedded credit. Breza, Kanz, and Klapper (2026) show that using a new financial technology requires learning. Han and Wang (2026) show that economies with weaker card adoption can leapfrog to card-substituting mobile payments and identify China as a leading case. Consistent with that path, Worldpay reports that digital wallets accounted for 84% of China's e-commerce transaction value in 2024, compared with 5% for credit cards, 4% for debit and prepaid cards, and 1% for cash-like methods. The innovation in our setting is therefore not digital checkout itself. Alipay and WeChat Pay are established wallet-based methods, whereas TikTok BNPL adds deferred credit to a familiar mobile-checkout environment. Platform-disclosed installment fees imply effective annualized borrowing costs around 20%. In a developing consumer-credit market, high-cost credit can scale quickly when presented as a payment choice.

China's experience with platform credit also shapes the regulatory setting. The wave of peer-to-peer (P2P) platform failures in 2018 showed how digital credit could expand faster than disclosure, consumer understanding, and risk controls, leaving financial losses and broader social disruption. That experience made regulators particularly alert to risks from platform-based financial innovation. The concern now extends to checkout design. Under the 2026 *Measures for the Administration of Online Marketing of Financial Products*, effective September 30, 2026, non-bank payment institutions will no longer be permitted to list loans or other financial products among payment-tool options or market such products; the measures also restrict one-sided promotion of installment-fee discounts.[6] Recent reforms

[5]Sources are Worldpay, *Global Payments Report 2025*; Federal Reserve Board, *Economic Well-Being of U.S. Households in 2025* (2026); Consumer Financial Protection Bureau, *The Buy Now, Pay Later Market* (2025); and Financial Conduct Authority, *Financial Lives 2024 Survey: Payments* (2026) and "New Protections Confirmed for Buy Now Pay Later Borrowers" (2026).

[6]Sources are the Internet Financial Risk Special Rectification Work Leading Group Office and the P2P Network Lending Risk Special Rectification Work Leading Group Office, statement on P2P risk resolution (August 6, 2018), and People's Bank of China et al., *Measures for the Administration of Online Marketing of Financial Products*, Announcement No. 9 [2026].

in the European Union, the United Kingdom, and Australia instead regulate BNPL through disclosure, affordability, and licensing requirements.[7]

Our setting is TikTok BNPL, officially branded in China as Douyin Monthly Payment. The product is an in-app "consume-first, pay-later" option on TikTok E-commerce; in the default billing configuration studied here, statements are generated on the 1st of the month and repayment is due on the 6th.[8] The key institutional feature for our question is the integration of credit, payment, and returns. Online purchases are not final at checkout because the platform's after-sales system allows consumers to reverse transactions and trigger refunds. With deferred payment, a return can also cancel or offset future repayment obligations. BNPL can therefore affect both the decision to buy and the later decision to keep the order. Figure 1 summarizes the order-level linkage among checkout, credit approval, merchant settlement, repayment, and after-sales processing.

[Figure 1 about here.]

The institutional flow creates different timing and exposure across participants. Consumers receive immediate possession and payment deferral, and the return option allows them to undo a transaction under platform rules. Merchants and platforms first observe completed checkout activity and GMV, whereas payment and credit providers intermediate the financed flow. The return margin is resolved only later. A returned order can reverse recorded revenue or settlement, use after-sales capacity, create reverse-logistics and inventory costs, and cancel or offset the consumer's repayment obligation.[9] Whether checkout activity becomes retained demand or returned activity with after-sales frictions thus determines what remains for merchants and platforms after checkout. This distinction motivates our separate

[7]Sources are European Commission, "Consumer Credit" (2026); Financial Conduct Authority, "Regulating Buy Now Pay Later" (2026); and Australian Securities and Investments Commission, "Buy Now Pay Later Credit Contracts: Credit Licensing" (2025).

[8]First-time BNPL users complete activation and credit-related authorizations. Repayment is executed through the platform's payment service and can be manual or automatic.

[9]McKinsey & Company, *Deconstructing Silos to Discover Savings: The End-to-End Excellence Playbook for Retailers* (2024); Deloitte, *Consumer Industry Briefing: Fewer Fashion Returns with AI* (2026).

measures of realized product returns, returned transaction value, and retained demand. Section 2.3 develops the mental-accounting interpretation of the separation between purchase and repayment.

2.2 Related Research

The closest literature studies BNPL as embedded checkout credit. Di Maggio, Williams, and Katz (2022) document that BNPL access increases spending and shifts expenditure toward retail goods. Bian, Cong, and Ji (2024) examine how e-wallet super-apps combine payments and BNPL to expand payment and credit access. On consumers' downstream financial health, deHaan, Kim, Lourie, and Zhu (2024) examine bank overdraft charges and credit-card costs after BNPL adoption, while BNPL demand also responds to contract design and credit-reporting features (Aidala, Koşar, Mangrum, and van der Klaauw, 2025). On the merchant side, Berg, Burg, Keil, and Puri (2025) study adoption and revenue. Merchant effects can vary with firm constraints and market structure (Alok, Ghosh, Kalda, and Mitra, 2025). We study what happens between checkout and repayment: whether a financed transaction is retained or unwound through a product return.

A related literature studies information and contract design in consumer credit. Pagano and Jappelli (1993) analyze how information sharing affects credit-market screening. Norden and Weber (2010) document that credit-line use and checking-account activity improve default prediction. Jansen, Nagel, Yannelis, and Zhang (2025) examine the welfare effects of borrower data. BNPL-specific work extends this question to credit reporting and bank use of BNPL histories. Dong et al. (2025) examine how national credit-reporting integration changes BNPL use and repayment, while Laudenbach, Molin, Roszbach, and Sondershaus (2025) show how banks use private BNPL histories in loan approval and pricing. Contract evidence shows in turn how liquidity constraints, screening, and repayment incentives shape financed purchases. Adams, Einav, and Levin (2009) study subprime auto credit. Field evidence separates selection from repayment behavior (Karlan and Zinman, 2009), while Einav,

Jenkins, and Levin (2012) analyze pricing and contract choice.

Consumer-credit research identifies countervailing effects of easier checkout credit. Payday credit can improve household resilience when liquidity shocks bind (Morse, 2011). Access to payday loans can also worsen household outcomes (Melzer, 2011); Gathergood, Guttman-Kenney, and Hunt (2019) find brief initial relief followed by persistently worse credit outcomes in the United Kingdom. Dollar-cost disclosures can reduce payday borrowing (Bertrand and Morse, 2011), and Di Maggio, Ma, and Williams (2025) show that less aggressive overdraft practices reduce payday borrowing and improve consumers' financial health. Consumer-protection interventions can also depend on context. In a preregistered experiment, Shema-Zlatokrilov (2026) finds that a risk prompt did not significantly reduce installment take-up, but users who learned that the prompt had been removed reported lower post-purchase satisfaction. Together, this evidence makes both liquidity relief and behavioral responses to delayed payment costs plausible in the BNPL setting. We develop their different predictions for retained demand below.

Mental accounting and payment design provide the theoretical foundation for our mechanism. Consumers mentally organize and track financial activity (Thaler, 1985). Prelec and Loewenstein (1998) formalize consumption–payment coupling, and payment immediacy and rehearsal also affect spending (Soman, 2001). Evidence from taxation shows how salience changes behavior. Finkelstein (2009) finds that electronic toll collection reduces drivers' responsiveness to tolls, while Chetty, Looney, and Kroft (2009) show that making the full tax-inclusive price salient reduces demand. In digital payments, Agarwal, Ghosh, Li, and Ruan (2024) show that a demonetization-induced shift from cash to digital payments raises spending, while spending patterns also reflect mental accounts (Gelman and Roussanov, 2024). These studies motivate reduced payment salience as the empirical channel through which BNPL may affect purchase completion and later reversal.

Several adjacent behavioral frameworks also make delayed costs consequential. Limited attention affects overdraft fees (Stango and Zinman, 2014), while Ponce, Seira, and Zamar-

ripa (2017) study costly credit-card debt allocation. Bu, Hanspal, Liao, and Liu (2022) find that self-control and budgeting interventions can reduce online borrowing and delinquency charges. Present-bias models make immediate benefits more influential than delayed costs (Laibson, 1997), and related work characterizes the resulting timing choices (O'Donoghue and Rabin, 1999). Gabaix and Laibson (2006) show how firms can obscure add-on costs, while related credit models study naive borrowers (Heidhues and Kőszegi, 2010). Keil and Burg (2023) provide a direct BNPL benchmark by studying impulse shopping. These frameworks identify related channels through which delayed costs shape borrowing; we use mental accounting to connect payment timing to purchase completion and later reversal.

Research on product returns supplies the post-purchase resolution margin. Petersen and Kumar (2009) place product returns within the broader firm-customer exchange process and document their antecedents and consequences. Altug, Aydinliyim, and Jain (2021) analyze opportunistic returns, while Nageswaran, Hwang, and Cho (2025) examine how return-channel design shapes operating frictions. After delivery, returns allow consumers to update on fit or quality and reverse low-match purchases; low return frictions can also induce opportunistic behavior and impose reverse-logistics and inventory costs. We complement this literature by allowing the payment arrangement at checkout to vary and examining how embedded credit changes the use of the return option and the conversion of checkout activity into retained demand.

2.3 Payment Decoupling and Retained Demand

The economic question is whether activity initiated with BNPL remains after product returns are resolved. At the market-cell level, gross orders consist of returned orders and retained orders, while gross transaction value consists of returned transaction value and retained transaction value. This accounting distinction separates activity recorded at checkout from the portion retained after product returns. To explain how payment design can change this gross-to-retained split, we turn to the double-entry mental-accounting framework of

Prelec and Loewenstein (1998). The framework treats consumption benefits and payment costs as connected mental entries but allows payment design to change the strength of that connection. Immediate settlement keeps the monetary outflow close to the product choice. BNPL instead separates order completion from repayment and can weaken the payment entry evoked at checkout. We refer to this reduction in consumption–payment coupling as payment decoupling; reduced payment salience, or a weaker pain of paying, is its observable channel.

The return option makes the consequences of decoupling visible along the transaction path. If BNPL weakens payment coupling, purchases that would not clear a more salient assessment of their full cost can cross the checkout margin and later be reversed. The first prediction is therefore that BNPL shifts completed purchases toward orders with a lower probability of retention.

Repayment cues sharpen this prediction. A salient repayment landmark can restore the connection between a current purchase and its deferred cost and discourage the most marginal orders. The second prediction is that the BNPL return effect weakens for orders placed when repayment becomes especially salient.

This logic also sharpens the market decomposition above. The retained component of incremental checkout activity depends on how many induced orders are retained after the return decision. When weaker coupling primarily admits purchases near the checkout margin, returned activity can rise substantially while the retained increment remains small. The third prediction is therefore a wedge between gross and returned activity on one side and retained demand on the other. Transaction value and purchase environments help diagnose this mechanism, although their signs also depend on fit uncertainty and return frictions.

The mental-accounting channel is behavioral: it departs from a frictionless rational benchmark by allowing payment timing and presentation to affect checkout even when lifetime resources and product attributes are unchanged. Under the rational benchmark, shifting payment timing alone should not change real purchase decisions (Hall, 1978). Two rational

accounts imply different transaction paths. First, short-run liquidity relief lets constrained consumers obtain products they prefer but could not otherwise finance. Cash on hand and available credit matter when constraints bind (Gross and Souleles, 2002), and credit rationing can prevent otherwise desirable purchases (Stiglitz and Weiss, 1981). This channel should improve match quality and make BNPL-induced purchases more likely to be retained, lowering rather than raising the return rate. Second, rational product-fit learning treats the return option as a way to resolve uncertainty after purchase. By reducing the immediate financing cost of trial, BNPL may encourage consumers to order items whose fit is uncertain, including multiple variants of the same product. This multi-item practice is commonly termed bracketing: consumers order several sizes, colors, or related products, keep the preferred item, and return the rest. It predicts post-delivery learning, more simultaneous candidate bundles, and selective rather than complete bundle returns. Figure 2 summarizes the three paths and their distinct implications for transaction resolution and retained demand.

[Figure 2 about here.]

The empirical analysis maps these predictions to separate stages of the transaction path. Section 4 first establishes the order-level relation between BNPL and transaction reversal, addresses selection into payment method, and tests whether incremental checkout activity becomes retained demand. Section 5.1 examines whether the return effect changes when a recurring repayment landmark restores the salience of deferred payment. Section 5.2 relates checkout speed to eventual returns and separates reversals before and after shipment. Section 5.3 tests whether BNPL creates candidate-order bundles and selective retention, while Section 5.4 uses cross-sectional patterns to locate boundary conditions and settings in which marginal purchases are easier to complete.

3 Data and Sample

The empirical unit is an order that links payment choice at checkout to payment completion, fulfillment, and the final after-sales outcome of the same transaction. This complete transaction path lets us study post-purchase reversals conditional on checkout conditions and fulfillment commitments recorded before the return decision.

3.1 Data Source

We obtain proprietary transaction data from a large third-party digital service provider that supports e-commerce transaction certification and automation for online merchants. For merchants connected to the provider's digital infrastructure, the data follow TikTok E-commerce orders from checkout through final after-sales resolution.

The raw data record the payment method used at checkout, including BNPL, and connect it to product and shop identifiers, price and discount components, shipping fees, logistics variables, and high-frequency timestamps for order submission, payment completion, and shipment. Product and shop identifiers are the platform's product and merchant identifiers. Using product identifiers, names, and specifications, we map products to two-digit consumption-purpose categories from the United Nations' Classification of Individual Consumption According to Purpose (COICOP); each order inherits its product's category. Street denotes the destination-street identifier, and date×hour cells use the calendar date and hour of order submission. The platform's after-sales records reveal whether the order is ultimately kept or reversed.

The platform records after-sales outcomes at the order level, matching the empirical unit to its operational record. The raw dataset spans June 2024 to March 2026 and contains about 20 million orders across a broad set of shops, products, categories, and destinations within China.

3.2 Variables and Measurement

Our primary outcome, *Return*, equals one when the platform's after-sales system records a finalized consumer-initiated return or refund. As described in Section 2.1, $Return = 1$ denotes an order-level reversal recorded by the platform, including both refund-only cases and cases involving a physical product return. Within the finalized-outcome sample, $Return = 0$ denotes an order that is kept or reaches a resolved after-sales state without reimbursement.[10] We multiply the indicator by 100 in all regressions and tables while retaining the notation *Return*, so coefficients are measured in percentage points.

Our key explanatory variable, *BNPL*, equals one when an order is paid using TikTok BNPL and zero otherwise. The order-level flag is available for essentially all observations in the cleaned data. The data identify the checkout channel recorded for each order, but not the funding source selected within an external wallet. An Alipay-coded order may therefore be funded by a wallet balance, a linked bank or credit card, or Huabei; the data do not separate these sources. TikTok BNPL, by contrast, is directly identified as a distinct order-level channel. Section 4.2 evaluates this measurement boundary by excluding Alipay from the comparison group.[11] Figure 3 shows how BNPL appears alongside other payment methods at checkout.

[Figure 3 about here.]

We validate the payment-method flag around two public TikTok promotional openings in 2025, covering the 618 and Double 11 campaigns (Internet Appendix Table IA.1). Platform materials and public reports describe both openings as including BNPL installment support or event-specific interest-free installment discounts. Around both dates, measured BNPL take-up rises more in province–product cells with greater predetermined BNPL exposure.

[10] We exclude unresolved after-sales cases rather than code them as nonreturns, because product-return outcomes are observed only after the platform closes the after-sales process.

[11] Sources are Alipay Open Platform documentation for merchant checkout and Huabei, and TikTok Mall, *User Service Agreement*.

This pattern links the order-level flag to the platform's embedded BNPL product rather than to a generic payment-code artifact.[12]

The order-level controls account for transaction characteristics observed at checkout that may jointly shape payment choice and the probability of a return. Unit price and quantity capture the scale of the purchase. Higher-priced orders carry greater financial stakes, while multi-unit orders offer more scope to retain only part of the purchase. Shipping fees and discounts capture the transaction's effective terms and promotional intensity; both can alter product selection and the economic stakes of reversing an order. Promised shipping lag and local shipping capture expected fulfillment speed and shipping distance. Longer waits and more distant fulfillment increase delivery uncertainty and the likelihood that an item arrives after its value to the consumer has declined, raising cancellation or return risk independently of payment choice. Price and quantity enter as *Log Price* and *Log Quantity*; shipping fees and discounts enter as *Log Shipping Fee* and *Log Discount*; and fulfillment enters as *Log Shipping Lag* and *Local Shipping*. Appendix A gives the transformations and definitions and describes the additional measures used in the mechanism tests.

3.3 Sample Construction

We construct the estimation sample sequentially from the provider's panel of 19,469,439 orders. First, we require strictly positive unit prices and quantities. This screen removes 1,758,708 orders, including zero-price gifts, extreme promotional items, and records that are not comparable to standard purchases, leaving 17,710,731 orders.

Second, we require the payment-method indicator, baseline controls, product identifier, destination-street identifier, and date×hour cell to be observed. Before applying this complete-case requirement, we set negative shipping fees, discounts, and payable amounts to missing because these monetary components should be weakly nonnegative. We likewise set

[12]Contemporaneous Chinese reports and campaign materials identify May 13 and October 9 as the two campaign openings and describe the officially branded Douyin Monthly Payment installment benefits.

negative or implausibly large promised shipping lags to missing. These requirements remove another 1,400,342 orders and leave 16,310,389 otherwise eligible orders. Among retained observations, the promised shipping lag is winsorized at the 1st and 99th percentiles to limit the influence of timestamp irregularities. Table 1 reports the resulting lag in hours; the regressions use the natural logarithm of one plus this value.

Third, we require the order's after-sales outcome to be finalized. The platform's return and refund process may remain open after purchase, especially for orders submitted near the end of the sample period. We therefore exclude 3,419,605 otherwise eligible orders whose final return status is unresolved rather than coding them as retained purchases. This final screen leaves 12,890,784 orders from about 390 online shops and roughly 45,600 products. Internet Appendix Table IA.2 reports the sequential screens and outcome-resolution rates separately for BNPL and non-BNPL orders. The older-cohort specification in Table 3 is a separate robustness check: restricting the finalized sample to orders submitted on or before January 10, 2026, removes 713,596 recent orders and leaves 12,177,188 observations.

3.4 Descriptive Statistics

Panel A of Table 1 describes the outcomes, payment choice, and transaction characteristics used in the baseline analysis. The table reports the transaction characteristics in their original units, while the regressions use the logarithmic transformations defined in Appendix A. Product returns occur in 41.61% of finalized orders, and BNPL is used in 26.21%. Orders are typically single-unit purchases, shipping fees are usually zero, and discounts vary considerably. The promised shipping lag has a median of 48 hours and a mean of 111 hours, indicating a long right tail.

[Table 1 about here.]

Panel B of Table 1 compares observable characteristics across payment choices. Before controls, the product-return rate is 10.07 percentage points higher for BNPL orders than for non-BNPL orders. BNPL orders also have higher unit prices, larger discounts, and promised

shipping lags that are 37 hours longer on average. Differences in quantity, shipping fees, and local shipping are statistically significant but economically small. Across specific checkout methods, product-return rates are 40.29% for the platform's payment service, 40.06% for Alipay, and 34.52% for WeChat Pay, compared with 49.04% for TikTok BNPL. These differences motivate specifications that compare transactions within narrow product, location, and purchase-time cells.

Destination-city order volume closely follows China's population geography (Internet Appendix Figure IA.1). The city-level rank correlation between order volume and 2024 population is 0.94. Order volume is highest in the direct-administered municipalities, provincial capitals, and other large urban markets,[13] and is thinner in sparsely populated western regions. Together with the presence of orders across the country, this pattern documents broad national coverage rather than concentration in a small set of platform markets. The resulting geographic breadth permits comparisons within fine destination markets across China.

Panels A and B of Internet Appendix Figure IA.2 show that product returns and BNPL take-up both exhibit cross-city spatial clustering. Panel C plots the two city-level rates and shows a positive unconditional association across the 265 cities with at least 10,000 orders (Pearson correlation = 0.260; Spearman correlation = 0.227). Because local payment adoption and return behavior covary, unconditional payment-method comparisons may confound BNPL use with local demand and return environments. The order-level specifications therefore absorb destination-street fixed effects (FE) and compare transactions within narrow local markets.

The unconditional BNPL product-return rate is higher in most categories, but the gap varies across the product space (Internet Appendix Figure IA.3). This variation motivates the later tests of whether the BNPL effect is stronger where experimentation, fit uncertainty, or discretionary spending makes marginal purchases more likely.

[13] China's provinces are first-level administrative divisions, roughly comparable to U.S. states. Beijing, Shanghai, Tianjin, and Chongqing are municipalities at the provincial level; provincial capitals serve as the administrative seats of the provinces.

BNPL orders have higher unconditional product-return rates in every sample month (Internet Appendix Figure IA.4). The persistence of the raw gap is descriptive rather than identifying because payment choice, product mix, location, and platform conditions also change over time. The regressions below address these dimensions with product, destination-street, and date-hour FE.

4 BNPL, Transaction Reversal, and Retained Demand

This section examines whether observed orders are more likely to be reversed when paid with BNPL. The central empirical problem is that consumers choose the payment method at checkout, so an unconditional comparison can combine payment deferral with impulsivity, liquidity needs, expected match quality, and other determinants of product returns. We first measure how the estimate changes as the comparison becomes more restrictive and then use a shift–share IV design to address unobserved selection. Unless otherwise noted, the analysis uses the final estimation sample.[14] Section 4.4 complements the order-level estimates with a market-level analysis of retained demand, and Section 5 examines behavior consistent with the competing mechanisms.

4.1 Baseline Regression Results

The baseline specification is an order-level linear probability model. The dependent variable is *Return*, expressed in percentage points as defined in Section 3.2, and *BNPL* indicates payment with BNPL. The specification is

$$Return_i = \beta\, BNPL_i + \boldsymbol{\gamma}'\mathbf{X}_i + \text{Fixed Effects} + \varepsilon_i, \tag{1}$$

[14] Throughout the regression tables, observations refer to valid input observations before fixed-effect singleton groups are removed by the estimator. Differences across specifications therefore reflect outcome, regressor, instrument, or sample availability rather than singleton deletion.

where $\mathbf{X}_i$ contains the six transaction characteristics defined in Section 3.2. Price and promotion controls are *Log Price*, *Log Quantity*, *Log Shipping Fee*, and *Log Discount*. Fulfillment controls are *Log Shipping Lag* and *Local Shipping*. The fixed effects absorb geographic, time, product, and seller heterogeneity.

Table 2 reports estimates from Eq. (1). Column (1) expresses the raw gap from Panel B of Table 1 as a regression coefficient: the product-return rate is 10.07 percentage points higher for BNPL transactions than for non-BNPL transactions. This difference establishes the scale of the unconditional relation, but it does not separate payment choice from composition across products, sellers, and purchase occasions.

[Table 2 about here.]

Column (2) adds the six transaction controls without fixed effects. The BNPL coefficient falls to 3.56 percentage points but remains statistically significant, indicating that observed transaction characteristics account for a substantial part of the raw gap. Column (3) further absorbs destination-street and date-hour FE. The estimate remains similar at 3.50 percentage points.

Columns (4)–(6) add category, shop, and product fixed effects, respectively, to the destination-street and date-hour effects. The estimates remain between 3.3 and 3.6 percentage points, and the shop- and product-FE specifications show separately that the relation survives comparisons within sellers and within products. Column (6), which also absorbs destination-street and purchase-time heterogeneity, yields 3.32 percentage points, equal to about 8% of the 41.6% sample mean. Product returns reflect many aspects of the firm-customer exchange, including fit, quality, fulfillment, post-purchase information, and return-channel design (Petersen and Kumar, 2009); analytical work also shows how return policies can induce opportunistic use (Altug, Aydinliyim, and Jain, 2021). That a single payment arrangement shifts the final status of transactions by this much makes payment design economically consequential. It also raises the question taken up below: does BNPL stabilize purchases that consumers value and retain, or does it admit marginal transactions that are

later reversed? The shift–share design targets remaining unobserved selection, while the market-level and transaction-path analyses in Sections 4.4 and 5 distinguish these economic interpretations.

The control coefficients describe the composition margins absorbed by the specification. In the preferred product-FE specification, product returns are more prevalent for higher-priced and larger-quantity purchases and for orders with longer promised shipping lags, whereas higher discounts and local shipping are associated with fewer product returns. These conditional associations reinforce the need for narrow comparisons. The preferred coefficient remains a conditional relation; the specification checks and IV design below tighten the comparison and introduce additional identifying variation.

4.2 Alternative Samples and Specifications

Table 3 organizes the robustness checks around five threats: observed platform context, time-varying composition, sample comparability, outcome measurement or fulfillment, and product composition. Each threat implies a different reason why BNPL orders could appear more likely to be returned. The specifications change the relevant comparison while leaving the payment indicator unchanged and, except for the fulfillment placebo, retain realized product returns as the outcome.

[Table 3 about here.]

The first set of checks asks whether platform placement or direct checkout incentives jointly raise BNPL use and returns. Excluding orders with explicit payment-specific discounts removes transactions directly subsidized for using a particular method and leaves a BNPL coefficient of 3.46 percentage points in Column (1). Absorbing traffic source, traffic type, traffic channel, and advertising channel yields a coefficient of 3.21 percentage points in Column (2). The small movement in these estimates indicates that observed payment subsidies and measured traffic routing do not account for the baseline relation.

External-wallet funding creates a related measurement concern. An Alipay-coded order may be funded by a wallet balance, a bank or credit card, or Huabei, while explicit platform discounts and major shopping campaigns may jointly affect payment choice and returns. Internet Appendix Table IA.3 isolates these channels. Excluding Alipay and the separately coded pay-later method yields a fixed-effects estimate of 3.11 percentage points; comparisons of TikTok BNPL only with TikTok Pay or WeChat Pay yield 2.58 and 4.28 percentage points. Removing Alipay, observable payment discounts, and the three major promotion windows produces an estimate of 3.28 percentage points. When both initial exposure and the national shift are reconstructed within the restricted samples, the corresponding 2SLS estimates are 3.24 and 3.27 percentage points, with K-P Wald F statistics above 1,400. Thus, potential Huabei use within Alipay and observable platform promotions do not account for the main result. These estimates identify TikTok BNPL as offered, including the checkout placement and product design that define it, rather than a credit contract detached from the platform environment.

The next specifications absorb increasingly local time-varying shocks. Product×date, shop×date, and category×date×hour fixed effects absorb daily product shocks, shop-day campaigns or fulfillment conditions, and high-frequency category demand shocks in Columns (3)–(5). The resulting estimates range from 3.22 to 3.30 percentage points, close to the baseline estimate of 3.32 percentage points. Across Columns (2)–(5), the baseline relation remains when the comparison absorbs measured platform routing and unusually high-return days for particular products, shops, or categories.

The overlap restriction in Column (6) addresses a different concern: the baseline estimate may compare cells in which only one payment method is observed. Requiring both BNPL and non-BNPL orders within product×street×date×hour cells produces an estimate of 4.22 percentage points. The smaller sample reflects the overlap requirement rather than a change in the outcome definition, and the larger estimate shows that the baseline coefficient does not depend on one-sided payment-method cells.

The final two specifications examine outcome measurement and seller-side fulfillment. Column (7) retains orders submitted on or before January 10, 2026, giving every observation at least a 60-day window before the panel ends, and produces a BNPL coefficient of 3.35 percentage points. Replacing product returns with late shipment yields a coefficient of −0.01 percentage points in Column (8), statistically indistinguishable from zero. The stability of the older-cohort estimate and the late-shipment placebo weigh against incomplete resolution of recent orders or measured deterioration in merchant fulfillment as explanations for the return estimate.

Column (9) addresses a specific composition concern. Women's apparel represents a large share of e-commerce transactions and has particularly high return rates, raising the possibility that the aggregate BNPL estimate is driven by a narrow segment of female-oriented goods. We classify products as male- or female-oriented from explicit signals in titles and specifications and exclude mixed or ambiguous products. In this classified sample, the BNPL coefficient for male-oriented products is 3.94 percentage points, while the female-minus-male interaction is −0.07 percentage points and statistically insignificant. The aggregate relation therefore does not appear to be specific to female-oriented goods.

More broadly, the estimates in this subsection are stable across checks targeted at observable composition and measurement. The next subsection uses the shift–share design to address remaining unobserved selection into BNPL.

4.3 Shift–Share Instrumental-Variables Evidence

The remaining identification challenge is endogenous payment choice. Even within product, destination-street, and date×hour cells, unobserved liquidity, product fit, or purchase intent may affect both BNPL use and product returns. We therefore use a shift–share instrument that combines predetermined segment exposure with lagged platform-wide BNPL conditions.

Instrument construction. The instrument has three components: held-fixed initial exposure,

a lagged platform-wide shift, and their interaction. Initial exposure captures a province–product segment's BNPL adoption propensity and is neither a response elasticity nor an exogenous shock in its own right. Consistent with evidence that contract terms affect BNPL take-up (Aidala et al., 2025), we measure exposure for each province–product pair (p, g) during its first seven days in the data:

$$BNPL_{pg}^{\text{base}} = \frac{1}{|\mathcal{T}_{pg}^{0}|} \sum_{t \in \mathcal{T}_{pg}^{0}} BNPL_{pgt}, \tag{2}$$

where $\mathcal{T}_{pg}^{0}$ contains date×hour cells in the initial window and $BNPL_{pgt}$ is average take-up in cell (p, g, t). This exposure measure is held fixed throughout estimation.

The second component is the national BNPL take-up rate at the date×hour level, lagged by 24 hours:

$$BNPL_{t-24}^{\text{nat}} = \frac{1}{N_{t-24}} \sum_{i \in \Omega_{t-24}} BNPL_{i,t-24}, \tag{3}$$

where Ω_{t-24} contains transactions in time cell $t - 24$. This series proxies for platform-wide availability and capacity as well as payment salience and promotion.[15] The lag establishes temporal ordering relative to the focal checkout decision; persistent aggregate conditions remain part of the exclusion analysis below.

The third step interacts the lagged aggregate shift with initial exposure:

$$Z_{pgt}^{\text{SS}} = BNPL_{t-24}^{\text{nat}} \times BNPL_{pg}^{\text{base}}. \tag{4}$$

Date-hour FE absorb the common aggregate movement, so identification comes from differential take-up responses across segments with different initial exposure. Product and street FE absorb additive product and location differences. Initial exposure also predicts subsequent BNPL take-up monotonically (Internet Appendix Figure IA.5); the first stage below

[15] We have observed that TikTok changes payment-method ordering, BNPL discounts, and campaign promotion over time.

evaluates the relevance of the complete instrument.

Identification and estimation. Identification requires relevance, exclusion, and monotonicity. Relevance requires aggregate conditions to shift BNPL use more strongly in higher-exposure segments. Exclusion requires these conditions to affect returns differentially across segments only through BNPL use; date-hour FE remove common national movements but not shocks whose loadings covary with initial exposure. Monotonicity requires favorable aggregate conditions not to move BNPL use in opposite directions across responsive orders. Under these conditions, 2SLS identifies a weighted local effect for orders whose payment choice responds to the instrument.

Following Di Maggio, Williams, and Katz (2022), we report the first stage, reduced form, and 2SLS in sequence:

$$BNPL_{ipgst} = \pi Z^{\mathrm{SS}}_{pgt} + \mathbf{X}'_{ipgst}\boldsymbol{\delta} + \alpha_g + \gamma_s + \lambda_t + u_{ipgst}, \tag{5}$$

$$Return_{ipgst} = \rho Z^{\mathrm{SS}}_{pgt} + \mathbf{X}'_{ipgst}\boldsymbol{\kappa} + \alpha_g + \gamma_s + \lambda_t + \eta_{ipgst}, \tag{6}$$

$$Return_{ipgst} = \beta \widehat{BNPL}_{ipgst} + \mathbf{X}'_{ipgst}\boldsymbol{\theta} + \alpha_g + \gamma_s + \lambda_t + \varepsilon_{ipgst}. \tag{7}$$

All three equations include the baseline controls and product, destination-street, and date-hour FE; standard errors are clustered at the product level.

Results. Table 4 reports the central results.

[Table 4 about here.]

The first-stage coefficient is 1.16 (standard error 0.03), and the Kleibergen–Paap rk Wald F statistic is 1458.83. The reduced-form coefficient is 4.30 (standard error 0.47), yielding a 2SLS estimate of 3.71 percentage points (standard error 0.40). Thus, exposure-weighted variation that strongly raises BNPL take-up also raises product returns. The IV effect equals about 8.9% of the 41.6% mean return rate and is close in economic magnitude to the preferred fixed-effects estimate of 3.32 percentage points. The estimands nevertheless differ:

the fixed-effects coefficient is a conditional relation, whereas 2SLS identifies a weighted local effect for instrument-responsive orders.

Following Degryse, Roukny, and Tielens (2026), Columns (4)–(5) hold aggregate composition fixed and remove the focal category or both the focal category and province from the shift. These variants retain within-segment take-up movements while limiting contemporaneous reweighting and mechanical own-market contributions; their construction is detailed in Internet Appendix Table IA.5. The corresponding estimates are 4.00 and 4.08 percentage points, with K-P rk Wald F statistics of 1055.02 and 969.79. Requiring at least five initial-window orders yields 2.82 percentage points. The similarity in economic magnitude across these designs reduces concern that the baseline estimate is driven by shifting composition, own-market contributions, or sparse exposure cells.

Threat-based exclusion diagnostics. We first address aggregate composition, public campaigns, and own-market contamination. Excluding major promotion windows gives estimates of 4.14 and 4.20 percentage points, while trimming the fixed-category shift yields 3.86–3.99 percentage points. Predetermined category weights also reduce the shift's correlations with lagged national prices and discounts (Internet Appendix Table IA.5). Together with the leave-one-out results, these patterns make contemporaneous product mix, campaigns, extreme aggregate movements, and mechanical own-market contributions less plausible explanations.

We next examine shocks whose loadings may covary with initial exposure, the central concern in formula-based instruments (Borusyak and Hull, 2023). With baseline-exposure, province-specific, and category-specific trends, reduced-form-to-first-stage ratios remain between 3.65 and 3.78 (Internet Appendix Table IA.7). A lead-innovation diagnostic is statistically imprecise, while placebo-outcome tests show little relation between the instrument and checkout attributes, payment discounts, or realized fulfillment (Internet Appendix Tables IA.6 and IA.8). Following Danieli et al. (2026), we also verify that historical product return measures predict current returns before using them as negative controls. The fixed-

category leave-one-out instrument predicts neither historical measure; both coefficients round to 0.00, with standard errors of 0.01 (Internet Appendix Table IA.9).

Finally, first-stage contribution weights are dispersed across many shift and exposure components rather than concentrated in a small influential set (Goldsmith-Pinkham et al., 2020). Alternative exposure windows and 12-, 24-, and 48-hour lags yield estimates of 3.51–3.71 percentage points, and alternative support, leave-one-out, and clustering choices preserve the positive inference (Internet Appendix Tables IA.10, IA.11, and IA.12). A bounded-direct-effect analysis following Conley, Hansen, and Rossi (2012) gives an Anderson–Rubin 95% confidence set of [2.92, 4.50] under strict exclusion; the set first includes zero only when a positive direct effect reaches 78.7% of the reduced form. Calibrating the direct effect to the observed late-shipment placebo association, 6.96% of the return reduced form, leaves a 95% confidence set of [2.67, 4.50].

Interpretation. Under the maintained relevance, exclusion, and monotonicity conditions, the baseline, predetermined-composition, and diagnostic evidence indicates that marginal BNPL adoption increases the probability that a completed purchase is subsequently reversed. The next section examines whether the associated checkout activity is retained as additional orders or transaction value.

4.4 Market-Level Gross, Returned, and Retained Demand

This section complements the order-level estimates with a market-level test of whether BNPL creates enough additional purchases to raise retained demand. The gross-to-retained accounting decomposition in Section 2.3 separates checkout activity into returned and retained components. Gross orders and gross transaction value record activity at checkout; retained orders and retained transaction value measure what remains after after-sales outcomes are resolved. The market-level analysis therefore measures how much BNPL-induced checkout activity remains after returns.

We aggregate orders to street×category×date×hour cells and apply the shift–share design

at the market level. Let m index a street×category market and t a date×hour cell. For market outcome Y_{mt}, we estimate

$$\begin{aligned} Y_{mt} &= \beta^M \, BNPL\,Share_{mt} + \bar{\mathbf{X}}'_{mt}\boldsymbol{\gamma} + \mu_m + \lambda_t + \varepsilon_{mt}, \\ BNPL\,Share_{mt} &= \pi^M Z^M_{mt} + \bar{\mathbf{X}}'_{mt}\boldsymbol{\delta} + \mu_m + \lambda_t + \nu_{mt}, \\ Z^M_{mt} &= BNPL^{\text{nat}}_{t-24} \times BNPL^{\text{base}}_{m}. \end{aligned} \tag{8}$$

Eq. (8) is the market-cell analogue of the order-level 2SLS specification. Its first line relates the market outcome to BNPL share, the second gives the first stage, and the third defines the excluded instrument. The common shift is the lagged national take-up rate used above, $BNPL^{\text{nat}}_{t-24}$, a reduced-form proxy for platform-wide BNPL availability and salience conditions. The exposure term, $BNPL^{\text{base}}_{m}$, is the initial BNPL propensity of street×category market m. The vector $\bar{\mathbf{X}}_{mt}$ contains cell-level means of the baseline controls; μ_m and λ_t denote market and date-hour FE.

[Table 5 about here.]

Table 5 first reports gross transaction value at checkout. It then groups the return outcomes—returned orders, the product-return rate, and returned transaction value—before reporting retained orders and retained transaction value. Because the coefficients correspond to a 100-percentage-point change in market BNPL share, we discuss magnitudes for a 10-percentage-point increase.

Table 5 uses the street×category markets in Eq. (8) at the date×hour frequency. A 10-percentage-point increase in BNPL share raises gross transaction value by RMB 20.63 per cell. It is also associated with 0.026 more returned orders, a 1.26-percentage-point increase in the product-return rate, and RMB 25.21 more returned transaction value. All four estimates are statistically significant.

By contrast, the same increase in BNPL share is associated with 0.008 fewer retained orders and RMB 4.59 less retained transaction value; neither estimate is statistically dis-

tinguishable from zero. Precision is informative in this aggregation: the upper endpoints of the two-sided 95% confidence intervals are 0.007 additional retained orders and RMB 1.84 of additional retained value per cell. Relative to mean retained outcomes of 0.70 orders and RMB 63.23, these bounds equal approximately 1.0% and 2.9%, ruling out a commensurate retained-demand response to the increase in returned activity.

Internet Appendix Tables IA.13 and IA.14 report estimates using coarser market definitions and alternative outcome transformations.

Table 5 decomposes cells containing observed orders. To add the extensive margin, we fix the market risk set before estimation and complete the street×category×date grid (Internet Appendix Table IA.15). The risk set contains 1,962 markets with at least 50 orders and activity on at least 10 days during a common 60-day baseline window. Of the resulting 1,006,506 market-date cells, 41.7% contain no orders. Because BNPL share is undefined in these cells, the endogenous variable is BNPL order flow. The instrument interacts predetermined market exposure and baseline order flow with a one-day-lagged national BNPL take-up shift constructed from fixed baseline province×category weights after excluding the focal province×category cell. The specification absorbs market, province×date, and category×date FE and clusters standard errors by market and date.

Panel A shows that an instrument-induced BNPL order raises returned orders by 1.60 (standard error 0.17), whereas the retained-order estimate is −0.79 (standard error 0.67). The first-stage K-P Wald F statistic is 15.99. Nominal gross and returned value rise by RMB 423.76 and RMB 339.90; the conventional retained-value estimate is RMB 83.86. Anderson–Rubin inference gives a retained-order confidence set of [−3.72, 0.10] with a p-value of 0.11. For nominal retained value, the confidence set is [RMB −10.60, RMB 142.57] and the p-value is 0.07. Panel B holds price composition fixed by valuing each order at its market's baseline average order value. Returned value rises by RMB 228.26 (standard error 28.02), while retained value changes by RMB −92.48 (standard error 83.05). The zero-inclusive estimates thus reinforce the increase in returned activity and the absence of a detectable

retained-order response. The nominal retained-value response is less definitive and sensitive to price composition.

The market-level estimates expose an incentive tension hidden by gross transaction measures. Campbell and Ramadorai (2025) emphasize that household finance concerns not only household choices but also how financial suppliers respond to consumer behavior. In our setting, returned orders and returned transaction value rise across the baseline and zero-inclusive aggregations, but retained orders do not. Evidence on retained transaction value is weaker: it is insignificant in Table 5 and becomes sensitive to price composition and weak-IV-robust inference in the zero-inclusive panel. Berg, Burg, Keil, and Puri (2025) document merchant-side effects of BNPL, and Alok, Ghosh, Kalda, and Mitra (2025) show that these effects can vary with merchant constraints and market structure. Our decomposition shows why gross merchant activity and activity retained after returns need not move one-for-one. Merchants and platforms benefit when checkout activity remains after returns, whereas credit providers can earn financing revenue when a retained purchase shifts from immediate payment to BNPL even without an increase in aggregate retained demand. Combined with the relatively high borrowing costs documented in Section 2.1, this asymmetry suggests that credit providers may be better positioned to monetize the change in payment method.

A market-level counterfactual accounting exercise rescales the merchant×product×date estimates to a 10 percentage-point increase in market BNPL share and decomposes the implied change into gross activity, returned activity, and retained demand (Internet Appendix Table IA.16). Retained orders and retained value remain imprecisely estimated in the full merchant×product×date sample. Among merchant×product markets with below-median baseline product-return rates, the retained-value estimate is also statistically indistinguishable from zero. This subgroup result shows that the absence of a detectable retained-value increase is not confined to markets with high initial return rates, although the coarser estimates remain less precise than the baseline estimates in Table 5.

5 Mental Accounting and Transaction Reversal

The market-level results establish how BNPL-financed transactions end: returned activity rises, while retained orders and retained transaction value do not increase detectably. Mental accounting provides a natural organizing framework. By separating current consumption from future payment, BNPL can reduce the weight placed on cost at checkout, admit marginal purchases, and leave some transactions vulnerable to reversal when payment or product information becomes salient. We trace this payment-decoupling mechanism through repayment cues, checkout behavior, shipment status, candidate-order bundles, and cross-sectional boundaries.

The tests locate different parts of that transaction path. Repayment-date timing provides the most targeted diagnostic of payment salience; checkout speed and return timing show when marginal transactions enter and when they are reconsidered. Candidate-order bundles and cross-sectional estimates then assess rational product-fit learning and short-run liquidity relief as benchmarks. These channels may coexist across transactions, but the joint evidence allows us to ask which account best organizes the aggregate pattern.

5.1 Repayment-Cue Salience at the Default Due Date

The platform's recurring default repayment due date provides a timing-based diagnostic of payment salience. Keys and Wang (2019) show that repayment cues and minimum-payment anchors affect credit-card debt paydown, while Kuchler and Pagel (2021) connect repayment behavior to present bias. Under consumption–payment decoupling, the due date can temporarily reconnect a new purchase to its deferred cost and discourage the most marginal orders. Neither simple liquidity relief nor product-fit learning after delivery naturally predicts the observed short-lived weakening in the BNPL–return relation for orders submitted around the due date.

We center each order on the nearest monthly due date, the 6th, and examine event time from day -15 through day $+15$. Except for the three-day reference period from day -3 to

day -1, adjacent days are pooled into two-day windows to retain a detailed path without relying on noisy day-specific estimates. The outcome is eventual product return, and the interactions compare returns on BNPL-financed orders submitted in each event-time window.

$$\begin{aligned} Return_o = \beta\, BNPL_o + \sum_{b \in \mathcal{B}} \delta_b \Big(BNPL_o \times EventWindow_o^b \Big) + \mathbf{X}_o'\boldsymbol{\gamma} + \mathbf{C}_o'\boldsymbol{\pi} \\ + \mu_{g(o),s(t(o))} + \lambda_{\ell(o)} + \tau_{h(o)} + \varepsilon_o, \end{aligned} \tag{9}$$

In Eq. (9), o indexes orders, b indexes event-time windows, and $EventWindow_o^b$ indicates that order o falls in window b. The vector $\mathbf{X}_o$ contains the baseline order controls. The product-specific time effect $\mu_{g(o),s(t(o))}$ is defined at either the product×month or product×week level; $\lambda_{\ell(o)}$ and $\tau_{h(o)}$ denote destination-street and date-hour FE. In the most stringent specification, $\mathbf{C}_o$ also includes interactions of BNPL with calendar month and day of the week. Standard errors are clustered at the product level.

The recurring due date makes time controls especially important. Event time is mechanically tied to the calendar, while promotion intensity, product demand, fulfillment conditions, and return composition can also change within products over time. Date-hour FE absorb platform-wide shocks at the time of purchase. Product×month FE then absorb product-specific monthly variation, and product×week FE tighten this comparison to orders for the same product in the same week. The final specification further allows the BNPL–return relation itself to vary across calendar months and days of the week. Estimating all three specifications on the same sample holds sample composition fixed as these time controls become progressively stricter.

Table 6 reports the two immediate post-due windows from these specifications. Figure 4 plots the central event-time paths.

[Table 6 about here.]

The two immediate coefficients are stable across the three columns. With product×month FE, the interactions for days 0 to +1 and days +2 to +3 are -0.35 and -0.38 percent-

age points, respectively. Product×week FE yield estimates of −0.34 and −0.42 percentage points. Adding the BNPL calendar controls produces estimates of −0.37 and −0.47 percentage points. The two post-due coefficients are jointly significant in the product×week specifications ($p = 0.036$ and $p = 0.012$), and each individual coefficient is significant at the 5% level or better in all three columns. The incremental BNPL effect on product returns therefore weakens during the first four days beginning with the repayment due date, even when the comparison is confined to the same product and week.

[Figure 4 about here.]

Figure 4 shows that the local decline is neither an anticipatory pattern nor a persistent day-of-month difference. Across the three specifications, the full set of estimated pre-event windows is jointly indistinguishable from the reference period, with p-values of 0.867, 0.789, and 0.570. The paths then fall in both windows covering days 0 to +3 and return close to zero by days +4 to +5; no subsequent displayed window through day +11 is significant at the 5% level. The common timing across increasingly stringent controls is consistent with a short-lived response to the repayment cue rather than a product-specific monthly or weekly cycle.

We assess whether the due-date pattern recurs around other dates by recentering a local day-specific specification on alternative days of the month (Internet Appendix Table IA.17). The average interaction for days +1 to +3 after the actual due date is −0.32 percentage points, with a standard error of 0.14. The corresponding estimates after the 3rd, 9th, 10th, 12th, 15th, 20th, and 25th range from −0.10 to 0.15 and are statistically insignificant. Among the tested dates, the temporary decline is therefore concentrated around the scheduled repayment landmark, consistent with a repayment cue briefly restoring the salience of the deferred obligation at checkout.

5.2 Checkout Speed and the Timing of Reversal

The next tests ask whether BNPL orders move through checkout more quickly and whether the additional reversals occur before or after consumers can experience the product. Table 7 reports both parts of the transaction path. Columns (1) and (2) decompose realized product returns into pre- and post-shipment reversals. Columns (3)–(6) use time to payment as a proxy for checkout deliberation and relate it to eventual returns within the same fixed-effects structure.

[Table 7 about here.]

Columns (1) and (2) show that BNPL orders are 1.45 percentage points more likely to be reversed before shipment and 1.86 percentage points more likely to be reversed after shipment, conditional on the baseline controls and fixed effects. The pre-shipment component accounts for approximately 44% of the 3.32-percentage-point baseline effect. Because these transactions are undone before delivery, learning from realized product fit or quality cannot explain this part of the response. Consumers may still acquire other information or reconsider price before shipment, so the result does not exclude every form of rational reassessment. The post-shipment estimate remains fully compatible with product-fit learning after delivery.

The positive return response on both sides of shipment is also difficult to reconcile with simple liquidity relief as the aggregate explanation. Financing an otherwise desired purchase should make that transaction easier to retain. Payment decoupling instead allows marginal purchases to enter at checkout and be reconsidered either before fulfillment or after product information arrives. The timing decomposition therefore leaves room for product-fit learning while showing that it cannot account for the full effect.

Columns (3)–(6) examine checkout speed. BNPL orders have a log time to payment that is about 0.09 lower and are 2.54 percentage points more likely to be paid within one minute. A longer time to payment is in turn associated with fewer realized product returns, consistent with slower checkout filtering out transactions that are more vulnerable to reversal. Adding

time to payment in Column (6) reduces the BNPL coefficient from 3.32 to 3.05 percentage points. Because time to payment is chosen jointly with the purchase, this attenuation is diagnostic evidence rather than a formal mediation estimate.

Taken together, faster payment and the substantial pre-shipment component point to a rapid-reconsideration margin, while the post-shipment component preserves a role for product learning. Table 3 also shows essentially no effect on late shipment, and the return estimate remains stable after accounting for payment-specific subsidies and measured traffic composition. These results make deterioration in observed merchant fulfillment and changes in measured platform context less plausible explanations for the broader transaction-path relation between BNPL and returns. We next examine the most directly observable multi-item form of rational product-fit learning.

5.3 Product-Fit Learning and Candidate-Order Bundles

Product-fit learning can occur for a single item, but a prominent multi-item form is bracketing: consumers order several alternatives, compare them after delivery, and return all but the preferred item. For bracketing to organize the aggregate return effect, BNPL should both create more simultaneous candidate bundles and shift their outcomes toward selective retention rather than complete reversal. We use orders submitted to the same destination street in the same second as a narrow proxy for such bundles; a stricter definition also requires the same merchant and at least two distinct products (Internet Appendix Table IA.18). Conditional on the baseline controls and fixed effects, BNPL orders are 0.18 percentage points less likely to enter a broad bundle and 0.11 percentage points less likely to enter a strict bundle. BNPL therefore does not create the additional simultaneous candidate sets required by the bracketing account.

The BNPL–product-return relation is larger within existing bundles, but their outcome composition does not support bracketing as an aggregate explanation. Partial returns and retaining exactly one item rise modestly only under the broad definition and are insignificant

under the strict definition. Complete reversal instead rises by 5.43 percentage points in the broad sample and 4.50 percentage points in the strict sample, while retained orders fall by 0.14 and 0.11. Because only 1.64% of orders belong to multi-order groups, weighting the incremental bundle coefficients by their sample shares accounts for about 0.04 percentage points, or roughly 1% of the baseline BNPL coefficient. The bundle evidence thus permits candidate shopping at the margin but rules out a quantitatively important shift toward selective retention.

5.4 Cross-Sectional Diagnostics and Boundary Conditions

The cross-section helps distinguish some predictions while locating the limits of others. Price and local economic conditions provide contrasts with the simple liquidity-relief benchmark. Product return risk can support either payment decoupling or product-fit learning, while livestreaming and late-evening purchases characterize environments in which marginal transactions may be easier to complete. Product categories and retail exposure reveal whether the relation is broad or confined to particular goods. We therefore treat these estimates as diagnostics and boundary conditions rather than as interchangeable tests of a single mechanism.

The conditional relation between BNPL and returns is positive across most two-digit consumption-purpose categories (Internet Appendix Table IA.19). It is 2.03 percentage points for food and non-alcoholic beverages, exceeds 3 percentage points in several groups, reaches 5.15 percentage points for restaurants and accommodation, and is 4.15 percentage points for personal care and miscellaneous. Education is the exception, with an implied estimate of -6.15 percentage points, so the relation has meaningful boundaries rather than being uniformly positive. A broad retail partition is less informative. BNPL orders are 2.76 percentage points more likely to be returned among non-retail goods and 3.34 percentage points more likely among retail goods, but the 0.58-percentage-point difference is imprecisely estimated. Consistent with Di Maggio, Williams, and Katz (2022), who show that BNPL

access shifts spending toward retail goods even among consumers without inferred liquidity constraints, the category evidence motivates looking beyond a single retail label to the settings in which marginal purchases are most likely to proceed.

Table 8 examines purchase environments and local conditions. The specifications use *Livestream*, *Late Evening*, and *Return Risk*, together with unit-price and district per capita gross domestic product (GDP) quartiles. All five specifications retain the baseline fixed effects, controls, and clustering, so the interactions are estimated within product, destination-street, and date×hour cells.

[Table 8 about here.]

Price provides the sharpest contrast with the simple liquidity-relief benchmark. Across price quartiles 1–4, the conditional BNPL–product-return differences are 2.16, 3.12, 3.97, and 3.87 percentage points. The increase is concentrated between the lowest-price group and the three higher-price groups rather than forming a strictly monotonic gradient. If liquidity relief primarily stabilized purchases, financing should be especially valuable for higher-priced purchases. Instead, all three upper quartiles have materially larger return differences, consistent with payment decoupling carrying greater consequences when the deferred monetary stake is larger, although higher prices may also accompany greater fit uncertainty. Local economic conditions provide a less direct comparison: the conditional difference is largest in the least economically developed destinations, but the upper three quartiles do not form a smooth gradient and district development is not an individual liquidity measure.

Product return risk does not distinguish payment decoupling from product-fit learning. A one-standard-deviation increase in the held-out risk measure is associated with a 0.40-percentage-point larger conditional return difference for BNPL orders, with a standard error of 0.08. The gradient is consistent both with payment decoupling admitting transactions already vulnerable to reconsideration and with rational learning being more important when product match is uncertain.

Livestreaming provides the sharper purchase-environment comparison. The conditional BNPL–product-return difference rises from 3.06 percentage points outside livestreaming traffic to 3.40 percentage points within it; the 0.34-percentage-point increment has a standard error of 0.14. Product presentation, social cues, and time-sensitive promotion occur close to payment in livestreaming transactions, making marginal purchases easier to complete, although livestreaming can also alter product information and fit uncertainty. The late-evening increment is smaller and only marginally significant, so it is a weaker diagnostic.

Following Chernozhukov, Demirer, Duflo, and Fernández-Val (2025), we use sample-split group average treatment effects (GATES) and classification analysis (CLAN) to assess the prespecified partitions jointly (Internet Appendix Table IA.20). The BNPL coefficient is 2.54 percentage points larger in the highest predicted-effect quintile than in the lowest. The high-effect group also has higher prices, more late-evening purchases, and more destinations with lower per capita GDP. These diagnostics recover the same broad economic profile using sample-split effect scores rather than one partition at a time.

The cross-section therefore identifies both gradients and boundaries. Price offers the clearest contrast with simple liquidity relief; return risk is consistent with both payment decoupling and product-fit learning; livestreaming and local economic conditions provide suggestive environmental gradients; and late-evening purchases and retail exposure offer weaker distinctions. Education shows that the BNPL–product-return relation is not uniformly positive across all goods.

Across the transaction-path tests, pure short-run liquidity relief does not organize the aggregate pattern: retained demand does not increase commensurately with returned activity, and reversals rise both before and after shipment. Product-fit learning may account for some post-shipment returns, but BNPL does not create more simultaneous candidate bundles, and complete bundle reversal dominates selective retention. The temporary due-date decline, faster checkout, and substantial pre-shipment response provide the most direct evidence consistent with payment decoupling. Together, the evidence is most consistent

with BNPL admitting marginal transactions that are reconsidered when deferred payment or product information becomes salient.

6 Conclusion

Embedded consumer credit is commonly evaluated when a loan is originated or a purchase is completed. In returnable e-commerce, this timing can overstate the demand created by PayTech because both gross merchandise value and credit originations are recorded before returns are resolved. The economically relevant distinction is therefore between transactions initiated at checkout and transactions ultimately retained. Our evidence shows that BNPL expands the former without a comparable increase in the latter.

Payment decoupling offers an economic interpretation of this wedge. By separating current consumption from future repayment, BNPL lowers the perceived cost at checkout and allows marginal transactions to proceed. The response to repayment cues, faster payment, and the substantial pre-shipment return effect connect this interpretation to the transaction path, while the limited retained-demand response shows its market consequence. PayTech can therefore change the timing and reversibility of consumption without creating an equivalent amount of retained demand. Merchants and platforms then face more returned activity, while credit providers may be better positioned to monetize the shift toward financed payment.

References

Adams, William, Liran Einav, and Jonathan Levin (2009). Liquidity constraints and imperfect information in subprime lending. *American Economic Review 99*(1), 49–84.

Agarwal, Sumit, Pulak Ghosh, Jing Li, and Tianyue Ruan (2024). Digital payments and consumption: Evidence from the 2016 demonetization in India. *Review of Financial Studies 37*(8), 2550–2585.

Aidala, Felix, Gizem Koşar, Daniel Mangrum, and Wilbert van der Klaauw (2025). Understanding consumer demand for "Buy Now, Pay Later". Staff Report 1167, Federal Reserve Bank of New York.

Alok, Shashwat, Pulak Ghosh, Ankit Kalda, and Chirasree Mitra (2025). The Buy Now Pay Later divide: Merchant heterogeneity and market structure. Working paper, Indiana University.

Altug, Mehmet Sekip, Tolga Aydinliyim, and Aditya Jain (2021). Managing opportunistic consumer returns in retail operations. *Management Science 67*(9), 5660–5678.

Berg, Tobias, Valentin Burg, Jan Keil, and Manju Puri (2025). The economics of "Buy Now, Pay Later": A merchant's perspective. *Journal of Financial Economics 171*, 104093.

Bertrand, Marianne and Adair Morse (2011). Information disclosure, cognitive biases, and payday borrowing. *Journal of Finance 66*(6), 1865–1893.

Bian, Wenlong, Lin William Cong, and Yang Ji (2024). The rise of e-wallet super-apps and Buy-Now-Pay-Later. NBER Working Paper No. 33178.

Borusyak, Kirill and Peter Hull (2023). Nonrandom exposure to exogenous shocks. *Econometrica 91*(6), 2155–2185.

Breza, Emily, Martin Kanz, and Leora Klapper (2026). Learning to navigate a new financial technology. *Journal of Finance, Early View*.

Bu, Di, Tobin Hanspal, Yin Liao, and Yong Liu (2022). Cultivating self-control in FinTech: Evidence from a field experiment on online consumer borrowing. *Journal of Financial and Quantitative Analysis 57*(6), 2208–2250.

Campbell, John Y. and Tarun Ramadorai (2025). Household finance in retrospect and prospect. Working paper, Harvard University.

Chernozhukov, Victor, Mert Demirer, Esther Duflo, and Iván Fernández-Val (2025). Fisher–Schultz lecture: Generic machine learning inference on heterogeneous treatment effects in randomized experiments, with an application to immunization in India. *Econometrica 93*(4), 1121–1164.

Chetty, Raj, Adam Looney, and Kory Kroft (2009). Salience and taxation: Theory and evidence. *American Economic Review 99*(4), 1145–1177.

Conley, Timothy G., Christian B. Hansen, and Peter E. Rossi (2012). Plausibly exogenous. *Review of Economics and Statistics 94*(1), 260–272.

Cornelli, Giulio, Leonardo Gambacorta, and Livia Pancotto (2023, December). Buy now, pay later: A cross-country analysis. *BIS Quarterly Review*, 61–75.

Danieli, Oren, Daniel Nevo, Itai Walk, Bar Weinstein, and Dan Zeltzer (2026). Negative control falsification tests for instrumental variable designs. *American Economic Review 116*(4), 1380–1414.

Degryse, Hans, Tarik Roukny, and Joris Tielens (2026). Asset overhang and technological change. *Review of Financial Studies 39*(7), 2115–2178.

deHaan, Ed, Jungbae Kim, Ben Lourie, and Chenqi Zhu (2024). Buy now pay (pain?) later. *Management Science 70*(8), 5586–5598.

Di Maggio, Marco, Angela Ma, and Emily Williams (2025). In the red: Overdrafts, payday lending, and the underbanked. *Journal of Finance 80*(3), 1691–1738.

Di Maggio, Marco, Emily Williams, and Justin Katz (2022). Buy Now, Pay Later credit: User characteristics and effects on spending patterns. NBER Working Paper No. 30508.

Dong, Yanfei, Jiayin Hu, Yiping Huang, Han Qiu, and Yingguang Zhang (2025). The use and disuse of FinTech credit: When Buy Now, Pay Later meets credit reporting. BIS Working Paper No. 1239.

Einav, Liran, Mark Jenkins, and Jonathan Levin (2012). Contract pricing in consumer credit markets. *Econometrica 80*(4), 1387–1432.

Finkelstein, Amy (2009). E-ZTAX: Tax salience and tax rates. *Quarterly Journal of Economics 124*(3), 969–1010.

Gabaix, Xavier and David Laibson (2006). Shrouded attributes, consumer myopia, and information suppression in competitive markets. *Quarterly Journal of Economics 121*(2), 505–540.

Gathergood, John, Benedict Guttman-Kenney, and Stefan Hunt (2019). How do payday loans affect borrowers? Evidence from the U.K. market. *Review of Financial Studies 32*(2), 496–523.

Gelman, Michael and Nikolai Roussanov (2024). Managing mental accounts: Payment cards and consumption expenditures. *Review of Financial Studies 37*(8), 2586–2624.

Goldsmith-Pinkham, Paul, Isaac Sorkin, and Henry Swift (2020). Bartik instruments: What, when, why, and how. *American Economic Review 110*(8), 2586–2624.

Gross, David B. and Nicholas S. Souleles (2002). Do liquidity constraints and interest rates matter for consumer behavior? Evidence from credit card data. *Quarterly Journal of Economics 117*(1), 149–185.

Hall, Robert E. (1978). Stochastic implications of the life cycle-permanent income hypothesis: Theory and evidence. *Journal of Political Economy 86*(6), 971–987.

Han, Pengfei and Zhu Wang (2026). Technology adoption and leapfrogging: Racing for mobile payments. *Management Science*. Articles in Advance.

Heidhues, Paul and Botond Kőszegi (2010). Exploiting naivete about self-control in the credit market. *American Economic Review 100*(5), 2279–2303.

Jansen, Mark, Fabian Nagel, Constantine Yannelis, and Anthony Lee Zhang (2025). Data and welfare in credit markets. *Journal of Financial Economics 174*, 104171.

Karlan, Dean and Jonathan Zinman (2009). Observing unobservables: Identifying information asymmetries with a consumer credit field experiment. *Econometrica 77*(6), 1993–2008.

Keil, Jan and Valentin Burg (2023). "Buy Now, Pay Later" and impulse shopping. Working paper, Indian School of Business.

Keys, Benjamin J. and Jialan Wang (2019). Minimum payments and debt paydown in consumer credit cards. *Journal of Financial Economics 131*(3), 528–548.

Kuchler, Theresa and Michaela Pagel (2021). Sticking to your plan: The role of present bias for credit card paydown. *Journal of Financial Economics 139*(2), 359–388.

Laibson, David (1997). Golden eggs and hyperbolic discounting. *Quarterly Journal of*

Economics 112(2), 443–477.

Laudenbach, Christine, Elin Molin, Kasper F. Roszbach, and Talina Sondershaus (2025). Buy Now Pay (Less) Later: Leveraging private BNPL data in consumer banking. Working paper, Goethe University.

Melzer, Brian T. (2011). The real costs of credit access: Evidence from the payday lending market. *Quarterly Journal of Economics 126*(1), 517–555.

Morse, Adair (2011). Payday lenders: Heroes or villains? *Journal of Financial Economics 102*(1), 28–44.

Nageswaran, Leela, Elina H. Hwang, and Soo-Haeng Cho (2025). Offline returns for online retailers via partnership. *Management Science 71*(1), 279–297.

Norden, Lars and Martin Weber (2010). Credit line usage, checking account activity, and default risk of bank borrowers. *Review of Financial Studies 23*(10), 3665–3699.

O'Donoghue, Ted and Matthew Rabin (1999). Doing it now or later. *American Economic Review 89*(1), 103–124.

Pagano, Marco and Tullio Jappelli (1993). Information sharing in credit markets. *Journal of Finance 48*(5), 1693–1718.

Petersen, J. Andrew and V. Kumar (2009). Are product returns a necessary evil? Antecedents and consequences. *Journal of Marketing 73*(3), 35–51.

Ponce, Alejandro, Enrique Seira, and Guillermo Zamarripa (2017). Borrowing on the wrong credit card? Evidence from Mexico. *American Economic Review 107*(4), 1335–1361.

Prelec, Drazen and George Loewenstein (1998). The red and the black: Mental accounting of savings and debt. *Marketing Science 17*(1), 4–28.

Shema-Zlatokrilov, Lilah (2026). Pay now, regret later? When disclosure fails to deter: Choice persistence and post-purchase satisfaction in installment payments. Working paper, Hebrew University of Jerusalem.

Soman, Dilip (2001). Effects of payment mechanism on spending behavior: The role of rehearsal and immediacy of payments. *Journal of Consumer Research 27*(4), 460–474.

Stango, Victor and Jonathan Zinman (2014). Limited and varying consumer attention: Evidence from shocks to the salience of bank overdraft fees. *Review of Financial Studies 27*(4), 990–1030.

Stiglitz, Joseph E. and Andrew Weiss (1981). Credit rationing in markets with imperfect information. *American Economic Review 71*(3), 393–410.

Thaler, Richard (1985). Mental accounting and consumer choice. *Marketing Science 4*(3), 199–214.

Appendix A. Variable Definitions

This appendix defines the principal variables in the main analysis. Measures used only in the Internet Appendix are defined where they are introduced.

Variable	Definition
Panel A. **Order-level measures**	
Return	Indicator equal to one when the platform records a finalized product return or refund and zero for finalized kept orders or resolved cases without reimbursement. Unresolved after-sales cases are excluded. The indicator is multiplied by 100 and expressed in percentage points in the regressions and tables.
BNPL	Indicator equal to one when the order is paid with TikTok BNPL, officially branded in China as Douyin Monthly Payment. It is measured on a zero-to-one scale unless a table explicitly reports $BNPL(\%) = 100 \times BNPL$.
Log Price	Natural logarithm of unit price.
Log Quantity	Natural logarithm of order quantity.
Log Shipping Fee	Natural logarithm of the shipping fee after adding one.
Log Discount	Natural logarithm of the total discount after adding one.
Log Shipping Lag	Natural logarithm of one plus the hours from order submission to promised shipment. Implausible values are removed, and the hours are winsorized at the 1st and 99th percentiles before the transformation.
Local Shipping	Indicator equal to one when shipping origin and destination are in the same province.
Payment Discount Amount	Payment-method-specific discount amount, in RMB.
Explicit Pay Discount	Indicator equal to one when *Payment Discount Amount* is strictly positive.
Late Shipment	Indicator equal to one when actual shipment occurs after the promised shipment time. The variable is expressed in percentage points in the regressions.
Panel B. **Instrument and market-level measures**	
$BNPL_{pg}^{\text{base}}$	Initial BNPL exposure, measured as mean take-up during the first seven days after province–product pair (p, g) enters the sample and expressed as a zero-to-one fraction.
$BNPL_{t-L}^{\text{nat}}$	National BNPL take-up in the lagged time cell, measured as a zero-to-one fraction. The lag is 24 hours in hourly specifications and one day in daily specifications.
Z_{pgt}^{SS}	Shift–share instrument equal to the product of lagged national BNPL take-up and initial province–product BNPL exposure.
$BNPL\,Share_{mt}$	Fraction of completed orders in market cell (m, t) paid with BNPL, measured from zero to one.
Market order outcomes	$Orders_{mt}$ is the number of completed orders. *Returned orders*$_{mt}$ counts completed orders ending in a realized return; *Retained orders*$_{mt}$ is the difference between the two. *Return rate*$_{mt}$ is the percentage of completed orders ending in a realized return and is measured in percentage points.
Market value outcomes	*Gross value*$_{mt}$ is the sum of order payable amounts. *Returned value*$_{mt}$ sums payable amounts for returned orders; *Retained value*$_{mt}$ is the difference between gross and returned value.
Z_{mt}^{M}	Market-level shift–share instrument, equal to initial market BNPL exposure multiplied by lagged national BNPL take-up.

Appendix A. Continued

Variable	Definition
Panel C. **Mechanism and heterogeneity measures**	
EventWindow^{b}	Indicator equal to one when the order is submitted in event-time window b relative to the nearest recurring default repayment date on the sixth. Figure 4 displays two-day windows from day −11 through day +11, except that days −3 to −1 form the three-day reference period.
Pre-Shipment Return, *Post-Shipment Return*	Return indicators, in percentage points, for orders without and with an observed shipment timestamp, respectively.
Time to Payment	Minutes from order submission to payment completion, restricted to values between zero and 1,440 and winsorized at the 1st and 99th percentiles.
Log Time to Payment	Natural logarithm of one plus *Time to Payment*.
Fast Payment	Indicator equal to one when payment is completed within one minute of order submission. The indicator is multiplied by 100 in the reported regression.
Retail	Indicator equal to one for the prespecified consumer-facing two-digit retail categories.
Livestream	Indicator equal to one when the traffic-format field identifies livestreaming.
Late Evening	Indicator equal to one for orders submitted from 21:00 through 23:59.
Return Risk	Standardized held-out product return-risk measure constructed from a random training half of the sample; sparse products use category-level risk.
Price quartiles	Indicators for unit-price quartiles Q2–Q4; Q1 is omitted.
Local development quartiles	Indicators for quartiles Q2–Q4 of destination-district economic development; Q1 is omitted.
Product gender orientation	Male- and female-oriented products are classified from explicit gender signals in product titles and specifications. Products with mixed or ambiguous signals are excluded from the split-sample analysis.

Figure 1: **Institutional and transaction flows of TikTok BNPL**

The diagram traces checkout, funding, repayment, and return flows among the buyer, platform, payment provider, lender, and merchant. Dashed arrows denote information and authorization flows; solid arrows denote funding and settlement flows. The representation is schematic.

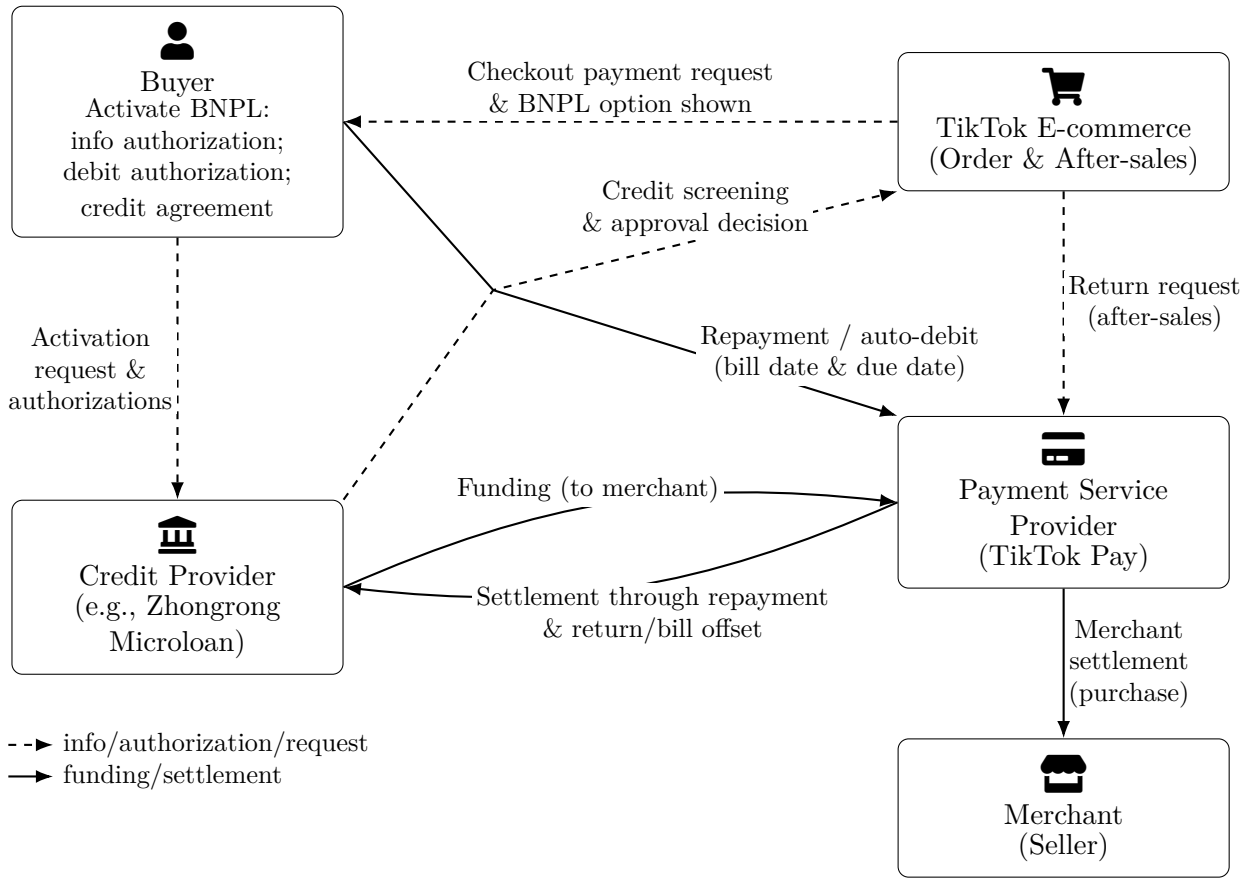


Figure 2: **Two-stage conceptual framework for BNPL and retained demand**

The figure contrasts three transaction paths. Under payment decoupling, BNPL weakens consumption–payment coupling at checkout, admits marginal purchases, and makes later reversal more likely when product information or repayment cues become salient. Under short-run liquidity relief, BNPL finances an otherwise desired purchase that is more likely to be retained. Under rational product-fit learning, BNPL lowers the immediate cost of trial; its multi-item form, bracketing, predicts post-delivery comparison and selective retention of the preferred item.

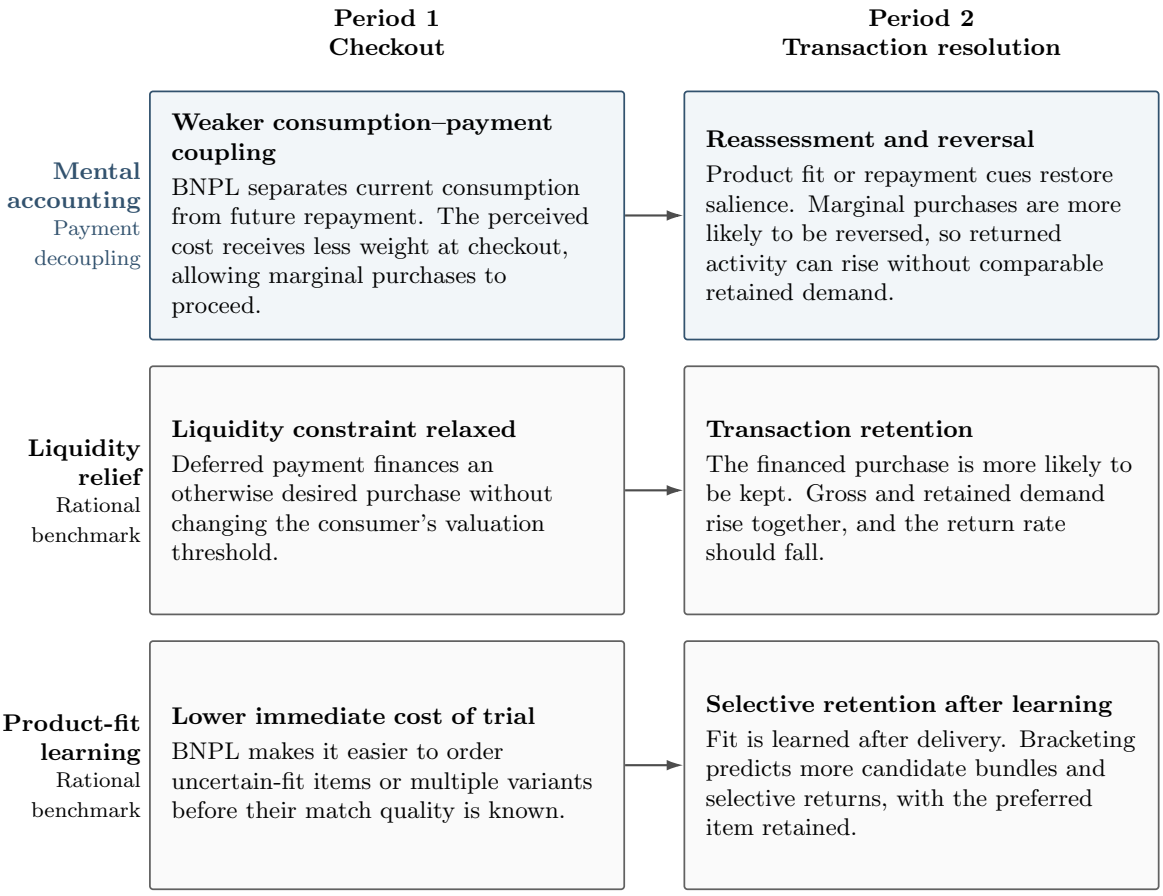


Figure 3: **TikTok checkout screen with BNPL**

A representative TikTok E-commerce checkout screen displays the platform's officially branded Douyin Monthly Payment option alongside alternative payment methods. The BNPL option offers either full repayment in the next billing cycle or installment plans, and the screen displays BNPL-specific discounts. Amounts are in Chinese renminbi (RMB). The displayed methods are front-end checkout choices; the order data do not reveal the funding source selected inside an external wallet. The screenshot is from the authors. The original interface is in Chinese; all on-screen text shown in the figure has been translated into English by the authors.

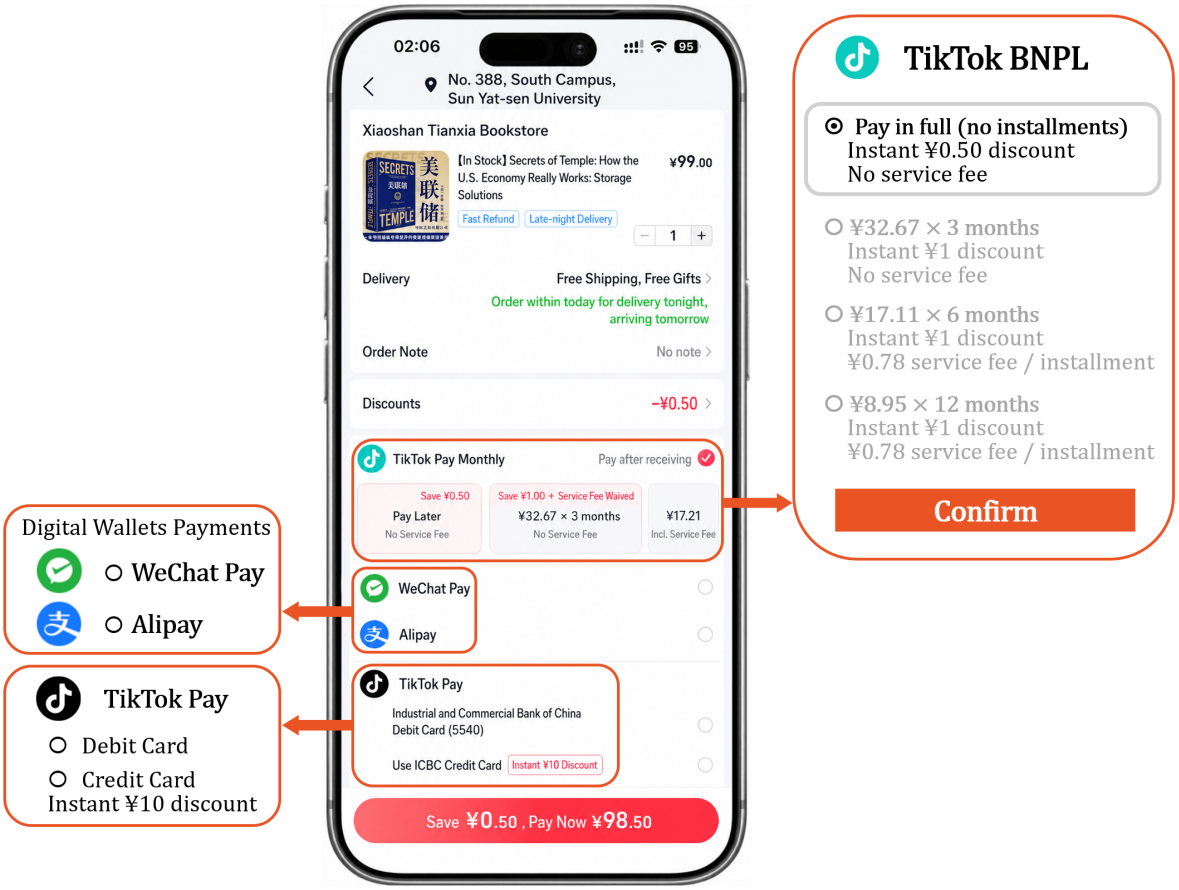


Figure 4: **BNPL and product returns around the default repayment date**

The lines plot coefficients on interactions between *BNPL* and event-time windows around the platform's recurring default repayment date, the 6th of each month; vertical bars give 95% confidence intervals. Orders are centered on the nearest due date. The figure displays two-day windows from day −11 through day +11, except that days −3 to −1 form the reference period and are normalized to zero. All specifications begin with the same 12,890,784 valid input orders and use the baseline controls, destination-street FE, date-hour FE, and product-clustered standard errors. The estimates use the common support that remains after singleton removal under the product×month and product×week FE structures. The first line absorbs product×month FE, the second absorbs product×week FE, and the third adds interactions of BNPL with calendar month and day of the week to the product×week specification.

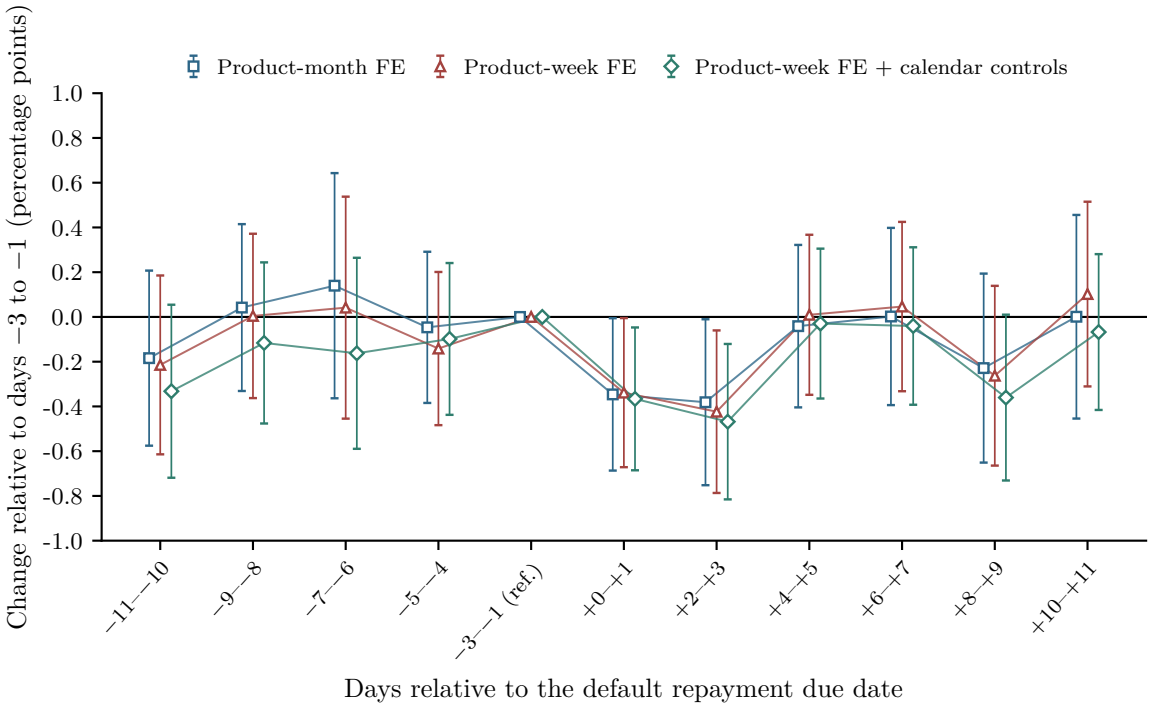


Table 1: **Sample statistics and differences by BNPL status**

This table reports full-sample descriptive statistics and differences by BNPL status. Panel A summarizes finalized orders with non-missing key variables; SD denotes standard deviation, and Min and Max report the sample extrema. Panel B reports group means and their differences. Difference is the BNPL mean minus the non-BNPL mean; significance is based on an unequal-variance two-sample test. Unit price, shipping fee, and discount are in Chinese renminbi; quantity and shipping lag are in units and hours, respectively. The table reports all transaction characteristics in their original units, whereas the regressions use the logarithmic transformations defined in Appendix A. Return is expressed in percentage points in both panels. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

Panel A. **Full-sample descriptive statistics (Obs. = 12,890,784)**

Variable	Mean	SD	Min	Median	Max
Return	41.61	49.29	0.00	0.00	100.00
BNPL	0.26	0.44	0.00	0.00	1.00
Unit Price (RMB)	150.39	410.07	0.01	69.90	100,000.00
Quantity	1.03	0.50	1.00	1.00	200.00
Shipping Fee (RMB)	0.11	1.83	0.00	0.00	820.00
Discount (RMB)	17.66	80.51	0.00	0.00	99,200.00
Shipping Lag (hours)	110.78	127.78	25.51	48.00	600.00
Local Shipping	0.09	0.29	0.00	0.00	1.00

Panel B. **Differences by BNPL status**

Variable	Non-BNPL (Obs. = 9,511,799)	BNPL (Obs. = 3,378,985)	Difference
Return	38.97	49.04	10.07***
Unit Price (RMB)	135.06	193.54	58.47***
Quantity	1.03	1.03	-0.01***
Shipping Fee (RMB)	0.10	0.12	0.01***
Discount (RMB)	16.26	21.59	5.33***
Shipping Lag (hours)	101.02	138.26	37.24***
Local Shipping	0.09	0.10	< 0.01***

Table 2: **BNPL and product returns**

Order-level regressions relate *Return* to *BNPL*. Controls are *Log Price*, *Log Quantity*, *Log Shipping Fee*, *Log Discount*, *Log Shipping Lag*, and *Local Shipping*. Column (1) gives the unconditional comparison. Columns (2) and (3) report robust standard errors. Standard errors are clustered by category, shop, and product in Columns (4)–(6), respectively. For specifications with fixed effects, observations report valid input orders before the estimator removes singleton groups. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

	Dependent variable = *Return*					
	(1)	(2)	(3)	(4)	(5)	(6)
BNPL	10.07***	3.56***	3.50***	3.54***	3.53***	3.32***
	(0.03)	(0.03)	(0.03)	(0.67)	(0.19)	(0.11)
Log Price		10.70***	10.47***	8.62***	8.48***	4.43***
		(0.01)	(0.01)	(1.52)	(0.89)	(0.60)
Log Quantity		-0.21**	0.13	3.73**	4.38***	3.42***
		(0.11)	(0.10)	(1.38)	(0.81)	(0.39)
Log Shipping Fee		-0.60***	0.38***	1.42	0.25	0.69**
		(0.06)	(0.07)	(1.27)	(0.65)	(0.31)
Log Discount		-2.21***	-2.02***	-2.39**	-0.65***	-0.64***
		(0.01)	(0.01)	(0.86)	(0.22)	(0.12)
Log Shipping Lag		12.77***	13.54***	8.61***	5.81***	3.69***
		(0.02)	(0.02)	(1.04)	(0.75)	(0.45)
Local Shipping		-3.11***	-1.03***	-1.25	-0.97*	-0.38***
		(0.04)	(0.05)	(0.73)	(0.51)	(0.11)
Obs.	12,890,784	12,890,784	12,890,784	12,890,784	12,890,784	12,890,784
R^2	0.01	0.17	0.21	0.24	0.28	0.31
Category FE	No	No	No	Yes	No	No
Shop FE	No	No	No	No	Yes	No
Product FE	No	No	No	No	No	Yes
Street FE	No	No	Yes	Yes	Yes	Yes
Date-hour FE	No	No	Yes	Yes	Yes	Yes

Table 3: **Alternative samples and specifications for BNPL and product returns**

This table reports alternative specifications and sample checks for the relation between BNPL use and product returns. In Panel A, Column (1) removes orders with payment-specific discounts, Column (2) adds traffic FE, and Columns (3)–(4) absorb product×date and shop×date FE. In Panel B, Column (5) absorbs category×date×hour FE, Columns (6)–(7) restrict the sample, and Column (8) uses late shipment as a fulfillment placebo. Column (9) restricts the sample to products classified from explicit male- or female-oriented signals, excludes mixed or ambiguous products, and interacts BNPL with the female-oriented indicator. The Column (8) outcome is multiplied by 100. Controls are as in Table 2. Observations report valid input orders before singleton removal. Standard errors are clustered at the product level. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

Panel A. **Alternative specifications**

	No pay discount	Traffic controls	Product date	Shop date
	(1)	(2)	(3)	(4)
BNPL	3.46***	3.21***	3.22***	3.27***
	(0.11)	(0.11)	(0.11)	(0.11)
Obs.	12,617,430	12,890,784	12,890,784	12,890,784
R^2	0.31	0.31	0.34	0.32
Controls	Yes	Yes	Yes	Yes
Product FE	Yes	Yes	No	Yes
Street FE	Yes	Yes	Yes	Yes
Date-hour FE	Yes	Yes	Yes	Yes
Traffic FE	No	Yes	No	No
Product×date FE	No	No	Yes	No
Shop×date FE	No	No	No	Yes

Panel B. **Additional specifications and sample checks**

	Category date×hour	Payment overlap	Older cohorts	Late shipment	Classified products
	(5)	(6)	(7)	(8)	(9)
BNPL	3.30***	4.22***	3.35***	-0.01	3.94***
	(0.11)	(0.25)	(0.12)	(0.01)	(0.32)
BNPL × *Female-Oriented*					-0.07
					(0.40)
Obs.	12,890,784	432,368	12,177,188	9,372,817	3,273,604
R^2	0.32	0.32	0.31	0.28	0.27
Controls	Yes	Yes	Yes	Yes	Yes
Product FE	Yes	Yes	Yes	Yes	Yes
Street FE	Yes	Yes	Yes	Yes	Yes
Date-hour FE	No	Yes	Yes	Yes	Yes
Category×date×hour FE	Yes	No	No	No	No

Table 4: **Shift–share IV estimates of BNPL and product returns**

Shift–share instrumental-variable (IV) estimates relate BNPL use to realized product returns. Column (1) gives the fixed-effects estimate for all 12,890,784 valid orders; Column (2) gives the reduced form; and Column (3) gives the baseline two-stage least squares (2SLS) estimate. The 24-hour-lagged national shift is unavailable during the opening 24-hour window, so Columns (2)–(3) exclude 2,134 orders. Columns (4) and (5) use fixed category and province×category weights, respectively, and begin after the predetermined weighting window; Column (6) requires at least five orders in the initial province–product exposure window. The baseline instrument is $Z^{\mathrm{SS}}_{pgt} = BNPL^{\mathrm{nat}}_{t-24} \times BNPL^{\mathrm{base}}_{pg}$. Controls are as in Table 2. Standard errors are clustered at the product level. Observations report valid input orders before the fixed-effect estimator removes singleton groups. *First-stage* gives the excluded-instrument coefficient; *K-P Wald F* denotes the Kleibergen–Paap (K-P) rk Wald F statistic. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

	Dependent variable = *Return*					
	(1) Fixed effects	(2) Reduced form	(3) 2SLS	(4) Category LOO	(5) Double LOO	(6) Thick cells ($n_{\mathrm{base}} \geq 5$)
BNPL	3.32***		3.71***	4.00***	4.08***	2.82***
	(0.11)		(0.40)	(0.43)	(0.42)	(0.68)
Z^{SS}_{pgt}		4.30***				
		(0.47)				
First-stage			1.16***	1.08***	1.14***	1.27***
			(0.03)	(0.03)	(0.04)	(0.03)
K-P Wald F			1458.83	1055.02	969.79	1494.52
Obs.	12,890,784	12,888,650	12,888,650	12,713,259	12,713,259	11,071,931
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Product FE	Yes	Yes	Yes	Yes	Yes	Yes
Street FE	Yes	Yes	Yes	Yes	Yes	Yes
Date-hour FE	Yes	Yes	Yes	Yes	Yes	Yes

Table 5: **BNPL and market-level gross, returned, and retained activity**

This table reports instrumental-variable estimation results for the effects of BNPL use on market-level gross, returned, and retained activity. Markets are defined at the street×category×date×hour level, and the common shift is lagged by 24 hours. Count outcomes are measured in orders, the return rate in percentage points, and value outcomes in RMB. Market BNPL share is instrumented with initial market exposure interacted with lagged national BNPL take-up; coefficients correspond to a one-unit increase in this zero-to-one share. Controls are cell-level means of the order-level controls in Table 2. Observations report valid input cells before singleton removal, and standard errors are clustered at the market level. *First-stage* gives the excluded-instrument coefficient; *K-P Wald F* is the Kleibergen–Paap rk Wald F statistic. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

	Dependent variable					
	(1) *Gross value*	(2) *Returned orders*	(3) *Return rate*	(4) *Returned value*	(5) *Retained orders*	(6) *Retained value*
BNPL Share	206.26**	0.26**	12.57**	252.11**	-0.08	-45.85
	(102.07)	(0.12)	(5.55)	(100.58)	(0.08)	(32.79)
First-stage	0.18***	0.18***	0.17***	0.18***	0.18***	0.18***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
K-P Wald F	363.94	363.94	220.82	363.94	363.94	363.94
Obs.	10,738,785	10,738,785	10,738,785	10,738,785	10,738,785	10,738,785
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Street×category FE	Yes	Yes	Yes	Yes	Yes	Yes
Date-hour FE	Yes	Yes	Yes	Yes	Yes	Yes

Table 6: **Billing-cycle salience around the default repayment date**

This table reports selected event-time coefficients plotted in Figure 4. Orders are centered on the nearest recurring repayment due date, the 6th of each month; days -3 to -1 form the reference period. Each column begins with 12,890,784 valid input orders and uses the common support remaining after singleton removal under the product×month and product×week FE structures. The dependent variable is multiplied by 100, so coefficients are in percentage points. Specifications include the baseline controls, street FE, and date-hour FE; Column (3) also allows the BNPL coefficient to vary by calendar month and day of the week. Observations report valid input orders before singleton removal, and standard errors are clustered at the product level. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

	Dependent variable = *Return*		
	(1)	(2)	(3)
BNPL × Days 0 to +1	-0.35**	-0.34**	-0.37**
	(0.17)	(0.17)	(0.16)
BNPL × Days +2 to +3	-0.38**	-0.42**	-0.47***
	(0.19)	(0.19)	(0.18)
Joint p-value, days 0 to +3	0.06	0.04	0.01
Pre-event joint p-value	0.87	0.79	0.57
Obs.	12,890,784	12,890,784	12,890,784
R^2	0.31	0.32	0.32
Controls	Yes	Yes	Yes
Product×month FE	Yes	No	No
Product×week FE	No	Yes	Yes
Street FE	Yes	Yes	Yes
Date-hour FE	Yes	Yes	Yes
BNPL calendar controls	No	No	Yes

Table 7: **Product return timing and checkout deliberation**

Columns (1) and (2) split realized product returns into pre- and post-shipment components, where *Shipped* = 1 if a shipment timestamp is observed. Columns (3)–(6) examine checkout deliberation. *Time to Payment* is the number of minutes between order submission and payment completion, and *Log Time to Payment* is the natural logarithm of one plus this value. The underlying duration is trimmed to $[0, 1440]$ and winsorized at the 1st and 99th percentiles. *Fast Payment* equals one when payment occurs within one minute. The pre-shipment return, post-shipment return, and fast-payment indicators are multiplied by 100, so their coefficients are in percentage points. Column (5) restates the baseline return specification; Column (6) adds time to payment in minutes and uses observations for which this duration is observed. Controls are as in Table 2. Observations report valid input orders before the fixed-effect estimator removes singleton groups. Standard errors are clustered at the product level. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

	Pre-shipment returns	*Post-shipment returns*	*Log time to payment*	*Pay within 1 minute*	*Return*	*Return*
	(1)	(2)	(3)	(4)	(5)	(6)
BNPL	1.45***	1.86***	-0.09***	2.54***	3.32***	3.05***
	(0.07)	(0.11)	(0.00)	(0.05)	(0.11)	(0.11)
Time to Payment						-2.21***
						(0.16)
Obs.	12,890,784	12,890,784	12,890,725	12,890,725	12,890,784	12,890,725
R^2	0.24	0.14	0.13	0.06	0.31	0.31
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Product FE	Yes	Yes	Yes	Yes	Yes	Yes
Street FE	Yes	Yes	Yes	Yes	Yes	Yes
Date-hour FE	Yes	Yes	Yes	Yes	Yes	Yes

Table 8: **BNPL and product returns across purchase environments, product risk, and local conditions**

Heterogeneity in the BNPL–product-return relation is estimated across purchase environments, product risk, price, and local economic conditions. *Livestream* = 1 if the order's traffic-format field identifies livestreaming, and *Late Evening* = 1 for orders submitted between 21:00 and 23:59. *Return Risk* is a standardized held-out product return-risk measure, constructed from a random training half of the sample and applied to the evaluation half; sparse products use category-level risk. Price quartiles are based on unit price, with Q1 omitted. District per capita gross domestic product quartiles are based on destination-district gross domestic product, with Q1 omitted; observations with missing district gross domestic product are excluded. Controls are as in Table 2. Observations report valid input orders before the fixed-effect estimator removes singleton groups. Standard errors are clustered at the product level. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

	Dependent variable = *Return*				
	Livestream	*Late evening*	*Return risk*	*Price quartiles*	*District GDP quartiles*
	(1)	(2)	(3)	(4)	(5)
BNPL (base group)	3.06***	3.26***	3.28***	2.16***	3.62***
	(0.12)	(0.13)	(0.11)	(0.15)	(0.15)
BNPL × *Livestream*	0.34**				
	(0.14)				
BNPL × *Late Evening*		0.21*			
		(0.11)			
BNPL × *Return Risk*			0.40***		
			(0.08)		
BNPL × *Price Q2*				0.96***	
				(0.21)	
BNPL × *Price Q3*				1.81***	
				(0.29)	
BNPL × *Price Q4*				1.71***	
				(0.25)	
BNPL × *GDP Q2*					-0.36***
					(0.09)
BNPL × *GDP Q3*					-0.43***
					(0.10)
BNPL × *GDP Q4*					-0.40***
					(0.10)
Obs.	12,890,784	12,890,784	6,446,900	12,890,784	12,659,183
R^2	0.31	0.31	0.31	0.31	0.31
Controls	Yes	Yes	Yes	Yes	Yes
Product FE	Yes	Yes	Yes	Yes	Yes
Street FE	Yes	Yes	Yes	Yes	Yes
Date-hour FE	Yes	Yes	Yes	Yes	Yes

Internet Appendix for

Order Now at Will, Return Later with Remorse: The Dark Side of PayTech

Figure IA.1: **Distribution of order volume across China**

The figure reports destination-city order volume in the final estimation sample. Cities are sorted into deciles of order volume, with darker shading indicating a higher position in the distribution.

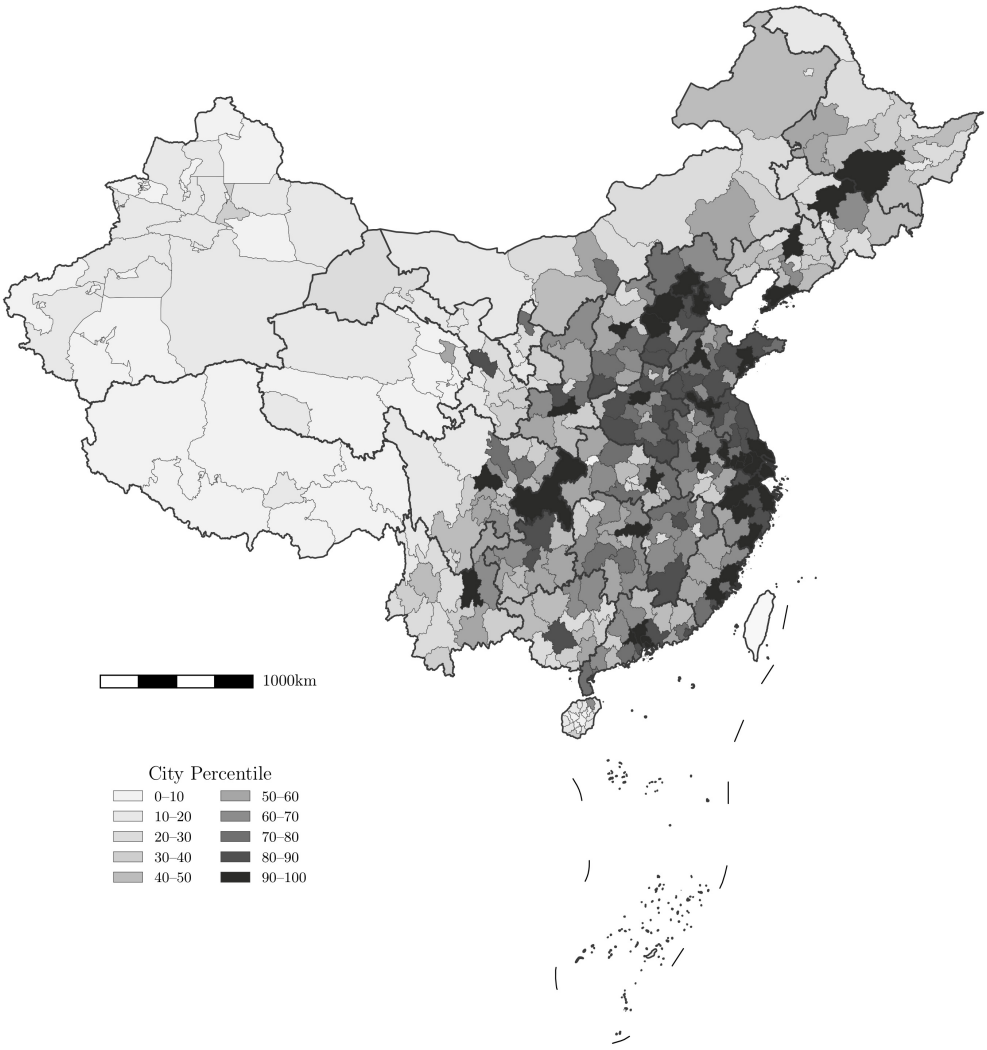
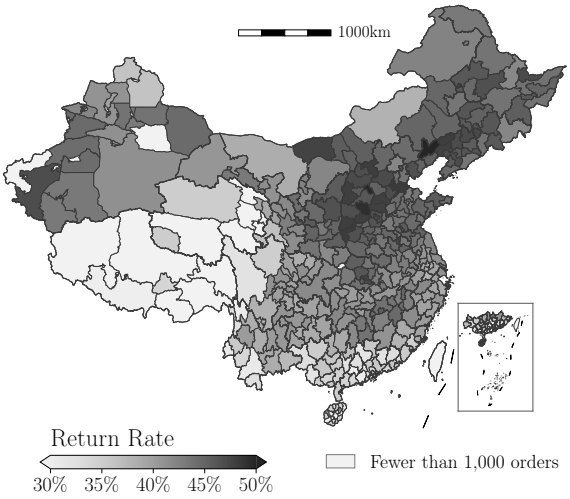


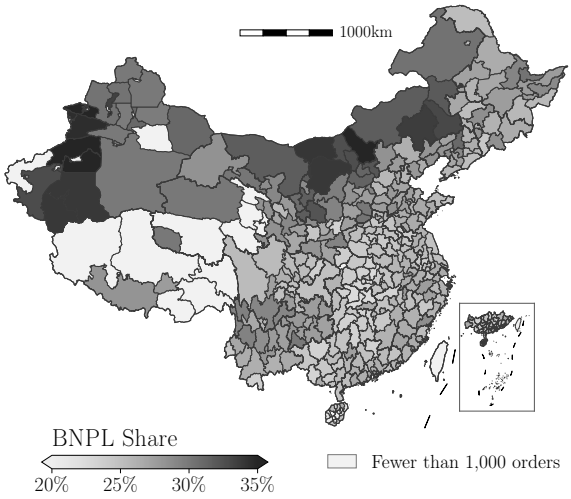
Figure IA.2: **Distribution of product returns and BNPL take-up**

This figure reports city-level product returns and BNPL take-up. Panels A and B map product-return rates and BNPL shares; cities with fewer than 1,000 orders are shown in light gray. Panel C plots the two rates for cities with at least 10,000 orders and includes an unweighted ordinary least squares fit. The Pearson and Spearman correlations are 0.260 and 0.227, respectively ($N = 265$).

Panel A. **Product Return Rate**



Panel B. **BNPL Share**



Panel C. **BNPL Share and Product Return Rate**

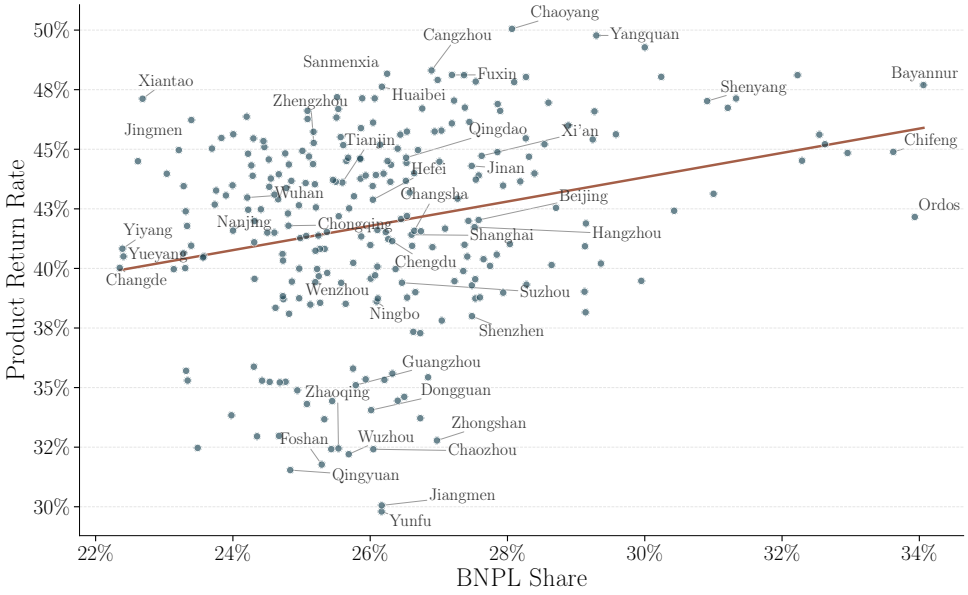


Figure IA.3: **Product return rates by product category and payment method**

Unconditional product-return rates for BNPL and non-BNPL orders are shown by two-digit consumption-purpose category. Categories are ordered from the highest to the lowest BNPL product-return rate; the unclassified category is omitted.

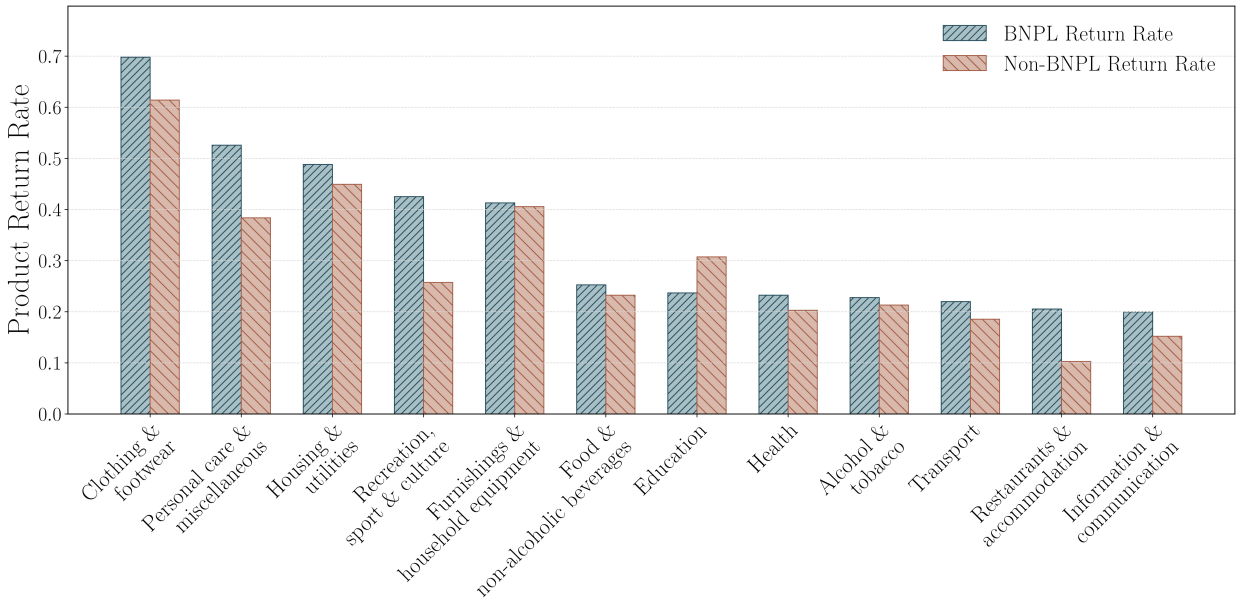


Figure IA.4: **Monthly product return rates by payment method**

Unconditional product-return rates for BNPL and non-BNPL orders are shown by order-submission month.

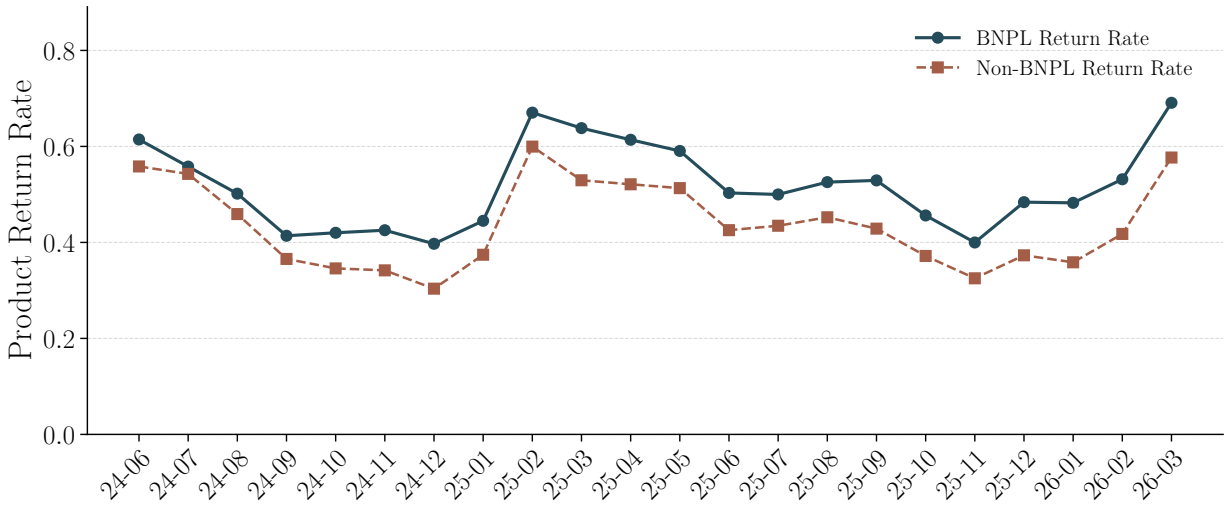


Figure IA.5: **Persistence of province–product BNPL exposure**

Province–product pairs are sorted into deciles by their initial seven-day BNPL share, restricting the sample to pairs with at least ten baseline-window orders. The dashed line shows the mean baseline BNPL share, and the solid line shows the mean post-baseline BNPL share. The Pearson and Spearman correlations between the initial and post-baseline shares are 0.405 and 0.463, respectively. Table 4 uses a looser cutoff of $n_{\text{base}} \geq 5$.

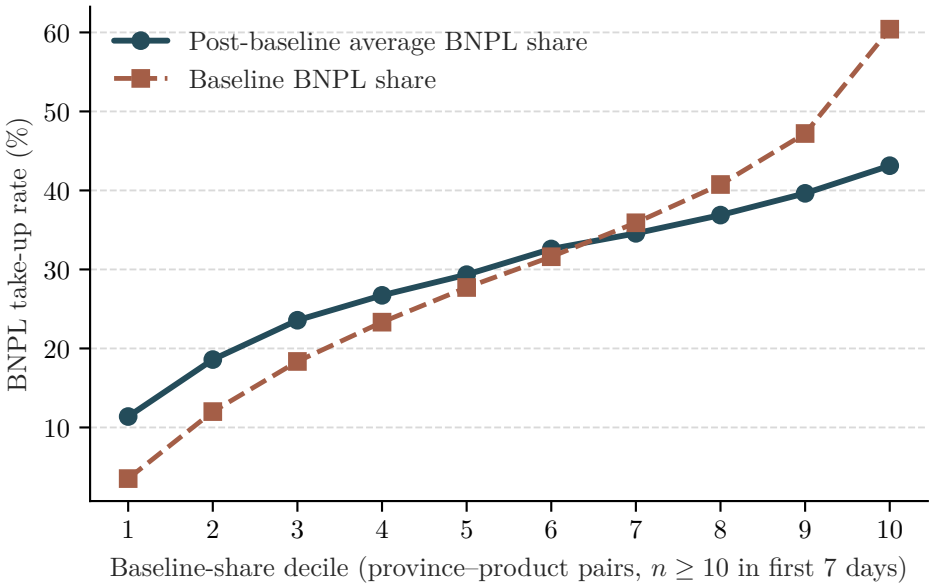


Table IA.1: **Public promotions and BNPL take-up**

This table reports order-level BNPL take-up around two public TikTok promotional openings in 2025 that featured installment support through the product officially branded as Douyin Monthly Payment. Panel A uses the 2025 618 campaign opening and reports 7-, 14-, 21-, and 30-day windows in Columns (1)–(4). Panel B uses the 2025 Double 11 campaign opening and reports the same windows in Columns (5)–(8). Each column gives a separate regression of BNPL take-up on the interaction between a post-promotion indicator and predetermined province–product BNPL exposure. The sample is restricted to province–product cells with at least 50 baseline-window orders. Coefficients and standard errors are in percentage points. Controls are as in Table 2. Observations report valid input orders before the fixed-effect estimator removes singleton groups, and standard errors are clustered at the product level. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

Panel A. **2025 618 campaign**

	Dependent variable = *BNPL*(%)			
	(1) 7-day	(2) 14-day	(3) 21-day	(4) 30-day
Post-promotion	12.00**	7.52*	6.46	6.90*
× *baseline exposure*	(4.95)	(4.24)	(4.07)	(4.03)
Obs.	124,369	285,180	472,122	655,151
R^2	0.18	0.12	0.10	0.09
Controls	Yes	Yes	Yes	Yes
Product FE	Yes	Yes	Yes	Yes
Street FE	Yes	Yes	Yes	Yes
Date-hour FE	Yes	Yes	Yes	Yes

Panel B. **2025 Double 11 campaign**

	Dependent variable = *BNPL*(%)			
	(5) 7-day	(6) 14-day	(7) 21-day	(8) 30-day
Post-promotion	22.32***	19.41***	17.31***	15.64***
× *baseline exposure*	(5.83)	(5.09)	(4.42)	(4.37)
Obs.	186,338	425,134	620,728	876,476
R^2	0.14	0.10	0.09	0.08
Controls	Yes	Yes	Yes	Yes
Product FE	Yes	Yes	Yes	Yes
Street FE	Yes	Yes	Yes	Yes
Date-hour FE	Yes	Yes	Yes	Yes

Table IA.2: **Sample construction and finalized product returns**

This table reports sample construction and diagnostics for finalized product-return outcomes. Panel A documents the main screens used to form the estimation sample. Panel B calculates resolved-outcome rates among orders with non-missing payment method, baseline controls, product identifiers, destination-street identifiers, and date×hour cells. Resolved shares and return rates are reported in percent. *Return* is observed only for orders with finalized after-sales outcomes; unresolved cases are not coded as nonreturns. Older cohorts are orders submitted on or before January 10, 2026.

Panel A. **Sample construction**

Screen	Orders retained	Orders removed
Provider order panel	19,469,439	
Unit price and quantity both positive	17,710,731	1,758,708
Non-missing payment method, controls, and identifiers	16,310,389	1,400,342
Finalized after-sales outcome	12,890,784	3,419,605
Older-cohort robustness sample	12,177,188	713,596

Panel B. **Finalized-outcome diagnostics among eligible orders**

Subsample	Eligible orders	Resolved orders	Resolved share	Return rate
All orders	16,310,389	12,890,784	79.03	41.61
Non-BNPL orders	12,315,492	9,511,799	77.23	38.97
BNPL orders	3,994,897	3,378,985	84.58	49.04
Older cohorts	15,290,575	12,177,188	79.64	41.64
Recent cohorts	1,019,814	713,596	69.97	41.06

Table IA.3: **Payment channels, promotions, and the BNPL return relation**

This table reports estimation results for the relation between TikTok BNPL use and product returns under payment-channel and promotion restrictions. Panel A presents fixed-effects comparisons. Column (1) excludes Alipay and the separately coded pay-later channel; Columns (2) and (3) compare TikTok BNPL only with TikTok Pay and WeChat Pay, respectively. Column (4) additionally excludes orders with explicit payment discounts, and Column (5) also removes the 2024 Double 11, 2025 618, and 2025 Double 11 promotion windows. Panel B presents instrumental-variable results after reconstructing the initial province–product exposure and the lagged national shift within each restricted sample. The Anderson–Rubin (AR) p-value tests a zero structural effect; with one endogenous regressor and one excluded instrument, it equals the same-sample reduced-form test. Controls are as in Table 2. Standard errors are clustered at the product level. Observations report valid input orders before the fixed-effect estimator removes singleton groups. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

Panel A. **Fixed-effects comparisons**

	(1) No Alipay	(2) TikTok Pay	(3) WeChat Pay	(4) No payment discounts	(5) No major promotions
BNPL	3.11***	2.58***	4.28***	3.25***	3.28***
	(0.11)	(0.09)	(0.24)	(0.12)	(0.12)
Obs.	10,383,814	8,292,961	5,448,736	10,110,615	8,904,497
R^2	0.31	0.32	0.32	0.31	0.31
Controls	Yes	Yes	Yes	Yes	Yes
Product FE	Yes	Yes	Yes	Yes	Yes
Street FE	Yes	Yes	Yes	Yes	Yes
Date-hour FE	Yes	Yes	Yes	Yes	Yes

Panel B. **Restricted-sample IV estimates**

	(6) No Alipay	(7) No Alipay, payment discounts, or major promotions
BNPL (2SLS)	3.24***	3.27***
	(0.38)	(0.33)
Reduced form	3.15***	3.34***
	(0.39)	(0.35)
First-stage	0.97***	1.02***
	(0.03)	(0.02)
K-P Wald F	1433.18	1741.65
AR p-value	0.00	0.00
Obs.	10,382,158	8,796,970
Controls	Yes	Yes
Product FE	Yes	Yes
Street FE	Yes	Yes
Date-hour FE	Yes	Yes

Table IA.4: **Order characteristics associated with BNPL use**

Linear probability models relate BNPL use to order characteristics. The dependent variable is $BNPL(\%)$, equal to $100 \times BNPL$. Observations report valid input orders before the fixed-effect estimator removes singleton groups. Standard errors are robust in Columns (1) and (2) and clustered at the product level in Column (3). *, **, and *** denote significance at the 10%, 5%, and 1% levels.

	Dependent variable = $BNPL(\%)$		
	(1)	(2)	(3)
Log Price	2.47***	2.21***	1.06***
	(0.01)	(0.01)	(0.22)
Log Quantity	-1.67***	-0.86***	-0.13
	(0.09)	(0.09)	(0.39)
Log Shipping Fee	-0.06	-0.27***	-0.17
	(0.06)	(0.06)	(0.16)
Log Discount	0.24***	0.28***	0.55***
	(0.01)	(0.01)	(0.08)
Log Shipping Lag	5.26***	4.55***	0.23**
	(0.02)	(0.02)	(0.09)
Local Shipping	0.58***	-0.21***	-0.44***
	(0.04)	(0.05)	(0.12)
Obs.	12,890,784	12,890,784	12,890,784
R^2	0.02	0.04	0.07
Street FE	No	Yes	Yes
Date-hour FE	No	Yes	Yes
Product FE	No	No	Yes

Table IA.5: **Fixed-composition and extreme-shift diagnostics for the shift–share IV**

This table reports fixed-composition and extreme-shift diagnostics for the shift–share instrumental variable (IV). Panel A fixes category weights to August 20–September 18, 2024, and excludes the focal category; the double leave-one-out version also fixes province×category weights and excludes the focal province. Estimation begins after the weighting window, and Columns (3) and (4) exclude the three major public-promotion windows. Panel B trims the stated percentage from each tail of the fixed-category shift. Across category–date-hour cells, correlations of the raw and fixed-category shifts are, respectively, 0.49 and 0.15 with average price, 0.11 and 0.00 with average discount, 0.03 and 0.02 with order volume, and 0.04 and -0.06 with livestream share. Each column reports a separate two-stage least squares specification. Observations report valid input orders before singleton removal, and standard errors are clustered at the product level. *First-stage* gives the excluded-instrument coefficient; *K-P Wald F* is the Kleibergen–Paap rk Wald F statistic. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

Panel A. **Fixed composition and promotion exclusion**

	(1) Fixed-category LOO	(2) Double LOO	(3) Fixed-category LOO No promotions	(4) Double LOO No promotions
BNPL (2SLS)	4.00***	4.08***	4.14***	4.20***
	(0.43)	(0.42)	(0.44)	(0.43)
First-stage	1.08***	1.14***	1.01***	1.07***
	(0.03)	(0.04)	(0.03)	(0.04)
K-P Wald F	1055.02	969.79	841.32	758.66
Obs.	12,713,259	12,713,259	9,510,541	9,510,541
Controls	Yes	Yes	Yes	Yes
Product FE	Yes	Yes	Yes	Yes
Street FE	Yes	Yes	Yes	Yes
Date-hour FE	Yes	Yes	Yes	Yes

Panel B. **Influence of extreme fixed-category shifts**

	(1) Trim 1% from each tail	(2) Trim 2.5% from each tail	(3) Trim 5% from each tail
BNPL (2SLS)	3.99***	3.86***	3.90***
	(0.43)	(0.43)	(0.43)
First-stage	1.09***	1.10***	1.12***
	(0.03)	(0.03)	(0.04)
K-P Wald F	1016.45	993.42	996.50
Obs.	12,629,670	12,365,947	11,783,405
Controls	Yes	Yes	Yes
Product FE	Yes	Yes	Yes
Street FE	Yes	Yes	Yes
Date-hour FE	Yes	Yes	Yes

Table IA.6: **Exposure persistence and additional diagnostics for the shift–share IV**

This table reports exposure-persistence and additional diagnostics for the shift–share instrumental variable (IV). Panel A gives the correlation between the initial seven-day BNPL share and the post-baseline share for province–product cells with progressively larger baseline samples. Panel B reports the thicker-cell 2SLS estimate in Column (1). Column (2) interacts the exposure share with the national BNPL innovation 24 hours ahead; the innovation is the residual from projecting national take-up on its 24-, 48-, and 168-hour lags and hour-of-day and day-of-week indicators. The top row in Column (2) is the reduced-form return coefficient, and the next row is the coefficient from the corresponding BNPL regression. Observations report valid input orders before the fixed-effect estimator removes singleton groups, and standard errors are clustered at the product level. *K-P Wald F* denotes the Kleibergen–Paap (K-P) rk Wald F statistic. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

Panel A. **Exposure persistence in thicker province–product cells**

Minimum baseline orders	Pearson	Spearman	Province–product cells
$n_{\text{base}} \geq 5$	0.32	0.37	67,601
$n_{\text{base}} \geq 10$	0.40	0.46	42,310
$n_{\text{base}} \geq 20$	0.46	0.54	24,748

Panel B. **Additional instrumental-variables diagnostics**

	(1) Thicker cells $n_{\text{base}} \geq 10$	(2) Future innovation placebo
Two-stage least squares /	2.18**	3.49
reduced-form coefficient	(0.94)	(2.75)
First-stage /	1.33***	0.18***
BNPL coefficient	(0.04)	(0.04)
K-P Wald F	1210.72	
Obs.	10,085,730	12,869,885
Controls	Yes	Yes
Product FE	Yes	Yes
Street FE	Yes	Yes
Date-hour FE	Yes	Yes

Table IA.7: **Shift–share IV reduced-form and first-stage estimates with trend controls**

Reduced-form and first-stage diagnostics add trend controls to the shift–share instrumental-variable (IV) design. Column (1) adds the baseline exposure share interacted with a linear time trend. Columns (2) and (3) absorb province-specific or category-specific linear trends. All columns use the full IV sample; the 24-hour-lagged shift is unavailable for the first 2,134 orders. Observations report valid input orders before the fixed-effect estimator removes singleton groups. Standard errors are clustered at the product level. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

	(1) Exposure linear trend	(2) Province trends	(3) Category trends
Reduced form:	3.79***	4.33***	4.23***
Z^{SS}_{pgt}	(0.86)	(0.45)	(0.47)
First-stage:	1.00***	1.16***	1.16***
Z^{SS}_{pgt}	(0.06)	(0.03)	(0.03)
Reduced-form / first-stage ratio	3.78	3.73	3.65
First-stage F	313.57	1462.91	1458.59
Obs.	12,888,650	12,888,650	12,888,650
Controls	Yes	Yes	Yes
Trend controls	Exposure share × time	Province linear trends	Category linear trends
Product FE	Yes	Yes	Yes
Street FE	Yes	Yes	Yes
Date-hour FE	Yes	Yes	Yes

Table IA.8: **Shift–share IV placebo outcomes**

This table reports placebo-outcome regressions using the baseline shift–share instrument. Panel A uses the core predetermined order attributes from the baseline controls as outcomes in Columns (1)–(6). Panel B uses the payment-method-specific discount amount in RMB in Column (7), the late-shipment indicator multiplied by 100 in Column (8), and the nonnegative difference between actual and promised shipment times, measured in hours, in Column (9). Each column reports a separate regression on the full instrumental-variable (IV) sample. The baseline instrument is unavailable for the first 2,134 orders; Columns (8)–(9) additionally require observed shipment timing. Observations report valid input orders before the fixed-effect estimator removes singleton groups. Standard errors are clustered at the product level and reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

Panel A. **Predetermined order attributes**

	Log Price	*Log Quantity*	*Log Shipping Fee*	*Log Discount*	*Log Shipping Lag*	*Local Shipping*
	(1)	(2)	(3)	(4)	(5)	(6)
Z_{pgt}^{SS}	0.00	0.00	0.00	0.01	0.01**	-0.03
	(0.00)	(0.00)	(0.01)	(0.01)	(0.00)	(0.04)
Obs.	12,888,650	12,888,650	12,888,650	12,888,650	12,888,650	12,888,650
R^2	0.98	0.18	0.37	0.80	0.92	0.32
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Product FE	Yes	Yes	Yes	Yes	Yes	Yes
Street FE	Yes	Yes	Yes	Yes	Yes	Yes
Date-hour FE	Yes	Yes	Yes	Yes	Yes	Yes

Panel B. **Payment and fulfillment outcomes**

	Payment Discount Amount (RMB)	*Late Shipment*	*Shipment Delay (hours)*
	(7)	(8)	(9)
Z_{pgt}^{SS}	-0.31	-0.30*	-0.07
	(0.37)	(0.17)	(0.13)
Obs.	12,888,650	9,371,252	9,371,252
R^2	0.80	0.28	0.42
Controls	Yes	Yes	Yes
Product FE	Yes	Yes	Yes
Street FE	Yes	Yes	Yes
Date-hour FE	Yes	Yes	Yes

Table IA.9: **Negative-control-outcome diagnostics for the shift–share instrument**

This table reports negative-control-outcome tests based on historical outcomes for the same product. Columns (1) and (2) use the overall return rate among orders submitted 31–60 days before the focal order; columns (3) and (4) use the pre-shipment return rate over the same window. Columns (1) and (3) regress current product returns, measured in percentage points, on the corresponding standardized historical measure. Columns (2) and (4) regress that measure on the fixed-category leave-one-out shift–share instrument. Observations report valid input orders before the fixed-effect estimator removes singleton groups; each regression uses 4,396,625 orders after these removals. Standard errors are clustered at the product level and reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

	(1) *Product Return*	(2) *Historical Product Return Rate*	(3) *Product Return*	(4) *Historical Pre-Shipment Return Rate*
Historical Product Return Rate	-2.38***			
	(0.91)			
Historical Pre-Shipment Return Rate			-1.63**	
			(0.66)	
Z_{pgt}^{SS}		0.00		0.00
		(0.01)		(0.01)
Obs.	4,399,754	4,399,754	4,399,754	4,399,754
R^2	0.25	0.98	0.25	0.95
Controls	Yes	Yes	Yes	Yes
Product FE	Yes	Yes	Yes	Yes
Street FE	Yes	Yes	Yes	Yes
Date-hour FE	Yes	Yes	Yes	Yes

Table IA.10: **Weight concentration and exclusion sensitivity in the shift–share IV**

This table reports contribution-weight concentration and bounded-direct-effect sensitivity for the shift–share instrumental variable (IV). Using a fixed 10% diagnostic sample, Panel A residualizes the instrument, treatment, and outcome on the baseline controls and product, street, and date-hour FE, then aggregates absolute first-stage contributions by date×hour shock and province–product exposure cell. Entries give cumulative absolute weights, the Herfindahl index in percent, and contributor counts. Panel B allows a positive direct reduced-form effect $\gamma \in [0, \delta]$ on product returns. *Point-identified set* gives the range of $(\hat{\rho} - \gamma)/\hat{\pi}$; *AR 95% confidence set* inverts the Anderson–Rubin statistic using the joint covariance of the reduced-form and first-stage estimates. The baseline reduced form is 4.30, the first stage is 1.16, and the two-stage least squares estimate is 3.71, based on 12,888,650 valid input observations; 2,134 opening-window orders lack the 24-hour-lagged shift. A direct effect equal to 46.0% of the reduced form lowers the point estimate to 2.00, while the AR confidence set first includes zero at 78.7%; the point estimate reaches zero only when the direct effect equals the entire reduced form. The placebo calibration sets δ to the absolute late-shipment coefficient of 0.30 in Table IA.8, whose estimated sign is negative. Effects are measured in return-rate percentage points.

Panel A. **First-stage contribution-weight concentration**

Level	Top 1	Top 5	Top 10	Top 25 / Top 100	Herfindahl index (%)	Count
Date×hour shock	0.18%	0.76%	1.37%	2.96%	0.02	13,624
Province–product exposure cell	0.37%	1.27%	1.72%	5.43%	0.01	166,298

Panel B. **Sensitivity to bounded direct effects**

Sensitivity case	Maximum direct effect δ	Point-identified set	AR 95% confidence set
Strict exclusion (0%)	0.00	[3.71, 3.71]	[2.92, 4.50]
5% of reduced form	0.21	[3.52, 3.71]	[2.74, 4.50]
10% of reduced form	0.43	[3.34, 3.71]	[2.55, 4.50]
25% of reduced form	1.07	[2.78, 3.71]	[2.00, 4.50]
50% of reduced form	2.15	[1.85, 3.71]	[1.07, 4.50]
75% of reduced form	3.22	[0.93, 3.71]	[0.14, 4.50]
100% of reduced form	4.30	[0.00, 3.71]	[-0.79, 4.50]
Placebo calibration (6.96%)	0.30	[3.45, 3.71]	[2.67, 4.50]

Table IA.11: **Shift–share IV estimates with alternative exposure windows and lags**

Two-stage least squares (2SLS) estimates of Eq. (7) use $Z^{SS}_{pgt} = BNPL^{nat}_{t-L} \times BNPL^{base}_{pg}$. Columns (1)–(3) vary the exposure window, using 7, 14, and 30 days with $L = 24$; Columns (4)–(5) vary the lag length, using 12 and 48 hours with a 7-day exposure window. Standard errors are clustered at the product level. *Reduced form* and *First-stage* give the coefficient on Z^{SS}_{pgt}; *K-P Wald F* denotes the Kleibergen–Paap (K-P) rk Wald F statistic. Observations report valid input orders before singleton removal. They differ from the full sample only because the relevant lagged national BNPL series is unavailable near the beginning of the sample. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

	Dependent variable = *Return*				
	(1) 7-day $L = 24$	(2) 14-day $L = 24$	(3) 30-day $L = 24$	(4) 7-day $L = 12$	(5) 7-day $L = 48$
BNPL (2SLS)	3.71***	3.56***	3.65***	3.51***	3.67***
	(0.40)	(0.34)	(0.25)	(0.43)	(0.38)
Reduced form: Z^{SS}_{pgt}	4.30***	5.95***	8.75***	4.11***	4.22***
	(0.47)	(0.59)	(0.61)	(0.50)	(0.44)
First-stage: Z^{SS}_{pgt}	1.16***	1.67***	2.40***	1.17***	1.15***
	(0.03)	(0.04)	(0.05)	(0.03)	(0.03)
K-P Wald F	1458.83	2113.52	2568.97	1155.07	1440.57
Obs.	12,888,650	12,888,650	12,888,650	12,887,106	12,887,339
Controls	Yes	Yes	Yes	Yes	Yes
Product FE	Yes	Yes	Yes	Yes	Yes
Street FE	Yes	Yes	Yes	Yes	Yes
Date-hour FE	Yes	Yes	Yes	Yes	Yes

Table IA.12: **Exposure-support thresholds and inference for the shift–share IV**

This table reports support and inference diagnostics for the shift–share instrumental variable (IV). Panel A gives one two-stage least squares (2SLS) specification per column after imposing progressively thicker province–product baseline exposure cells; the leave-one-out (LOO) specifications recompute the national BNPL shift excluding the focal province. Panel B reports the baseline 2SLS estimate under alternative clustering choices; its 24-hour-lagged shift is unavailable for the first 2,134 orders. Observations report valid input orders before the fixed-effect estimator removes singleton groups. Standard errors are in parentheses. *K-P Wald F* denotes the Kleibergen–Paap (K-P) rk Wald F statistic. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

Panel A. **Support diagnostics across thicker exposure cells**

	(1) $n_{\text{base}} \geq 5$	(2) $n_{\text{base}} \geq 5$ LOO	(3) $n_{\text{base}} \geq 10$	(4) $n_{\text{base}} \geq 10$ LOO	(5) $n_{\text{base}} \geq 15$	(6) $n_{\text{base}} \geq 20$
BNPL (2SLS)	2.82***	2.79***	2.18**	2.11**	2.14*	2.33
	(0.68)	(0.68)	(0.94)	(0.94)	(1.20)	(1.42)
First-stage	1.27***	1.26***	1.33***	1.32***	1.36***	1.37***
	(0.03)	(0.03)	(0.04)	(0.04)	(0.04)	(0.05)
K-P Wald F	1494.52	1480.00	1210.72	1194.90	920.75	702.16
Obs.	11,071,931	11,032,059	10,085,730	10,053,458	9,422,090	8,937,859
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Product FE	Yes	Yes	Yes	Yes	Yes	Yes
Street FE	Yes	Yes	Yes	Yes	Yes	Yes
Date-hour FE	Yes	Yes	Yes	Yes	Yes	Yes

Panel B. **Alternative clustering for the baseline two-stage least squares estimate**

	(1) Product cluster	(2) Date-hour cluster
BNPL (2SLS)	3.71***	3.71***
	(0.40)	(0.26)
First-stage	1.16***	1.16***
	(0.03)	(0.01)
K-P Wald F	1458.83	9253.73
Obs.	12,888,650	12,888,650
Controls	Yes	Yes
Product FE	Yes	Yes
Street FE	Yes	Yes
Date-hour FE	Yes	Yes

Table IA.13: **Robustness of market-level retained-demand estimates**

This table reports robustness results for the merchant×product×date instrumental-variable analysis of retained demand. Coefficients give the effect of a 10 percentage-point increase in market BNPL share. Panel A uses retained orders and Panel B retained transaction value in RMB. Columns (1)–(3) use, respectively, a category leave-one-out national shift, two-way clustering by market and date, and a seven-day-lagged shift. Columns (4)–(6) require at least 30, 60, and 90 days to resolve returns. Specifications include the cell-level controls in Table 5. Observations report valid input cells before singleton removal. *First-stage* gives the excluded-instrument coefficient; *K-P Wald F* is the Kleibergen–Paap rk Wald *F* statistic. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

	Specification					
	(1) LOO category shift	(2) Two-way clustering	(3) Seven-day lag	(4) 30-day window	(5) 60-day window	(6) 90-day window
Panel A. **Dependent variable:** ***Retained orders***						
BNPL Share (+10 pp)	-1.08	1.26	-1.50	1.35	0.76	0.67
	(2.29)	(4.01)	(8.30)	(4.16)	(3.73)	(3.18)
First-stage	1.11***	0.82***	0.40***	0.80***	0.92***	1.08***
	(0.12)	(0.22)	(0.11)	(0.13)	(0.13)	(0.14)
K-P Wald *F*	85.28	14.01	12.69	38.55	47.28	62.28
Obs.	408,710	409,447	409,429	406,080	390,574	366,245
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Merchant×product FE	Yes	Yes	Yes	Yes	Yes	Yes
Category×date FE	Yes	Yes	Yes	Yes	Yes	Yes
Date clustering	No	Yes	No	No	No	No
Panel B. **Dependent variable:** ***Retained transaction value***						
BNPL Share (+10 pp)	149.54	434.96	563.98	452.10	283.29	148.57
	(190.41)	(346.98)	(582.09)	(326.79)	(288.20)	(223.11)
First-stage	1.11***	0.82***	0.40***	0.80***	0.92***	1.08***
	(0.12)	(0.22)	(0.11)	(0.13)	(0.13)	(0.14)
K-P Wald *F*	85.28	14.01	12.69	38.55	47.28	62.28
Obs.	408,710	409,447	409,429	406,080	390,574	366,245
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Merchant×product FE	Yes	Yes	Yes	Yes	Yes	Yes
Category×date FE	Yes	Yes	Yes	Yes	Yes	Yes
Date clustering	No	Yes	No	No	No	No

Table IA.14: **Alternative transformations of retained market outcomes**

This table reports merchant×product×date instrumental-variable results for alternative transformations of retained market outcomes. Coefficients give the effect of a 10 percentage-point increase in market BNPL share. Columns (1) and (2) apply the inverse hyperbolic sine transformation to retained orders and retained transaction value; Column (3) caps retained value at its 1st and 99th percentiles. Market BNPL share is instrumented with initial merchant×product exposure interacted with lagged national BNPL take-up. Controls are cell-level means of the order-level controls in Table 2. Observations report valid input cells before singleton removal, and standard errors are clustered at the merchant×product level. *First-stage* gives the excluded-instrument coefficient; *K-P Wald F* is the Kleibergen–Paap rk Wald F statistic. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

	Dependent variable		
	(1) IHS *retained orders*	(2) IHS *retained value*	(3) Capped *retained value*
BNPL Share (+10 pp)	0.08	0.13	75.09
	(0.05)	(0.12)	(93.48)
First-stage	0.82***	0.82***	0.82***
	(0.13)	(0.13)	(0.13)
K-P Wald F	40.97	40.97	40.97
Obs.	409,447	409,447	409,447
Controls	Yes	Yes	Yes
Merchant×product FE	Yes	Yes	Yes
Category×date FE	Yes	Yes	Yes

Table IA.15: **Zero-inclusive market-level estimates of gross, returned, and retained activity**

This table reports instrumental-variable results from a zero-inclusive street×category×date panel. The predetermined risk set contains markets with at least 50 orders and activity on at least 10 days during a common 60-day baseline window. Crossing 1,962 eligible markets with post-baseline dates yields 1,006,506 cells, 41.71% of which have no orders; estimation retains dates at least 60 days before the sample ends. Because BNPL share is undefined in zero-order cells, the endogenous variable is the number of BNPL orders. The instrument interacts predetermined market exposure and baseline order flow with a one-day-lagged, fixed-weight national BNPL shift that excludes the focal cell. Panel A reports order counts and nominal values; Panel B values each order at its market's baseline average. The AR rows report two-way-clustered Anderson–Rubin inference for retained outcomes on a common residualized sample of 879,919 cells, and the corresponding p-values determine their coefficient stars. Standard errors are clustered by market and date. Observations report valid input cells before singleton removal. *First-stage* gives the excluded-instrument coefficient; *K-P Wald F* is the Kleibergen–Paap rk Wald F statistic. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

Panel A. **Order counts and nominal transaction value**

	(1) *Orders*	(2) *Returned orders*	(3) *Retained orders*	(4) *Gross value*	(5) *Returned value*	(6) *Retained value*
BNPL orders	0.81	1.60***	-0.79	423.76***	339.90***	83.86*
	(0.67)	(0.17)	(0.67)	(60.97)	(54.53)	(31.29)
First-stage	2.68***	2.68***	2.68***	2.68***	2.68***	2.68***
	(0.67)	(0.67)	(0.67)	(0.67)	(0.67)	(0.67)
K-P Wald F	15.99	15.99	15.99	15.99	15.99	15.99
AR p-value			0.11			0.07
AR 95%			[-3.72,			[-10.60,
confidence set			0.10]			142.57]
Obs.	888,786	888,786	888,786	888,786	888,786	888,786
Market FE	Yes	Yes	Yes	Yes	Yes	Yes
Province×date FE	Yes	Yes	Yes	Yes	Yes	Yes
Category×date FE	Yes	Yes	Yes	Yes	Yes	Yes

Panel B. **Transaction value at baseline market prices**

	(1) *Gross value*	(2) *Returned value*	(3) *Retained value*
BNPL orders	135.78	228.26***	-92.48
	(86.29)	(28.02)	(83.05)
First-stage	2.68***	2.68***	2.68***
	(0.67)	(0.67)	(0.67)
K-P Wald F	15.99	15.99	15.99
Obs.	888,786	888,786	888,786
Market FE	Yes	Yes	Yes
Province×date FE	Yes	Yes	Yes
Category×date FE	Yes	Yes	Yes

Table IA.16: **Market-level counterfactual accounting by subgroup**

This table reports market-level counterfactual accounting results across merchant×product subgroups. The instrumental-variable estimates give the implied change associated with a 10 percentage-point increase in market BNPL share. Count outcomes are measured in orders and value outcomes in RMB. Market BNPL share is instrumented with initial merchant×product exposure interacted with lagged national BNPL take-up; second-stage estimates are rescaled from a zero-to-one share, while first-stage coefficients retain their original units. Panel A includes all cells. Panel B includes markets below the median baseline return rate, measured over their first 30 days among markets with at least 50 baseline orders. Controls are cell-level means of the order-level controls in Table 2. Observations report valid input cells before singleton removal, and standard errors are clustered at the market level. *First-stage* gives the excluded-instrument coefficient; *K-P Wald F* is the Kleibergen–Paap rk Wald F statistic. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

	Dependent variable					
	(1) *Orders*	(2) *Returned orders*	(3) *Retained orders*	(4) *Gross value*	(5) *Returned value*	(6) *Retained value*
Panel A. **All merchant×product×date markets**						
BNPL Share (+10 pp)	4.40	3.13	1.26	942.20	507.23	434.96
	(5.82)	(2.19)	(4.03)	(610.23)	(356.66)	(315.71)
First-stage	0.82***	0.82***	0.82***	0.82***	0.82***	0.82***
	(0.13)	(0.13)	(0.13)	(0.13)	(0.13)	(0.13)
K-P Wald F	40.97	40.97	40.97	40.97	40.97	40.97
Obs.	409,447	409,447	409,447	409,447	409,447	409,447
Panel B. **Low baseline-return merchant×product markets**						
BNPL Share (+10 pp)	22.28	9.70	12.58	5,068.04	1,864.52	3,203.53
	(42.23)	(12.71)	(30.59)	(3,942.96)	(1,489.94)	(2,511.20)
First-stage	0.81***	0.81***	0.81***	0.81***	0.81***	0.81***
	(0.27)	(0.27)	(0.27)	(0.27)	(0.27)	(0.27)
K-P Wald F	8.78	8.78	8.78	8.78	8.78	8.78
Obs.	144,966	144,966	144,966	144,966	144,966	144,966
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Merchant×product FE	Yes	Yes	Yes	Yes	Yes	Yes
Category×date FE	Yes	Yes	Yes	Yes	Yes	Yes

Table IA.17: **Default repayment-date placebo tests**

This table reports default repayment-date placebo tests that recenter a local day-specific specification on alternative dates of the month. In Panel A, Column (1) centers the specification on the actual due date, the 6th, and Columns (2)–(4) use the 3rd, 9th, and 10th. Panel B uses the 12th, 15th, 20th, and 25th in Columns (5)–(8). The reported coefficient is the average *BNPL* interaction over days +1 to +3. The regressions use the full 12,890,784-order estimation sample, the baseline controls, and the fixed effects shown below. Standard errors are clustered at the product level. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

Panel A. **Due date and nearby placebo dates**

	Dependent variable = *Return*			
	(1)	(2)	(3)	(4)
	6th	3rd	9th	10th
	Actual due date	Placebo	Placebo	Placebo
Average	-0.32**	-0.10	0.04	-0.02
days +1 to +3	(0.14)	(0.15)	(0.14)	(0.15)
Obs.	12,890,784	12,890,784	12,890,784	12,890,784
R^2	0.31	0.31	0.31	0.31
Controls	Yes	Yes	Yes	Yes
Product FE	Yes	Yes	Yes	Yes
Street FE	Yes	Yes	Yes	Yes
Date-hour FE	Yes	Yes	Yes	Yes

Panel B. **Additional placebo dates**

	Dependent variable = *Return*			
	(5)	(6)	(7)	(8)
	12th	15th	20th	25th
Average	-0.10	0.04	0.15	-0.05
days +1 to +3	(0.16)	(0.16)	(0.14)	(0.12)
Obs.	12,890,784	12,890,784	12,890,784	12,890,784
R^2	0.31	0.31	0.31	0.31
Controls	Yes	Yes	Yes	Yes
Product FE	Yes	Yes	Yes	Yes
Street FE	Yes	Yes	Yes	Yes
Date-hour FE	Yes	Yes	Yes	Yes

Table IA.18: **Simultaneous candidate purchases and product return composition**

This table examines whether BNPL finances simultaneous candidate purchases and changes their return composition. Same-second groups contain orders submitted to the same destination street in the same second; unclassified products are excluded throughout. A multi-order group contains at least two orders; a strict bundle also requires the same merchant and at least two distinct products. Panel A estimates bundle membership in Columns (1)–(2) and interacts BNPL with group indicators in Columns (3)–(4), using singleton and non-strict orders as the respective base groups. Panel B compares all-BNPL with all-non-BNPL groups under the broad definition in Columns (1)–(4) and the strict definition in Columns (5)–(8). Binary outcomes are multiplied by 100, so their coefficients are in percentage points; the retained-orders outcome remains a count. All regressions include the baseline controls; bundle-level regressions also control for group size and product and merchant composition. Observations report valid input orders or bundles before the fixed-effect estimator removes singleton groups. Standard errors are clustered by product in Panel A and by street in Panel B. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

Panel A. **Bundle formation and return amplification**

	Multi-order group (%)	*Strict bundle* (%)	*Return*	*Return*
	(1)	(2)	(3)	(4)
BNPL	-0.18***	-0.11***	3.27***	3.28***
	(0.03)	(0.01)	(0.11)	(0.11)
BNPL× two-order group			2.48***	
			(0.32)	
BNPL× three-or-more group			1.80***	
			(0.50)	
BNPL× strict bundle				2.69***
				(0.32)
Obs.	12,843,553	12,843,553	12,843,553	12,843,553
R^2	0.12	0.12	0.31	0.31
Controls	Yes	Yes	Yes	Yes
Product FE	Yes	Yes	Yes	Yes
Street FE	Yes	Yes	Yes	Yes
Date-hour FE	Yes	Yes	Yes	Yes

Panel B. **Return composition by bundle definition**

	Broad groups				Strict bundles			
	Partial return	*Keep one*	*All returned*	*Retained orders*	*Partial return*	*Keep one*	*All returned*	*Retained orders*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
All-BNPL bundle	0.60*	0.72**	5.43***	-0.14***	0.39	0.59	4.50***	-0.11***
	(0.34)	(0.32)	(0.42)	(0.01)	(0.49)	(0.46)	(0.55)	(0.01)
Obs.	88,761	88,761	88,761	88,761	55,217	55,217	55,217	55,217
R^2	0.29	0.26	0.40	0.51	0.38	0.35	0.49	0.59
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Bundle controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Street FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Merchant FE	No	No	No	No	Yes	Yes	Yes	Yes
Date-hour FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table IA.19: **BNPL and product returns across product types**

This table reports estimates of heterogeneity in the BNPL–product-return relation across product types. Column (1) interacts *BNPL* with two-digit consumption-purpose categories; the omitted category is Food & non-alcoholic beverages, and unclassified products are excluded. Column (2) groups categories into retail and non-retail and retains the full sample; *BNPL* gives the non-retail effect and *BNPL* × *Retail* gives the incremental retail effect. Controls are as in Table 2. Observations report valid input orders before the fixed-effect estimator removes singleton groups. Standard errors are clustered at the product level. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

	Dependent variable = *Return*	
	(1)	(2)
BNPL× *Clothing & footwear*	1.85***	
	(0.21)	
BNPL× *Personal care & miscellaneous*	2.12***	
	(0.35)	
BNPL× *Housing & utilities*	1.87***	
	(0.21)	
BNPL× *Recreation, sport & culture*	1.23***	
	(0.30)	
BNPL× *Furnishings & household equipment*	1.63***	
	(0.38)	
BNPL (*Food & non-alcoholic beverages*; base category)	2.03***	
	(0.14)	
BNPL× *Education*	-8.18**	
	(3.22)	
BNPL× *Health*	0.64	
	(0.63)	
BNPL× *Alcohol & tobacco*	1.82**	
	(0.79)	
BNPL× *Transport*	1.03***	
	(0.37)	
BNPL× *Restaurants & accommodation*	3.12***	
	(0.90)	
BNPL× *Information & communication*	0.96***	
	(0.23)	
BNPL (*Non-retail goods*)		2.76***
		(0.42)
BNPL × *Retail*		0.58
		(0.44)
Obs.	12,843,553	12,890,784
R^2	0.31	0.31
Controls	Yes	Yes
Product FE	Yes	Yes
Street FE	Yes	Yes
Date-hour FE	Yes	Yes

Table IA.20: **Machine-learning diagnostics of heterogeneity in BNPL effects on product returns**

This table reports machine-learning heterogeneity diagnostics. Panel A estimates sample-split group average treatment effects (GATES) from five random splits. In each split, a training sample constructs a predicted BNPL-effect score from interactions with price quartiles, district per capita gross domestic product (GDP) quartiles, product category, and the night-order indicator; the holdout sample is sorted into quintiles by this score. Panel B averages classification analysis (CLAN) characteristics across splits. Return, BNPL share, and the remaining share characteristics are reported in percent; log price and log shipping fee are natural logarithms. Average observations report valid holdout-sample inputs before the fixed-effect estimator removes singleton groups. Standard errors are clustered at the product level. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

Panel A. **Group average treatment effects estimates**						
	Dependent variable = *Return*					
	Q1 Lowest	Q2	Q3	Q4	Q5 Highest	Q5–Q1 gap
BNPL	1.88***	2.84***	3.38***	3.77***	4.42***	2.54***
	(0.20)	(0.17)	(0.17)	(0.18)	(0.25)	(0.32)
Average Obs.	3,222,102	3,222,102	3,222,102	3,222,102	3,222,102	3,222,102
Sample splits	5	5	5	5	5	5
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Product FE	Yes	Yes	Yes	Yes	Yes	Yes
Street FE	Yes	Yes	Yes	Yes	Yes	Yes
Date-hour FE	Yes	Yes	Yes	Yes	Yes	Yes
Panel B. **Classification-analysis characteristics**						
Average group size	674,052	639,226	643,421	665,621	599,783	
Return	22.68	27.53	45.40	58.42	55.01	
BNPL Share	20.35	22.69	27.39	30.00	31.01	
Log Price	3.07	3.30	4.22	4.83	4.97	
High-Price Share	6.69	11.03	37.22	39.88	32.49	
Night-Order Share	22.17	19.88	24.05	27.71	34.49	
Low-GDP-District Share	15.59	23.35	22.36	25.01	37.80	
Log Shipping Fee	0.04	0.03	0.02	0.02	0.02	