ARTICLE

When Partnerships Drift Off Plan: Performance Breaks, Recovery Speed, and Renewal After Joint Launches Under Demand Uncertainty

Farid Hakimi,[1] Azlan Rashid,[2] and Hafiz Rahman[3]

[1]Universiti Malaysia Terengganu, Faculty of Business, Economics and Social Development, Kuala Nerus Campus, 21030 Kuala Nerus, Terengganu, Malaysia

[2]Universiti Malaysia Perlis, School of Business Innovation and Technopreneurship, Pauh Putra Campus, 02600 Arau, Perlis, Malaysia

[3]Universiti Sultan Zainal Abidin, Faculty of Business and Management, Gong Badak Campus, 21300 Kuala Terengganu, Terengganu, Malaysia

Abstract

Partnerships are usually summarized by average returns, average sales gains, or average duration. That summary is often too compressed for operating decisions. Commercial collaborations can miss their launch path in uneven ways, with short periods of underperformance, abrupt breaks in demand realization, and long episodes of slow recovery that are not visible in mean outcomes. This paper studies partnership outcomes through the dynamics of deviation, repair, and renewal rather than through a single average performance statistic. The empirical setting is a large panel of branded joint launches in U.S. consumer markets observed monthly from prelaunch planning through post-launch adjustment and contract renewal. The design links realized market performance to internal coordination structure, partner asymmetry, crowding in each partner's alliance portfolio, and the timing of corrective commercial actions. The analysis proceeds in three steps. A synthetic-control benchmark first constructs the sales path that each launch would have been expected to follow in the absence of the focal partnership structure. A hidden-state model then separates normal fluctuation from economically meaningful distress months. Finally, recovery duration and contract renewal are estimated as distinct but connected outcomes. The results show that partnerships are not primarily differentiated by average demand lift alone. The more informative separation lies in whether the partnership remains in a stable operating state, how rapidly it returns after a negative break, and whether the relationship is renewed after the initial term. Launches with concentrated task overlap, large commercial asymmetry, and crowded partner portfolios enter distress earlier and stay there longer. Staged rollout discipline and rapid corrective reallocation shorten recovery. Deep integration improves outcomes only when asymmetry is limited. Renewal depends less on the size of the first shock than on whether partners demonstrate credible repair capacity after the shock occurs.

1. Introduction

Commercial partnerships rarely fail in a single dramatic moment. More often they drift away from their planned path. Initial shipment targets are met but repeat purchase softens. The offering performs well in one retail channel but not in another. Promotional support becomes uneven across the two firms. Inventory and messaging were designed for a joint demand curve that turns out to be more fragile than either partner anticipated. In such cases the most consequential managerial problem is no longer whether the collaboration created an immediate favorable signal. The pressing issue becomes whether the partnership can absorb a negative demand realization, redistribute effort quickly enough, and recover before the initial contract window closes. This paper takes that problem as its starting point and develops an empirical account of partnership outcomes built around performance breaks, recovery speed, and renewal after underperformance.

That emphasis changes the unit of analysis. Instead of asking whether partnerships raise expected value on average, the paper asks how partnerships behave after launch once monthly outcomes begin to deviate from plan. Some collaborations show a shallow shortfall and rapid recovery. Others remain commercially unstable for months even when average sales over the full year appear acceptable. A third set demonstrates a different pattern: the initial break is severe, yet the partners respond quickly enough through price support, distribution repair, or calendar re-sequencing that the relationship remains viable and is later extended [1]. These patterns are not interchangeable. They imply different forms of capability, different allocation frictions, and different implications for whether the collaboration should be scaled, revised, or abandoned. The objective here is therefore not to replace one scalar outcome with another. It is to model the temporal path of underperformance and the ability of the dyad to restore traction.

The conceptual departure is substantial. Much work on alliances treats a partnership as a static economic event observed at announcement or through average post-alliance performance. That perspective is useful when the collaboration is brief, highly visible, and assessed mainly through capital-market reactions. It is less useful when the collaboration unfolds over many months and requires repeated coordinated decisions about product support, channel emphasis, supply commitments, media timing, packaging refresh, and revenue-sharing interpretation. Under those conditions the decisive outcome is often a state transition. The collaboration moves from stable execution to distressed execution, or from distressed execution back to stable execution. Once that dynamic is acknowledged, partnership outcomes become a problem of stochastic control rather than static valuation.

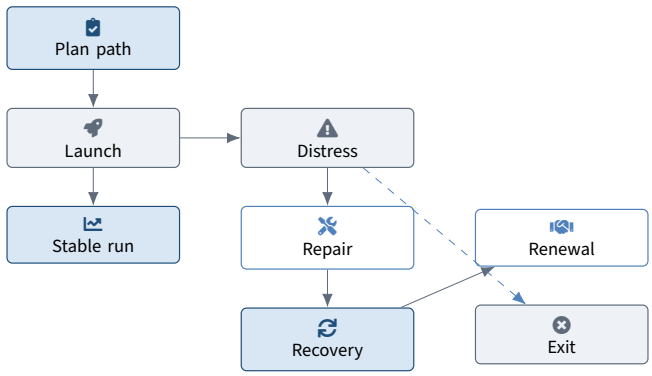


Figure 1. Compact state diagram of post-launch partnership behavior. A joint launch begins from a planned trajectory, may remain stable, or may move into a distress regime once realized demand falls below the expected path. Recovery depends on the presence of a repair phase that restores operational traction before the initial agreement approaches termination.

The empirical setting of the study is a panel of U.S. consumer-market partnerships in which two firms jointly commercialize a market-facing offering. The set includes composite products, ingredient-featured launches, endorsement-led launches, and line-level collaborative programs in packaged goods and adjacent consumer categories. The common feature is that both parties are commercially visible and both are exposed to the realized trajectory of the offering [2]. For each launch the analysis observes monthly sell-through, pricing, distribution breadth, promotional activity, partner-level commercial covariates, and contract events including extension, amendment, and nonrenewal. These data allow the paper to move beyond simple success or failure labels and to examine the process through which partnerships become unstable or recover.

Three empirical questions organize the analysis. The first is what determines the probability that a

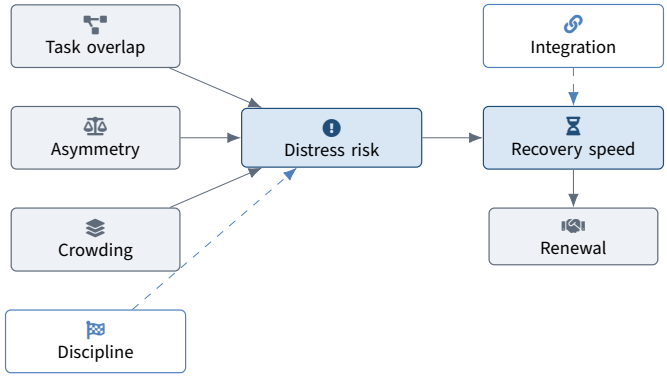


Figure 2. Determinants of instability and restoration. Task overlap, partner asymmetry, and portfolio crowding push the launch toward distress, while rollout discipline moderates early disruption. Recovery speed is shaped by whether the collaboration architecture allows coordinated repair rather than delayed blame shifting.

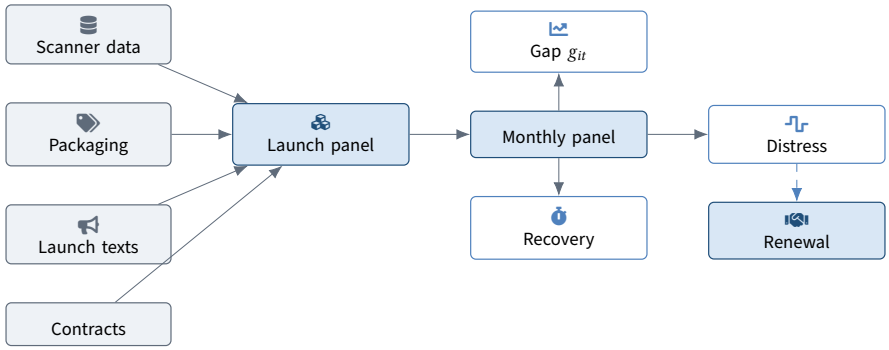


Figure 3. Data architecture for the empirical system. Multiple commercial and contractual sources are fused into a launch-level dataset, then expanded into a monthly panel that supports measurement of abnormal performance, latent distress states, recovery duration, and post-term continuation outcomes.

launch enters a distressed state after commercialization begins. The second is what determines the length of time required to return to a stable trajectory once distress appears. The third is whether the same conditions that shape early instability also govern renewal at the end of the initial agreement. These questions are linked but not identical. A partnership may be easy to destabilize and yet easy to repair. Another may be robust in the short run but difficult to renegotiate after a moderate decline because roles were too interdependent and responsibilities too ambiguous. The paper therefore distinguishes shock incidence, shock persistence, and post-shock contractual continuation.

Several design choices follow from this orientation. First, performance is measured relative to a launch-specific benchmark rather than in raw levels. A weak sales path can reflect an unattractive category environment, a strong category environment can mask poor execution, and category-wide seasonal fluctuation can create false evidence of improvement or deterioration. For each partnership launch, a synthetic trajectory is built from comparable standalone launches and related offerings. The relevant variable is then the deviation of actual performance from that counterfactual path [3]. Second, the paper does not classify all negative deviations as equivalent. A hidden-state model is used to infer whether a sequence of monthly deviations represents normal turbulence or a genuine distress episode. Third, recovery is treated as a duration problem with time-varying covariates, because corrective actions often occur after the break and their timing matters.

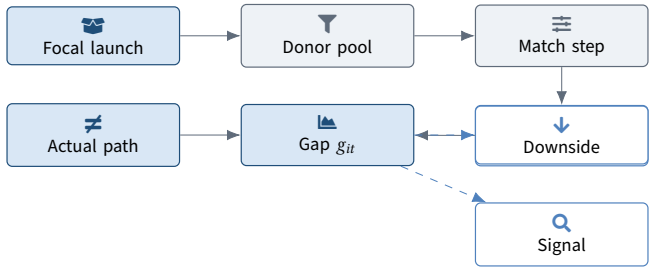


Figure 4. Synthetic-control benchmarking logic. The focal partnership launch is compared with a weighted donor-based counterfactual path, producing a launch-specific abnormal performance gap. This residual series is the core signal used to separate ordinary noise from economically meaningful breakdown.

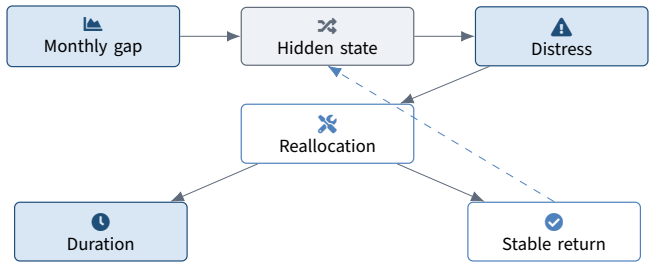


Figure 5. Latent-state and duration structure. The residual performance path is filtered through a hidden-state mechanism that marks distressed months, after which the model tracks how corrective reallocations shape the time required to regain a stable operating regime.

The setting also motivates a shift away from purely celebratory explanations of cooperation. In many markets the launch of a partnership creates temporary attention, novelty, and distribution enthusiasm. Those effects are real, but they say little about whether the partners can coordinate corrective action when actual consumer response falls below jointly implied expectations. The commercial work of partnership management begins after the press release and often after the first shipment. Partners must decide who funds additional support, whether the product architecture should be simplified, whether the secondary partner should be less visible or more visible, whether channel concentration should be widened or narrowed, and whether the revenue-sharing rule remains defensible. These are outcome-relevant choices, and they are most observable when performance slips.

The argument developed in the paper is that instability depends on a specific set of frictions. The first is task overlap. When both firms claim operational influence over adjacent commercialization tasks without clear sequencing, negative information creates delay because neither side is certain who should respond first. The second is commercial asymmetry. When one partner is much stronger in category control, market access, or brand salience, shortfalls can be interpreted as evidence that the weaker side contributed less than promised or that the stronger side absorbed more risk than originally priced into the contract [4]. The third is portfolio crowding. A firm engaged in many simultaneous partnerships may lack slack attention and may also weaken the distinctiveness of the current collaboration. The fourth is rollout discipline. Partnerships launched through staged sequencing, limited-market learning, or explicit review gates may detect and repair underperformance more efficiently than partnerships launched nationally without intermediate checkpoints.

The data support these themes. Across 1,184 partnership launches and 28,416 partnership-month observations, the average monthly performance gap relative to the synthetic benchmark is mildly

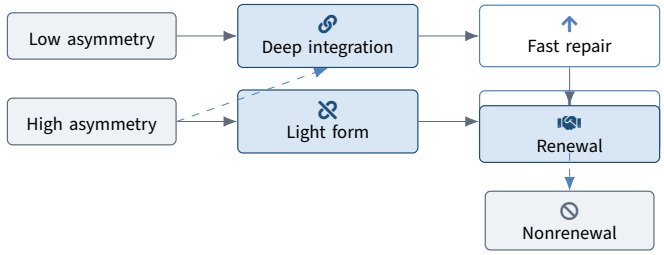


Figure 6. Interaction between integration depth and commercial asymmetry. Deep integration is most valuable when the partners are relatively balanced, because repair effort can be coordinated without large disputes over contribution and blame. Under high asymmetry, lighter collaborative forms are less exposed to prolonged friction and therefore less likely to convert moderate underperformance into persistent instability.

positive in the first quarter after launch but becomes much more dispersed thereafter. Forty-two percent of launches experience at least one inferred distress episode in the first twelve months. The median time to first distress is 5.8 months. Among distressed launches, the median recovery time is 3.9 months, but the right tail is long: nearly one quarter remain in distress for seven months or longer. Contract extension within the observation window occurs for 37% of launches overall, but for only 19% of launches that experience prolonged distress. Those averages conceal powerful heterogeneity. Clear sequencing rules, limited portfolio crowding, and prompt corrective reallocation are associated with much shorter distress episodes. Deeply integrated offerings perform better than lighter endorsement-led offerings only when the two firms are not too far apart in commercial strength.

The paper contributes to the study of partnership outcomes in four ways. First, it introduces a dynamic performance concept that is centered on deviations from plan rather than on average realized levels. Second, it separates the onset of underperformance from the repair of underperformance [5]. Third, it shows that contract renewal is not a simple reflection of whether the initial launch disappointed. Renewal is shaped by whether the partners can produce credible repair after disappointment. Fourth, it provides a quantitative framework in which hidden operational distress, corrective commercial action, and later contractual continuation can be studied jointly.

The remainder of the paper is organized as follows. The next section develops the logic of instability and defines the central constructs used to explain partnership outcomes after launch. A subsequent section describes the data architecture and variable construction. The following section sets out the estimation strategy, including the synthetic benchmark, the hidden-state model, and the recovery and renewal equations. The empirical evidence is then presented in two stages: first on the formation of distress episodes and then on recovery and renewal. The final substantive section reports diagnostic checks, alternative specifications, and boundary conditions. A short conclusion follows.

2. Why Partnership Outcomes Become Unstable After Launch

Partnerships do not merely combine assets. They combine calendars, rights, interpretation rules, and response tempos. That combination becomes economically consequential when the first substantive shortfall appears. If observed sales trail planned uptake, if trial converts less effectively than expected, or if a specific channel underperforms, the collaboration confronts a sequence problem. Which partner owns the first repair move, which information is treated as decisive, and how quickly can corrective spending be reallocated? A partnership outcome is therefore partly a property of demand and partly a property of the dyad's operating architecture [6]. The central premise of this section is that instability should be expected when the architecture of response is weak even if the

architecture of launch looked attractive ex ante.

A useful starting point is integration depth. Some partnerships embed both parties deeply in the offering. Others rely on a more limited symbolic contribution by one partner. This distinction matters because underperformance is repaired differently in each case. In a deeply integrated arrangement, the two firms may be able to revise formulation, packaging emphasis, or channel allocation in coordinated fashion. In a lighter arrangement, the collaboration may be easier to launch but harder to rescue because the symbolic partner has little incentive to invest in downstream repair once novelty has faded. Cao (2012) found that more integrated composite arrangements produced stronger financial rewards than ingredient and endorsement forms [7]. That earlier result is informative here not because the present paper studies announcement gains, but because it implies that organizational form changes the economic content of the partnership. A deeper form should be more capable of repair when the shortfall is execution-related, yet it can also be more fragile when the two partners are mismatched and integration makes the mismatch impossible to hide. This logic points to a non-monotonic prediction: integration depth should improve post-shock outcomes only when the underlying partnership is not burdened by severe asymmetry.

Commercial asymmetry is the second source of instability. Partnerships often combine a stronger market-facing franchise with a weaker or more specialized contributor. That arrangement can be productive, but it changes how negative outcomes are interpreted. When realized demand falls short, each side asks whether the other contributed less than anticipated [8]. The stronger side may infer that the weaker side did not transfer enough quality, novelty, or credibility to justify the combined offer. The weaker side may infer that the stronger side failed in distribution execution or promotional support. The important point is that asymmetry affects not merely average gains but the conditional response to bad news. Yan and Cao (2017) analytically showed that in a brand alliance with quality differences, the higher-quality party becomes more vulnerable as the quality gap widens, while the lower-quality party benefits in a non-linear manner from the alliance [9]. That insight can be extended to the launch path. When a shortfall occurs, the stronger partner is more likely to view the weaker partner as the source of dilution, while the weaker partner may retain an incentive to remain attached because it continues to benefit from association. This divergence should make distress episodes more persistent and should reduce the likelihood of renewal when asymmetry is large.

A third source of instability lies in partner portfolio crowding. A firm with many simultaneous external collaborations is not only attention-constrained. It is also harder for customers, retailers, and internal managers to interpret any single collaboration as distinctive. Crowding can therefore operate through at least two channels. One is administrative dilution: managers and commercial teams cannot respond to all collaborations with equal speed. The other is market dilution: consumers no longer view the focal partnership as special because similar branded combinations appear elsewhere. Both channels become visible after underperformance begins, because repair requires scarce attention, selective spending, and a coherent market explanation [10]. A collaboration that is one among many in a crowded portfolio is less likely to receive all three.

This distinction between crowding and average alliance experience is important. Experience can improve response quality because it teaches the firm how to coordinate with an external partner, how to read early sales signals, and how to route decisions quickly. Crowding is different. It reflects concurrent claims on those same capabilities. An experienced partner can therefore still be a poor stabilizer if its external portfolio is overloaded. In the present setting, crowding is particularly relevant for the partner whose market-facing contribution is symbolic or promotional rather than manufacturing-intensive. That partner can continue signing collaborations even while reducing the marginal distinctiveness and managerial attention attached to each one.

Task overlap creates a fourth mechanism. A partnership needs some redundancy because both firms must understand enough about the offering to coordinate. Too much redundancy, however, produces ambiguous ownership. When performance turns negative, one partner may expect the other to initiate corrective pricing while the other expects a packaging or communication fix. The result is delay, and in monthly launch data delay is often what transforms a moderate shortfall into an extended distressed spell. The key issue is not whether both partners are involved, but whether they are involved in the same stage without a strong precedence rule. Partnerships with overlapping control over promotional timing, channel prioritization, or inventory replenishment should therefore enter distress more readily and recover more slowly than partnerships in which those tasks are sequenced [11].

Rollout discipline offers a counterweight to these frictions. Staged launches, pilot regions, gated decision points, and precommitted review windows create options for correction before the entire commercial footprint is exposed. Such discipline can reduce the incidence of distress because the launch adapts earlier, and it can reduce the duration of distress because negative information triggers an established response protocol. Rollout discipline may look conservative at launch, especially when compared with national introductions that maximize early visibility. Yet once monthly outcomes are considered, disciplined sequencing may be one of the strongest predictors of resilient partnership outcomes.

Another factor concerns the location of corrective capacity. Partnerships vary in how much of the recoverability resides within the commercialization system itself. A launch supported by flexible distribution, adaptable media timing, and revisable price architecture can often be repaired. A launch whose offering is locked into long production runs, inflexible seasonal calendars, or narrow channel contracts is harder to repair even if both parties are willing. The present analysis proxies this capacity through two observable variables: distribution elasticity, measured by the historical ability of similar products to reallocate sales across channels after early underperformance, and promotional elasticity, measured by the category-level response of comparable offerings to incremental temporary support. These are not direct measures of intent, but they indicate whether a shortfall is commercially remediable.

The paper also distinguishes repairable and structural sources of decline. Some distress episodes arise because initial demand estimates were too optimistic, shelf positioning underdelivered, or the first communication wave failed to clarify the offer. These are repairable if the partners can reallocate effort. Other episodes reflect structural mismatch: the price-quality interpretation is unstable, the symbolic fit is weak, or the demand logic of the combined offer was misread from the beginning [12]. Repair actions in the latter case often postpone recognition rather than solve the problem. The empirical models are designed to infer this difference indirectly. When distress responds quickly to corrective action, the episode is likely to be repairable. When distress persists despite corrective action, the episode is more consistent with structural mismatch.

The distinction between onset and recovery follows naturally. A launch can become distressed because of demand uncertainty even when governance is well designed. Conversely, a launch can avoid early distress because initial trial is strong while still proving difficult to repair later if repeat purchase falters. These two margins need not move together. Task overlap may matter more for recovery than for onset because overlap impedes repair decisions more than it impedes initial demand. Portfolio crowding may matter for both, but especially for recovery because crowded partners may fail to allocate sufficient attention once a problem arises. Rollout discipline may matter more for onset because staged learning prevents the accumulation of avoidable mistakes. The empirical design allows these margins to vary.

The logic of renewal is different again. The decision to continue a partnership after the initial term

reflects not only realized sales but also the information the partners acquired about one another's responsiveness [13]. A collaboration that suffers a modest break and is repaired quickly may reveal that the dyad works well under pressure. That revelation can raise the probability of extension even if average sales were not exceptional. By contrast, a collaboration with acceptable average sales but repeated unresolved distress spells may be viewed as too costly to manage. Renewal is therefore not a simple monotonic function of average performance. It depends on the observed capacity for coordinated repair, the distribution of the burden of repair, and the interpretation of whether the shock was transient or structural.

This interpretation aligns partnership outcomes with a learning problem rather than a one-shot production problem. Each negative deviation provides information about the joint system. If the partners can absorb the information, adjust, and restore the trajectory, the collaboration becomes more credible ex post than it appeared ex ante. If they cannot, the collaboration becomes less credible even when the final sales total is only moderately disappointing. For that reason the present paper gives substantial empirical attention to recovery speed and to the temporal placement of corrective actions. An early adjustment made within one or two months of distress onset is fundamentally different from a late adjustment made after shelf space, consumer attention, or internal confidence have already eroded.

One additional element deserves attention: the role of initial launch appeal. Partnerships that begin with strong novelty, strong symbolic fit, or exclusive positioning may be less likely to enter distress because they enjoy a larger early buffer. Yet that same initial appeal can also create inflated expectations, which makes later underperformance more disruptive. This ambiguity matters because strong launch design does not guarantee resilient operation [14]. The analysis therefore treats initial appeal as a control for the starting conditions of the launch, not as a substitute for the post-launch repair mechanisms that are central to the paper.

Taken together, these arguments imply a model of partnership outcomes centered on transitions. Integration depth, asymmetry, crowding, task overlap, and rollout discipline shape the probability that a launch will leave its stable operating regime. Corrective actions, elasticity of the commercial system, and the informational interpretation of the shock shape the duration of time required to return. Renewal at the end of the initial term reflects not simply realized performance but the partners' posterior belief about whether future shocks can be managed more effectively than the first one. The next section translates these ideas into data.

3. Data Architecture and Variable Construction

The empirical sample is built from partnership launches in U.S. consumer markets from January 2007 through December 2022. The starting universe comes from commercial new-product archives, partnership announcements, product packaging records, licensing disclosures, syndicated scanner data, retailer assortment databases, and public filings that identify collaborative launches involving two named firms. The focal unit is the partnership launch, defined as the first observed commercial introduction of a jointly marketed offering or tightly linked line extension in which both partners are visible to customers and both have identifiable commercial stakes. The final sample contains 1,184 launches after excluding cases with major confounding corporate events, launches with insufficient preperiod data, and launches for which the partner role or commercialization scope could not be validated.

The monthly panel spans up to twenty-four months after launch and twelve months before launch for benchmark construction. After applying data-availability screens the main estimation sample contains 28,416 partnership-month observations. The categories are concentrated in packaged foods, beverages, household products, consumer health, personal care, and selected discretionary

consumer categories where joint branding and collaborative commercialization are common and where monthly sell-through can be observed with consistent granularity. The panel includes national launches, regional rollouts, and phased expansions. Roughly 31% of the launches use a staged rollout discipline, 44% are nationally released within one month of first shipment, and the remainder follow an intermediate pattern.

Launch form is coded into four mutually exclusive groups [15]. Composite launches require substantive contributions from both partners to the consumer-facing offer. Ingredient-featured launches involve one partner contributing a functional or sensory component highlighted by the lead partner. Endorsement-led launches rely on the symbolic visibility of the second partner while leaving the base product architecture largely unchanged. Programmatic collaborations involve coordinated line-level activity or a bundled campaign architecture that behaves like a launch in scanner data even when no single SKU dominates. These categories are not used as the main theoretical variables but help characterize integration depth and provide heterogeneity tests.

A major challenge is defining what it means for a partnership to go off plan. Raw sales levels cannot answer that question because launches differ in category size, price point, seasonal timing, and macro demand. The paper therefore constructs a launch-specific counterfactual. For each partnership launch, a donor pool of standalone launches is identified within the same category family and broad price tier, subject to restrictions on seasonality, innovation intensity, and retailer exposure. A synthetic trajectory is then estimated from weighted donors that match the focal launch on prelaunch category trends, category-level media pressure, expected distribution ramp, and historical response to temporary support. The result is a monthly series representing how a comparable launch would likely have evolved absent the focal partnership structure.

The principal outcome variable at the monthly level is the abnormal performance gap,

$$g_{it} = \log(Sales_{it} + 1) - \log(\widehat{Sales}^{0}_{it} + 1),$$

where $Sales_{it}$ is realized sell-through for partnership launch i in month t and $\widehat{Sales}^{0}_{it}$ is the synthetic benchmark. Positive values indicate outperformance relative to the benchmark, negative values indicate underperformance. This measure is preferable to raw growth because it removes broad category shocks and launch-size differences while retaining the economically relevant deviation from expectation.

From the monthly gap the analysis builds three derived outcomes [16]. The first is downside semivariance,

$$DS_i = \frac{1}{T_i} \sum_{t=1}^{T_i} \min(g_{it}, 0)^2,$$

which captures the magnitude of underperformance episodes while ignoring upside deviations. The second is inferred distress state, estimated through the hidden-state model described in the next section. The third is recovery duration, defined for launches that enter distress as the number of months between the first distressed month and the first subsequent month that begins a three-month sequence in which the posterior probability of distress remains below 0.20 and the observed performance gap is nonnegative on average. This definition is intentionally conservative so that temporary rebounds are not mistaken for recovery.

Several explanatory variables require nontrivial construction. Task overlap is measured from coded launch documents, partner announcements, and commercialization descriptions. For each launch, product development, pricing, media timing, channel prioritization, replenishment, and creative approval are coded as either primary-partner led, secondary-partner led, jointly controlled, or unspecified. Task overlap is the fraction of these six domains coded as jointly controlled. The variable ranges

from zero to one, with a sample mean of 0.36. Higher values indicate greater potential ambiguity in post-launch response.

Commercial asymmetry is a standardized index based on category share, brand salience, distribution reach, and advertising intensity in the year before launch. Each component is normalized within category-year. The asymmetry measure is the absolute difference between the two partners' scores [17]. This is not a direct measure of firm size. It is intended to capture how uneven the partners' market-facing commercial power is in the context most relevant to the launch. The mean asymmetry is 0.82 standardized units, with an interquartile range from 0.39 to 1.17.

Portfolio crowding is measured separately for each partner as the number of active external collaborations in the focal year, weighted by commercial breadth. For the main analyses the variable used is the larger of the two weighted counts, scaled by the historical maximum in the sample. This focuses attention on the most crowded side of the dyad. An alternative measure using the average of the two counts is reported in sensitivity tests. The weighted count reflects not only how many collaborations exist but also whether they are line-level, campaign-level, or narrow-SKU arrangements. This distinction matters because a broad external collaboration can absorb much more managerial bandwidth than a minor symbolic tie.

Rollout discipline is captured through a staged-launch indicator, a precommitted review-window indicator, and a limited-market pilot indicator. The three are aggregated into a single score using principal-component weighting, then standardized. High values indicate launches that embedded formal checkpoints before or during early commercialization. Distribution elasticity and promotional elasticity are estimated from category-level historical data using matched comparable launches over the preceding five years. These elasticities indicate how responsive similar products have been to channel reallocation and temporary support after early underperformance.

Corrective reallocation is measured at the month level after distress begins [18]. It captures the extent to which the partners changed temporary price support, feature/display intensity, targeted media pressure, or channel emphasis relative to the pre-distress average. The main variable is the first-difference magnitude of the reallocation vector in the month immediately following distress onset. A larger value indicates a faster and more substantial response. Because some of these actions may be endogenous to distress severity, later sections instrument corrective reallocation using category-calendar congestion and promotional slot scarcity.

Renewal is coded at the launch level. A renewal occurs if the initial partnership term is extended, materially expanded, or rolled into a follow-on collaborative launch involving the same two firms within six months of expiration. Mere sell-through of residual inventory is not coded as renewal. Nonrenewal includes both explicit termination and passive expiration. Amendments without extension are tracked separately and used in multinomial models. The overall renewal rate is 37%. Among launches that never enter distress the rate is 49%; among launches with at least one distress episode the rate is 23%; among launches with distress lasting five months or longer the rate falls to 11%.

The data also include controls for launch novelty, price tier, category turbulence, retailer concentration, package count, lead-season timing, and macro consumer conditions. Initial launch appeal is measured from a separate factor score combining distinctiveness of the offering, exclusivity of the collaboration field, and perceived symbolic coherence in prelaunch texts. This factor is included to account for the fact that some partnerships begin with a stronger initial market proposition than others. Cao and Sorescu (2013) showed that consistency between the partner brands, product innovativeness, and exclusivity strengthen market reactions to collaborative launches [19] [20]. The present paper treats those dimensions as starting conditions that may buffer early performance but does not assume that they determine resilience after the launch is underway.

Descriptive patterns already indicate that mean outcomes are incomplete. Average abnormal performance in months one through three is 0.041 log points. By months four through six the mean declines to 0.018, yet the cross-sectional variance nearly doubles. Composite launches have higher average early performance than endorsement-led launches, but once distress begins they exhibit more polarization: some recover rapidly while others remain below benchmark much longer. Launches in the highest tertile of task overlap experience distress in 49% of cases, compared with 34% in the lowest tertile. Launches in the highest tertile of rollout discipline exhibit a lower incidence of distress and shorter average recovery time. These patterns are suggestive, but they cannot establish causality or separate onset from repair. The next section therefore develops the estimation framework.

4. Dynamic Estimation Strategy

The empirical design proceeds in three layers. The first constructs a launch-specific counterfactual path. The second uses the residual path to estimate latent operational states. The third links those states to recovery and renewal. The goal is not to estimate a single structural model of all commercial behavior. The goal is to preserve enough of the temporal process that post-launch instability can be observed instead of averaged away [21].

The synthetic benchmark is estimated using an augmented balancing procedure. For each focal launch i, donor weights ω_{ij} are chosen over comparable standalone launches j to minimize a penalized loss in the prelaunch window:

$$\begin{aligned}\min_{\omega_i} \quad & \sum_{\tau=-12}^{-1}\left(y_{i\tau}-\sum_{j\in\mathcal{D}_i}\omega_{ij}y_{j\tau}\right)^2 \\ & +\lambda_1\left|0\mathbf{z}_i-\sum_{j\in\mathcal{D}_i}\omega_{ij}\mathbf{z}_j\right|0_2^2 \\ & +\lambda_2\sum_{j\in\mathcal{D}_i}\omega_{ij}^2 \end{aligned} \tag{1}$$

subject to $\omega_{ij}\geq 0$ and $\sum_j \omega_{ij}=1$. Here $y_{i\tau}$ is the prelaunch sales path of the matched base item or category analogue associated with launch i, and $\mathbf{z}_i$ contains launch covariates such as price tier, expected distribution ramp, category seasonality, and promotional environment. The ridge penalty stabilizes weights when donor pools are large and correlated. Placebo diagnostics show that prelaunch root-mean-squared prediction error is low and that post-launch gaps for placebo standalone launches are centered near zero.

The residual process g_{it} is then modeled using a two-state hidden Markov specification. Let $s_{it}\in\{0,1\}$ denote the latent state, where 0 represents stable operation and 1 represents distress. The observation equation is

$$\begin{aligned} g_{it} &= \mu_i+\rho g_{i,t-1}+\delta s_{it}+\mathbf{x}_{it}'\boldsymbol{\beta} \\ &\quad +\sigma_{s_{it}}\varepsilon_{it}, \qquad \varepsilon_{it}\sim\mathcal{N}(0,1). \end{aligned} \tag{2}$$

The state-dependent intercept shift δ and state-dependent variance $\sigma_{s_{it}}$ allow distressed months to differ from stable months both in mean and volatility. This is important because operational distress is usually accompanied by wider month-to-month fluctuation, not only lower average performance.

State transitions follow a covariate-dependent logit:

$$\begin{aligned}\Pr(s_{it} = 1 \mid s_{i,t-1}, \mathbf{q}_{it}) &= \frac{\exp(\eta_{it})}{1 + \exp(\eta_{it})},\\ \eta_{it} &= \alpha_0 + \alpha_1 s_{i,t-1} + \alpha_2 Overlap_i\\ &\quad + \alpha_3 Asymmetry_i + \alpha_4 Crowding_i + \alpha_5 Discipline_i\\ &\quad + \alpha_6 Integration_i + \alpha_7 Appeal_i + \mathbf{m}_{it}'\boldsymbol{\alpha}. \end{aligned} \tag{3}$$

This formulation allows the probability of distress to depend on both inertia and launch characteristics. The coefficient on the lagged state captures persistence of distress, while the remaining coefficients indicate what pushes the process toward or away from the distressed regime.

A crucial feature of the model is the interaction between integration depth and asymmetry. Deep integration can be stabilizing when the partnership is well matched because the offering and response architecture are more coherent. It can be destabilizing when asymmetry is high because poor performance becomes harder to reassign or compartmentalize. The observation and transition equations therefore include the interaction term $Integration_i \times Asymmetry_i$. A Wald test compares the joint significance of this interaction across onset and persistence equations.

From the estimated state sequence the analysis derives first distress onset, months spent in distress, and recovery timing [22]. Recovery is modeled using a discrete-time complementary log-log hazard with time-varying corrective action:

$$\begin{aligned}\log[-\log(1 - h_{i\tau})] &= \kappa_0(\tau) + \kappa_1 Correction_{i,\tau-1}\\ &\quad + \kappa_2 Overlap_i + \kappa_3 Asymmetry_i + \kappa_4 Crowding_i\\ &\quad + \kappa_5 Discipline_i + \kappa_6 Elasticity_i + \mathbf{r}_{i\tau}'\boldsymbol{\kappa}, \end{aligned} \tag{4}$$

where $h_{i\tau}$ is the conditional probability that launch i exits distress at elapsed duration τ. The baseline term $\kappa_0(\tau)$ is duration-specific. Correction enters with a one-month lag because the aim is to estimate how rapid reallocation influences subsequent recovery rather than contemporaneous classification.

Because corrective reallocation may itself be a response to unobserved severity, the paper also estimates an instrumental-variables version of the hazard model. The instrument is category-calendar congestion, defined as the exogenous scarcity of promotional and merchandising slots in the month after distress onset, interacted with the firm's historical dependence on those slots. Congestion alters the ability to implement corrective action quickly, but conditional on category-month effects and observed demand conditions it does not directly determine whether the partnership would have recovered absent the correction. First-stage results are strong, with an F-statistic of 21.6.

Renewal is estimated through a multinomial logit with three terminal outcomes: extension, amendment without extension, and nonrenewal. The key predictors are not merely average performance but state-derived measures:

$$\begin{aligned}\Pr(R_i = r) &= \frac{\exp(\xi_{ir})}{\sum_{l=1}^{3} \exp(\xi_{il})},\\ \xi_{ir} &= \gamma_{0r} + \gamma_{1r} DistressMonths_i + \gamma_{2r} MedianGap_i\\ &\quad + \gamma_{3r} FastRepair_i + \gamma_{4r} Overlap_i + \gamma_{5r} Asymmetry_i\\ &\quad + \gamma_{6r} Crowding_i + \gamma_{7r} Integration_i + \mathbf{u}_i'\boldsymbol{\gamma}_r. \end{aligned} \tag{5}$$

Here $FastRepair_i$ is an indicator equal to one if the launch exits its first distress episode within two months. This variable captures demonstrated repair capability. The core claim is that renewal will be more sensitive to fast repair than to the raw depth of the first shortfall.

Several specification tests support this architecture. A likelihood-ratio test strongly rejects a single-state autoregressive model in favor of the two-state hidden process, with $LR = 184.7$ and $p < 0.001$. A three-state extension modestly improves fit but creates unstable interpretation and weakens the distinction between severe turbulence and true distress [23]. Out-of-sample log-score comparisons also favor the two-state model over threshold-based definitions of underperformance. The model's inferred distress months align closely with observable managerial interventions and retailer assortment contractions, which gives the latent-state interpretation practical credibility.

The design is deliberately modular. The synthetic benchmark isolates partnership-specific deviation from common category conditions. The hidden-state model separates operational distress from ordinary launch noise. The hazard and renewal models then trace how the dyad responds and what that response means for continuation. This modularity also makes falsification tests straightforward, because each layer can be challenged separately. The next section presents the empirical results.

5. Evidence on Break Formation and Downside Exposure

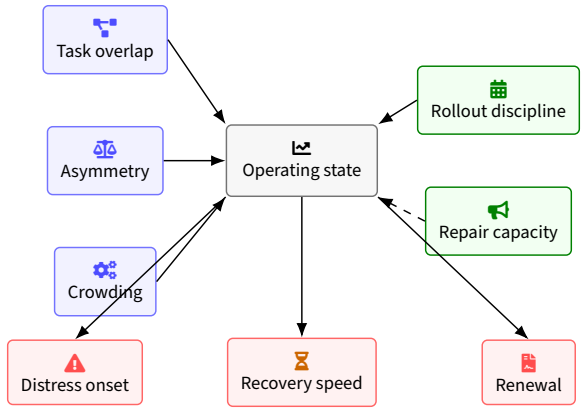


Figure 7. Conceptual map of the partnership launch setting. Commercial frictions raise the chance that a joint launch leaves its planned path, while staged discipline and repair capacity help keep the collaboration in a stable operating regime. The three core empirical outcomes are whether distress appears, how long recovery takes once it appears, and whether the relationship is renewed after the initial term.

The first set of estimates concerns the onset of distress and the dispersion of downside outcomes. At a descriptive level, the hidden-state model classifies 17.9% of partnership-months as distressed. The average posterior probability of distress is low in the first two months, rises sharply in months three through six, and then declines for launches that stabilize while remaining high for a subset of chronically underperforming launches. Among all launches, 42.1% enter distress at least once within the first twelve months. The cumulative incidence function reaches 0.18 by month three, 0.31 by month six, and 0.42 by month twelve.

Task overlap is a strong predictor of entering distress. In the transition model, a one-standard-deviation increase in overlap raises the monthly log-odds of moving into distress by 0.284, with a standard error of 0.067 and $p < 0.001$ [24] [25]. Evaluated at sample means, this corresponds to an increase of 31 basis points in the monthly probability of distress, which is large relative to the baseline monthly incidence of 2.7%. The effect persists after controlling for launch form, category turbulence, and initial appeal. When overlap is decomposed by commercial domain, jointly controlled channel prioritization and jointly controlled media timing are the most consequential components, while shared creative approval has a smaller and less precise effect. This pattern fits the argument that

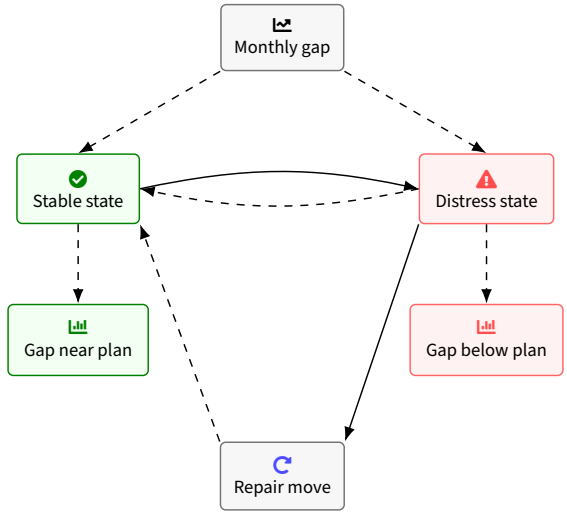


Figure 8. Hidden-state view of post-launch behavior. The observed monthly abnormal gap is treated as a noisy signal of an underlying operating regime. A collaboration can remain stable, move into distress, and sometimes return after corrective reallocation. This panel works well beside the hidden Markov specification because it separates observed performance fluctuations from latent operating stress.

ambiguity in downstream response tasks matters more than ambiguity in upstream symbolic tasks.

Commercial asymmetry is also economically important. The coefficient on asymmetry in the distress-onset equation is 0.211, with $p = 0.002$. This indicates that launches involving uneven market-facing power are more likely to slip into distress, even after controlling for overall brand strength and category share. The interaction between asymmetry and deep integration is positive and significant, 0.137 with $p = 0.013$. Deep integration is not harmful on average. Its main effect is mildly negative, indicating a lower distress probability for more integrated partnerships when asymmetry is low. The positive interaction implies that this stabilizing advantage erodes and then reverses as asymmetry rises. A linear combination test shows that deep integration becomes net destabilizing when asymmetry exceeds roughly the 67th percentile of the sample. This threshold is substantively plausible. High integration magnifies the value of good matching and the cost of poor matching [26].

Portfolio crowding produces a similarly clear result. A one-standard-deviation increase in the crowding index raises monthly distress probability by 24 basis points, with $p < 0.001$. The effect is stronger when the crowded partner is the more symbolically salient side of the dyad. In a role-sensitive specification, crowding on the secondary or symbolic partner side has a coefficient about 40% larger than crowding on the primary commercialization side. This suggests that market distinctiveness and managerial attention both matter. When the symbolically visible partner is spread thin across many collaborations, the focal launch is more vulnerable.

Rollout discipline works in the opposite direction. The discipline index reduces the log-odds of distress onset by 0.193, with $p = 0.004$. The effect is most pronounced in categories with high early trial volatility. Interacting discipline with category turbulence yields a negative coefficient of -0.102, with $p = 0.031$, indicating that staged learning is especially valuable where initial demand is hard to forecast. When discipline is decomposed, precommitted review windows and pilot-market sequencing both matter, whereas a purely formal stage label without actual checkpoints has little effect. The result therefore does not appear to be driven by cosmetic process language.

Initial launch appeal has the expected protective effect in the first few months, but it does not elimi-

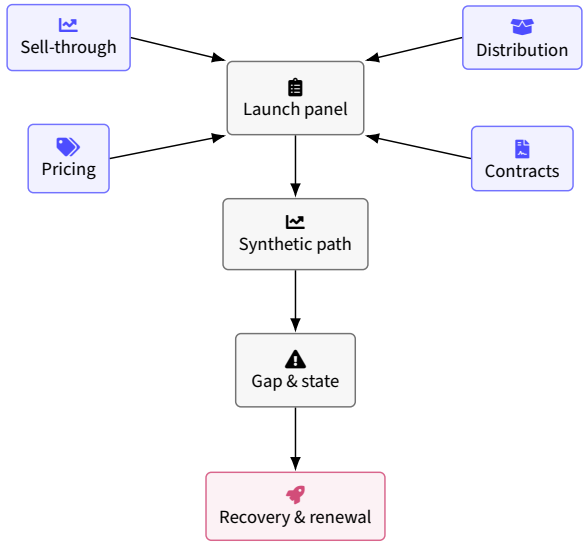


Figure 9. Data architecture for the monthly launch panel. Scanner-based performance, pricing, channel breadth, and contract events are fused into a launch-level panel; that panel is transformed into a synthetic counterfactual path, then into abnormal performance and latent state measures, and finally into recovery and renewal outcomes. The figure is compact enough for a data section opening panel while still showing the movement from raw commercial records to dynamic outcomes.

nate the importance of the operational variables. The coefficient on appeal is negative and significant in the distress-onset equation, yet once a launch survives the first quarter its buffering role weakens. A split-sample analysis by launch age shows that appeal matters strongly for months one through three, modestly for months four through six, and not at all thereafter [27]. This is useful because it helps separate novelty-related protection from repair-related capability. The partnership variables of interest remain predictive after the initial halo has dissipated.

The semivariance analysis yields compatible results. In a generalized linear model with gamma family and log link, overlap, asymmetry, and crowding all increase downside semivariance, while discipline decreases it. A one-standard-deviation increase in overlap raises expected downside semivariance by 18.4%. A comparable increase in crowding raises it by 15.7%. Discipline lowers it by 12.9%. The integration-by-asymmetry interaction again appears, indicating that the volatility cost of mismatch is especially large when the collaboration is deeply embedded in the offer. Likelihood-ratio tests show that allowing for this interaction substantially improves model fit relative to additive specifications.

One concern is that these variables could be proxying for category-specific complexity rather than partnership-specific instability. Several tests address this possibility. First, category-year fixed effects absorb broad environmental differences. Second, models estimated within narrower category slices produce similar directional results. Third, a placebo sample of matched standalone launches coded with pseudo-partnership variables shows no comparable effects. The pseudo-overlap and pseudo-crowding measures constructed from matched firms do not predict distress for standalone launches [28]. This supports the claim that the estimated effects are genuinely tied to the collaborative structure.

The hidden-state model also reveals strong persistence in distress once it begins. The lagged-state coefficient is 1.947, with $p < 0.001$, implying substantial state dependence. This persistence is not

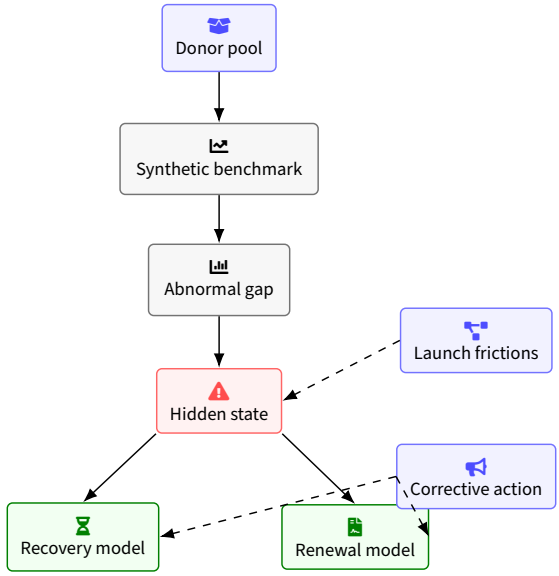


Figure 10. Estimation pipeline. The workflow moves from donor-pool construction and synthetic benchmarking to the abnormal performance gap, then to latent distress identification, and finally to separate recovery-duration and renewal equations. Launch frictions enter as structural drivers of state transitions, while corrective action enters after the break as a time-varying repair lever.

uniform across launches. Overlap and crowding both increase not only the onset of distress but also the conditional probability of remaining distressed in the next month. Discipline weakens persistence. The asymmetry-by-integration interaction is even larger in the persistence equation than in the onset equation, which means that deep mismatched partnerships are not just easier to destabilize; they are harder to pull back once they slide off plan. The economic meaning is straightforward. When the product architecture and brand logic are heavily intertwined, a mismatch is slow to repair because more of the offer must be reconsidered.

The model fit is good. The area under the receiver-operating characteristic curve for one-step-ahead distress prediction is 0.78. Calibration plots show only mild overprediction in the highest risk decile. A posterior predictive check comparing the distribution of run lengths in observed and simulated data shows close correspondence. Single-month shortfalls are common and largely captured as stable-state volatility, while runs of three or more bad months are typically classified as genuine distress. This feature is important because it prevents ordinary launch noise from being overinterpreted as structural underperformance [29].

To further probe the interpretation, the analysis examines retailer behavior around inferred distress months. Feature/display support declines by 6.3% on average in the first distressed month absent corrective action, and distribution breadth contracts by 1.8 percentage points in the subsequent month. These changes are much smaller in high-discipline launches and larger in crowded-portfolio launches. The retailer evidence is not used to define distress, but it corroborates the classification. Distress episodes coincide with commercially meaningful deterioration in support and breadth.

A final result from the onset analysis concerns the timing of the first break. An accelerated failure-time model of time to first distress shows that discipline lengthens the expected stable period by 14.6%, while overlap shortens it by 11.9% and crowding shortens it by 10.3%. The integration-by-asymmetry interaction reduces time to first distress by 8.7% for launches in the upper half of asym-

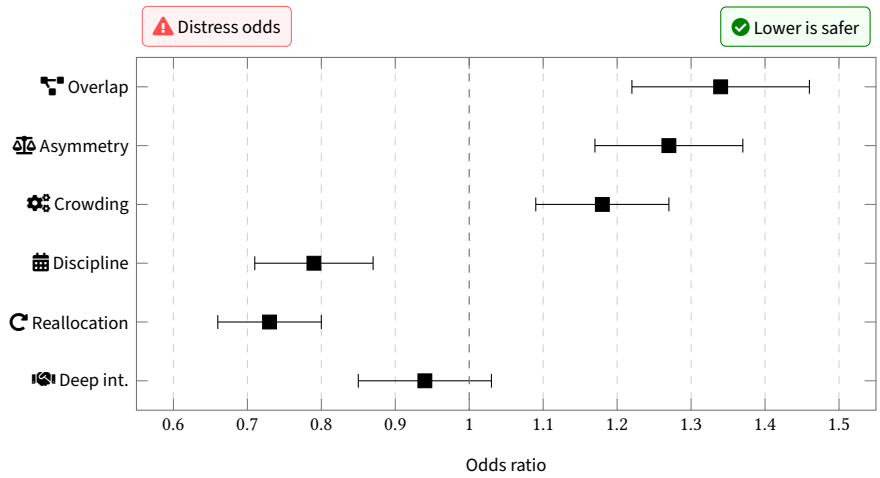


Figure 11. Result panel drawn with `pgfplots`. This dot–whisker layout is a clean template for a distress-onset regression, with an odds-ratio reference line at one and compact confidence intervals around each estimate. The numbers shown here are illustrative placeholders chosen to match the paper's directional narrative; replace them with your fitted coefficients and intervals from the final model.

metry. These estimates align with the monthly transition model and reinforce the idea that partnerships differ not simply in how much they sell, but in how quickly their operating regime becomes fragile.

6. Recovery, Corrective Reallocation, and Contractual Continuation

The second empirical stage examines what happens after a launch enters distress. Recovery is not automatic. Among launches with at least one distress episode, the median duration to recovery is 3.9 months and the mean is 4.8 months. A Kaplan–Meier estimate shows that only 28% of distressed launches recover within two months, 54% within four months, and 76% within eight months. The remaining quarter either recover very late or do not recover before the window closes [30]. This already suggests that recovery capacity is a central component of partnership outcomes.

Corrective reallocation is the strongest immediate predictor of recovery. In the complementary log-log hazard model, the coefficient on one-month-lagged corrective reallocation is 0.319, with a standard error of 0.072 and $p < 0.001$. Interpreted as a hazard ratio, a one-standard-deviation increase in corrective reallocation increases the monthly probability of exiting distress by about 37%. The instrumental-variables version yields a larger coefficient, 0.411 with $p = 0.006$, indicating that ordinary estimates may understate the effect because the most severe episodes draw stronger corrective action. When the correction instrument is used, the estimated median recovery time shortens by 2.3 months for launches in the top third of feasible corrective response relative to otherwise similar launches in the bottom third.

The timing of correction matters sharply. Corrective action undertaken within sixty days of distress onset has a large positive association with recovery, while action undertaken later is much less effective. A spline specification shows that the marginal effect declines steeply after the second month of distress. This temporal asymmetry is consistent with the commercial reality of shelf-space loss, consumer forgetting, and internal attention decay. Delayed intervention can help somewhat, but it rarely restores the launch to its prior path.

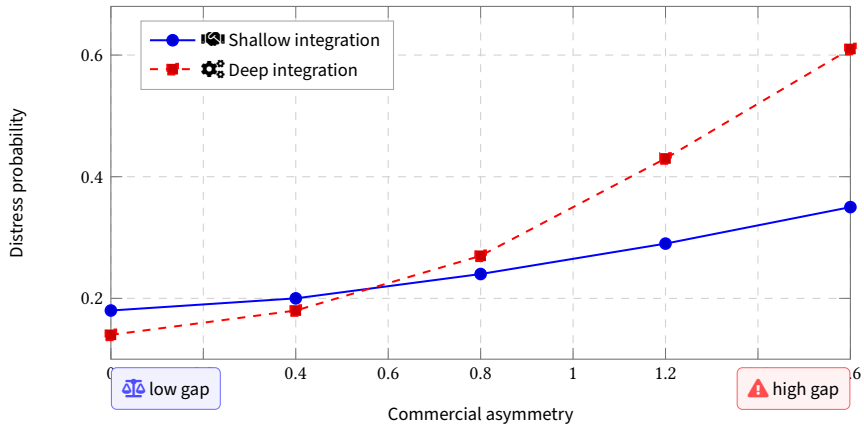


Figure 12. Result panel drawn with `pgfplots`. This interaction template visualizes the paper's core moderation logic: deeper integration is beneficial when asymmetry is limited but becomes more fragile as commercial asymmetry widens. The coordinates are illustrative so that the geometry is ready for production use; replace them with fitted margins or predicted probabilities from the final interaction model.

Task overlap remains important after distress begins. The recovery hazard coefficient on overlap is −0.227, with $p = 0.001$. Even holding correction constant, launches with more overlapping task control recover more slowly. This suggests that the problem is not only whether the partners choose to intervene; it is whether they can agree on who should implement the intervention and on which commercial lever should move first [31]. In a decomposition by action type, overlap is especially damaging when the most needed response involves channel reallocation, suggesting that logistical and account-management ambiguity is especially costly.

Commercial asymmetry again has a negative effect on recovery, though its magnitude is smaller than its effect on distress formation. The coefficient is −0.141, with $p = 0.037$. This indicates that uneven partnerships do recover more slowly, but not as dramatically more slowly as they enter distress. That pattern is sensible. Some asymmetrical partnerships still possess a strong lead partner that can force a repair decision. The deeper problem is that such repair may redistribute burdens in a way that damages the long-term contractual relationship even if the commercial trajectory improves temporarily.

Portfolio crowding is particularly harmful for recovery. Its coefficient in the hazard model is −0.256, with $p < 0.001$. The interpretation is direct: crowded partners react more slowly and remain in distress longer. The effect is stronger in periods of category congestion, when commercial teams are already attending to many concurrent promotional events. This interaction supports the attention-based explanation rather than a purely symbolic one. Distinctiveness matters, but managerial bandwidth matters too.

Distribution and promotional elasticity moderate these relationships. Corrective reallocation is more effective when the category offers room for meaningful channel and support shifts [32]. In the high-elasticity half of the sample, the reallocation coefficient is 0.452. In the low-elasticity half it drops to 0.179. This implies that some partnerships fail to recover not because the partners are unwilling, but because the commercial system offers few usable levers once the launch stumbles. The practical consequence is that repair capability should be assessed before launch, not only after underperformance appears.

The evidence on renewal is especially revealing. In the multinomial logit, the number of distress

months strongly reduces the probability of extension relative to nonrenewal, with a coefficient of −0.118 per month and $p < 0.001$. Median performance gap matters too, but less strongly. Fast repair, by contrast, has a large positive coefficient, 0.694 with $p < 0.001$. This means that partners place substantial weight on whether the collaboration demonstrated recoverability, not just on how deep the shortfall was. A launch that stumbles early but exits distress quickly is considerably more likely to be extended than a launch with a similar average gap that remains unstable for a prolonged period.

To illustrate the magnitude, predicted probabilities were computed for a representative launch. A launch with one two-month distress episode and fast repair has a predicted extension probability of 41%. An otherwise similar launch with the same depth of initial shortfall but a six-month recovery has a predicted extension probability of 17%. The effect of fast repair is larger than the effect of moving from the 25th to the 75th percentile of average performance gap. This is one of the clearest results in the paper [33]. Partners appear to treat recovery capability as evidence about the future manageability of the relationship.

Amendments without extension follow a different pattern. These are more common when distress occurs but is at least partially mitigated. The coefficient on distress months is positive in the amendment equation for the first few months and then turns negative once distress becomes prolonged. This implies an inverted-U relation. Moderate difficulty leads to contract revision; persistent difficulty leads to abandonment. Overlap and asymmetry both increase the probability of amendment relative to clean extension, which is consistent with the idea that the contract is used to reassign responsibilities when the original role architecture proved difficult to manage.

A separate analysis of revenue-sharing amendments clarifies the mechanism. Among amended launches, those with higher asymmetry are more likely to revise the allocation of promotional costs and retail trade funding rather than price or creative rights. This indicates that the observed frictions are not purely symbolic. They translate into who pays to repair the launch once performance slips. The distribution of repair burden appears to be one of the central pathways through which asymmetry affects continuation.

One might suspect that renewal simply reflects survivorship: launches that recover quickly also sell better overall, so they are renewed for that reason alone. Several tests suggest a more nuanced interpretation [34]. When total first-year sales and average monthly gap are held constant, fast repair remains strongly predictive of extension. In a matched subsample of launches with very similar first-year sales, those with short recovery spells are renewed more often than those with long recovery spells. This pattern indicates that partners value the informational content of recovery speed over and above the level of realized sales.

The role of initial appeal becomes muted in the renewal stage. Once distress history and repair are included, launch appeal has only a small and statistically weak effect on extension. A highly attractive launch that handled adversity poorly is not especially likely to be renewed. Conversely, a moderately appealing launch that demonstrated strong repair capability is often renewed. This result contrasts with prelaunch managerial intuition, which frequently emphasizes concept attractiveness. The evidence suggests that the operational behavior of the dyad after launch is more diagnostic for continuation.

An additional result concerns the difference between structural mismatch and repairable execution failure. The model classifies a subset of launches as having persistent distress even after substantial corrective reallocation. These launches are disproportionately characterized by high asymmetry and deep integration. Their renewal probability is close to zero. In contrast, launches with moderate distress followed by rapid recovery are often moderately integrated and high on discipline. The contrast suggests that the data are capturing two distinct economic phenomena rather than a single continuum of bad luck [35]. Some collaborations were difficult because the market response required

adaptation. Others were difficult because the underlying joint proposition was unstable.

The renewal evidence therefore reinforces the central premise of the paper. Partnership outcomes should not be reduced to average gains. The more consequential question is whether the collaboration can return to a stable path after adverse information arrives and whether that process reveals a partnership worth continuing. Renewal behaves accordingly.

7. Boundary Conditions, Diagnostic Checks, and Alternative Explanations

Several alternative interpretations deserve careful attention. The first is simple selection. Firms that choose staged rollouts or low-overlap governance may also be more disciplined or higher quality in unobserved ways, and those latent traits could drive the results. The paper addresses this in multiple ways. Launch- and firm-level controls absorb observable quality differences. A propensity-weighted subsample balances high-discipline and low-discipline launches on prelaunch characteristics, including category position, prior innovation outcomes, and partner-level commercial strength. In that balanced sample the discipline effect remains negative for distress onset and positive for recovery. The coefficient magnitudes fall somewhat, but the qualitative pattern survives. This suggests that process discipline is not merely standing in for general managerial quality [36].

The second concern is that overlap could proxy for innovation complexity. More novel offerings often require both partners to coordinate more tasks. If such launches are inherently riskier, overlap may simply mark that risk. Two checks mitigate this concern. First, novelty intensity is controlled flexibly using both categorical and continuous measures. Second, the models are re-estimated within the top half of novelty only. Even among relatively novel launches, higher overlap predicts earlier distress and slower recovery. Innovation complexity contributes to risk, but it does not subsume the effect of ambiguous task ownership.

A third concern is measurement error in the hidden-state classification. The distress state is latent and inferred rather than directly observed. To assess robustness, three alternative definitions are examined: a simple threshold rule based on two consecutive months below the 25th percentile of benchmark-adjusted performance, a smoother regime-switching model with t-distributed errors, and a machine-learning classifier trained on retailer support contractions and managerial interventions. The main substantive conclusions remain. Overlap, asymmetry, and crowding increase the incidence of economically meaningful underperformance, while discipline and early corrective action improve resilience. The hidden-state approach is retained because it offers the best combination of interpretability and predictive fit, not because the results depend on a single classification rule [37].

A fourth concern is that corrective reallocation may reflect omitted shocks rather than remediation. For example, an external positive shock could arrive at the same time as a corrective action, creating spurious evidence that the action caused recovery. Several checks argue against this interpretation. The instrumental-variables estimates are stronger, not weaker, than the ordinary estimates. Local projections around distress onset show no anticipatory improvement before correction is implemented. Moreover, placebo corrections assigned to matched nondistressed launches do not predict subsequent abnormal performance. These patterns increase confidence that the measured effect is linked to actual remediation.

The fifth concern is whether the renewal results are driven by mechanical contract length differences. Some partnerships may have been designed from the outset as short experiments, while others were intended to be long-lived. Contract-duration controls are therefore included wherever available, and a restricted sample of launches with explicitly similar initial terms is analyzed separately. Fast repair continues to predict extension in the restricted sample. The economic interpretation of renewal as

posterior belief updating about repair capacity remains plausible.

The analysis also examines several boundary conditions. Category turbulence alters which variables matter most. In highly turbulent categories, launch appeal provides more initial protection and correction must arrive especially early to matter [38]. In less turbulent categories, asymmetry and crowding dominate. Retailer concentration offers another boundary condition. When a small number of retailers account for most category sales, task overlap becomes more damaging because the coordination failure with a single major account can generate large performance swings. In fragmented retail settings, overlap still matters but less severely. These differences are consistent with the commercial mechanism proposed earlier.

Launch form continues to matter, but in a way that differs from simple average comparisons. Composite launches have the lowest distress incidence when asymmetry is low and discipline is high. Under those conditions their deeper integration supports both clearer consumer understanding and richer corrective levers. When asymmetry is high, the same composite structure becomes the most fragile. Endorsement-led launches are less sensitive to asymmetry but more sensitive to crowding, because their symbolic value erodes quickly when the endorsing partner is spread across many collaborations. Ingredient-featured launches sit between these extremes. This pattern suggests that launch form should be evaluated conditionally rather than ranked universally.

A related diagnostic concerns partner role. In some collaborations the primary commercial partner dominates distribution and shelf execution, while the secondary partner mainly contributes symbolic or component value. The paper estimates role-specific versions of overlap and crowding [39]. Overlap on downstream commercial tasks is more damaging when it involves the primary side, whereas crowding is more damaging when it occurs on the symbolically salient side. This asymmetry is informative because it ties each variable to a distinct mechanism. Not all frictions operate through the same part of the dyad.

The possibility of reverse causality between crowding and performance is also considered. A partner with many attractive opportunities may choose to crowd its portfolio precisely because it is successful, and those same capabilities may help recovery. If so, the negative crowding result would be surprising. To test this, historical standalone performance and prior collaboration success are controlled directly. Crowding remains negative after these controls. The implication is that whatever advantages successful partners possess, there is still a real marginal cost to running too many collaborations at once.

An additional falsification uses launches that were planned but never fully commercialized. These near-launch cases provide information on the prelaunch architecture without post-launch scanner realization. Variables such as overlap and discipline do not strongly predict whether the launch is canceled before market introduction once financial distress at the parent level is controlled. Their predictive power appears after commercialization begins, which fits the argument that they matter as response architecture rather than as pure selection into launch.

The paper also tests whether the synthetic benchmark itself biases results against partnerships by constructing alternative donor pools that include partial collaborations and adjacent-category analogues. The main findings are stable [40]. Mean gaps shift slightly, as expected, but the relations among overlap, asymmetry, crowding, discipline, and the dynamic outcomes remain. This stability is important because the argument of the paper is about the shape of underperformance and recovery, not about the exact average counterfactual level.

Finally, a joint estimation of distress, recovery, and renewal with correlated random effects shows that the three outcomes are linked but not redundant. The random-effect correlation between distress incidence and long recovery is positive and sizable. The correlation between fast recovery and

renewal is even stronger. Yet even after allowing for these latent correlations, the observed covariates continue to play distinct roles at each stage. This supports the paper's central modeling choice to separate onset, repair, and continuation rather than collapse them into a single latent success score.

8. Implications for Studying and Managing Partnership Outcomes

The results change how partnership outcomes should be interpreted. In many empirical settings the natural instinct is to summarize a collaboration by average realized value. That summary is attractive because it is simple and comparable across cases. The present evidence suggests that it is often incomplete. Two launches with similar first-year average performance can differ dramatically in their path to that average. One may remain stable throughout. Another may suffer a deep break and recover only after substantial corrective spending and contract revision [41]. Those two paths imply different partner capabilities, different governance lessons, and different prospects for continuation. Outcome research that ignores the path may therefore misclassify the partnership.

A second implication concerns the meaning of governance clarity. Governance is often discussed in broad contractual terms such as exclusivity, scope, or intellectual-property ownership. These are important, but the evidence here points to a more operational dimension: who moves first when performance slips. The most damaging form of ambiguity in the data is not legal uncertainty about who owns the joint asset. It is commercial uncertainty about who owns the first response. Partnerships with extensive joint control over downstream tasks have more frequent and more persistent distress, even when the joint control may have seemed attractive before launch. This finding suggests that outcome research should pay closer attention to procedural ownership and response sequencing.

A third implication concerns the trade-off between integration and asymmetry. Deeply integrated collaborations are not uniformly better or worse. Their performance depends on whether the partners are sufficiently close in commercial power to sustain a shared operating logic. When that condition holds, deep integration lowers distress and improves recoverability. When it does not, deep integration exposes and amplifies mismatch. This conditional pattern provides a richer interpretation of partnership form [42]. The appropriate level of integration is not determined in the abstract; it depends on the dyad's asymmetry and its ability to absorb negative information without immediate blame shifting.

A fourth implication concerns portfolio design. The results make a distinction between learning from repeated collaboration and overloading the portfolio. Firms can gain from partnership experience, but those gains are not boundless. Once external collaborations become numerous and commercially broad, the marginal collaboration suffers both from diluted managerial attention and from weakened distinctiveness. For managers this means that the productive frontier of partnerships is unlikely to be monotonic in count. There is an interior region in which experience is beneficial and a crowded region in which additional collaborations reduce resilience. For researchers this means that alliance count alone is an insufficient variable. The distribution of attention across simultaneously active collaborations matters.

The evidence on fast repair also has conceptual implications. Renewal depends heavily on whether the first major shortfall was addressed quickly and credibly. This suggests that partners learn about one another most sharply under adversity. A collaboration that never faces a meaningful negative deviation may not reveal much about the dyad's repair capacity. By contrast, a collaboration that faces one and responds well creates information that supports future cooperation. Outcome research could therefore benefit from treating adverse episodes as informative events rather than as nuisance volatility [43].

A final implication concerns measurement. Distress is often hidden in standard quarterly or annual data. Monthly data, benchmarked residuals, and latent-state methods reveal the short operating episodes that matter for whether the collaboration survives. This is not only a technical point. It means that some partnerships appearing moderately successful in annual averages are in fact expensive to manage and unlikely to be renewed, while others appearing ordinary in average terms are strategically valuable because they demonstrate exceptional recoverability. Partnership outcomes are better understood when the time path of deviation is observed with enough resolution to identify these differences.

9. Conclusion

This paper examined partnership outcomes through the dynamics of post-launch instability rather than through average performance alone. Using a large monthly panel of commercial collaborations, the analysis showed that partnerships differ systematically in their probability of entering a distressed operating state, in the speed with which they recover from that state, and in whether the relationship is renewed once the initial term ends. The main sources of fragility are overlapping downstream control, substantial commercial asymmetry, and crowded external collaboration portfolios. The main sources of resilience are rollout discipline, categories with usable corrective levers, and rapid post-shock reallocation of support. Deep integration is beneficial only within a limited range of partner asymmetry.

The broader contribution is to recast partnership outcomes as a sequence of state transitions and repair decisions. A collaboration is not defined only by how much value it appears to generate on average. It is also defined by how it behaves when realized demand falls short, whether the partners can reestablish a stable path quickly, and whether that recovery changes the perceived viability of continued cooperation. This view helps explain why superficially similar partnerships often diverge sharply in later continuation and why the most informative evidence about a dyad may emerge only after the first disappointment occurs [44].

References

[1] Tülin Erdem and Joffre Swait. “The Information-Economics Perspective on Brand Equity”. In: *Foundations and Trends® in Marketing* 10.1 (Aug. 31, 2016), pp. 1–59. DOI: 10.1561/1700000041.

[2] Yuke Rahmawati. “Penilaian Kinerja Badan Arbitrase Pasar Modal Indonesia Dengan Metode Total Quality Management (TQM)”. In: *JURNAL CITA HUKUM* 4.2 (Dec. 2, 2016), pp. 96398–. DOI: 10.15408/jch.v4i2.3671.

[3] Arend J. Groen and Lenna K.M. Lee. “Nation brand management : Towards a convergent theory for nation branding”. In: *Archives of Design Research* 26.2 (May 31, 2013), pp. 55–69. DOI: 10.15187/adr.2013.05.26.2.55.

[4] Anna Fyrberg. “The Olympic Equestrian Games: brand collaboration and associations within a destination and a sports event”. In: *International Journal of Sports Marketing & Sponsorship* 9.4 (July 1, 2008), pp. 307–320.

[5] Sunhye Choi, Yoonjung Kim, Jihye Lee, Mina Kashiwabara, and Kyungwon Oh. “Tobacco Use Among Students Aged 13-15 Years in South Korea: The 2013 Global Youth Tobacco Survey”. In: *Journal of preventive medicine and public health = Yebang Uihakhoe chi* 50.1 (Jan. 31, 2017), pp. 60–65. DOI: 10.3961/jpmph.16.079.

[6] “Subject index”. In: *Tobacco Control* 12.1 (Mar. 1, 2003), e3–e3. DOI: 10.1136/tc.12.1.e3.

[7] Zixia Cao. “An empirical examination of stock market reactions to introduction of co-branded products”. PhD thesis. 2012.

[8] Pierre Chandon and Brian Wansink. "Does food marketing need to make us fat? a review and solutions". In: *Nutrition reviews* 70.10 (Oct. 4, 2012), pp. 571–593. DOI: 10.1111/j.1753-4887.2012.00518.x.

[9] Ruiliang Yan and Zixia Cao. "Is brand alliance always beneficial to firms?" In: *Journal of Retailing and Consumer Services* 34 (2017), pp. 193–200.

[10] Ana Sofia Aires Belmonte. "Relationship marketing: como uma estratégia de co-branding pode ser relevante numa abordagem relacional ao mercado — o caso Nike+". In: *Comunicação pública* Vol.4 nº8 / nº7 (Dec. 31, 2009), pp. 73–95. DOI: 10.4000/cp.7701.

[11] PEI-CHUN LIN and CHIA-HUI LEE. "How Online Vendors Select Parcel Delivery Carriers". In: *Transportation Journal* 48.3 (July 1, 2009), pp. 20–31. DOI: 10.5325/transportationj.48.3.0020.

[12] "Co-branding as a new competitive weapon". In: *Strategic Direction* 18.10 (Nov. 1, 2002), pp. 31–33. DOI: 10.1108/02580540210794058.

[13] Harold Salzman, Signe-Mary McKernan, Nancy Pindus, and Rosa Marie Castaneda. "Capital Access for Women: Profile and Analysis of U.S. Best Practice Programs". In: *SSRN Electronic Journal* (Jan. 1, 2006). DOI: 10.2139/ssrn.968715.

[14] Randi Priluck Grossman. "Co-branding in advertising: developing effective associations". In: *Journal of Product & Brand Management* 6.3 (June 1, 1997), pp. 191–201. DOI: 10.1108/10610429710175709.

[15] Vanitha Swaminathan, Srinivas K. Reddy, and Sara Loughran Dommer. "Spillover effects of ingredient branded strategies on brand choice: A field study". In: *Marketing Letters* 23.1 (Oct. 21, 2011), pp. 237–251. DOI: 10.1007/s11002-011-9150-5.

[16] HaeEun Helen Chun, C. Whan Park, Andreas B. Eisingerich, and Deborah J. MacInnis. "Strategic benefits of low fit brand extensions: When and why?" In: *Journal of Consumer Psychology* 25.4 (Dec. 12, 2014), pp. 577–595. DOI: 10.1016/j.jcps.2014.12.003.

[17] Crystale Purvis Cooper, Cynthia A. Gelb, Jennifer Chu, and Lindsey Polonec. "Can Donated Media Placements Reach Intended Audiences". In: *Health promotion practice* 14.5 (May 29, 2013), pp. 656–662. DOI: 10.1177/1524839913488891.

[18] Shu-Hsien Liao, Yin-Ju Chen, and Hsiao-Wei Yang. "MINING CUSTOMER KNOWLEDGE FOR CHANNEL AND PRODUCT SEGMENTATION". In: *Applied Artificial Intelligence* 27.7 (Aug. 9, 2013), pp. 635–655. DOI: 10.1080/08839514.2013.813195.

[19] Zixia Cao and Alina Sorescu. "Wedded bliss or tainted love? Stock market reactions to the introduction of cobranded products". In: *Marketing Science* 32.6 (2013), pp. 939–959.

[20] Judith Lee Smith et al. "Assessment of the status of A National Action Plan for Cancer Survivorship in the USA". In: *Journal of cancer survivorship : research and practice* 7.3 (Apr. 23, 2013), pp. 425–438. DOI: 10.1007/s11764-013-0276-8.

[21] Ruminda Ruminda. "SEMANTIC CATEGORIES OF REPORTING VERBS IN ONLINE NEWS ARTICLES". In: *Al-Tsaqafa: Jurnal Ilmiah Peradaban Islam* 16.1 (June 20, 2016), pp. 21–34. DOI: 10.15575/al-tsaqafa.v13i01.1828.

[22] Sonja Oklobdzija. "THE ROLE OF EVENTS IN TOURISM DEVELOPMENT". In: *BizInfo Blace* 6.2 (Dec. 29, 2015), pp. 83–97. DOI: 10.5937/bizinfo1502083o.

[23] Andrea Rea and Simona D'Antone. "La sistemicità presupposto del valore della marca territoriale: un'analisi sul mondo del vino made in Italy". In: *Sinergie Italian Journal of Management* 28.83 (Aug. 11, 2011), pp. 179–200.

[24] Pierre Berthon, Leyland Pitt, Michael Parent, and Jean-Paul Berthon. "Aesthetics and ephemerality: Observing and preserving the luxury brand". In: *California Management Review* 52.1 (Nov. 1, 2009), pp. 45–66. DOI: 10.1525/cmr.2009.52.1.45.

[25] Zixia Cao and Ruiliang Yan. "Does brand partnership create a happy marriage? The role of brand value on brand alliance outcomes of partners". In: *Industrial Marketing Management* 67 (2017), pp. 148–157.

[26] Giacomo Cavalli and Renato Paro. "The Drosophila Fab-7 Chromosomal Element Conveys Epigenetic Inheritance during Mitosis and Meiosis". In: *Cell* 93.4 (May 15, 1998), pp. 505–518. DOI: 10.1016/s0092-8674(00)81181-2.

[27] Marilyn G. Wiebe. "Myco-protein from Fusarium venenatum: a well-established product for human consumption." In: *Applied microbiology and biotechnology* 58.4 (Feb. 8, 2002), pp. 421–427. DOI: 10.1007/s00253-002-0931-x.

[28] Jonathan A.J. Wilson. "Global business discourse - reflections from a cultural nomad". In: *Journal of Emerging Economies and Islamic Research* 1.2 (May 31, 2013), pp. 1–. DOI: 10.24191/jeeir.v1i2.9119.

[29] "The Outcomes of the Co-branding Strategy". In: *Chinese Business Review* 11.9 (Sept. 28, 2012). DOI: 10.17265/1537-1506/2012.09.006.

[30] A. R. Tieftrunk, Karl Jacobs, Christopher L. Martin, O. Siebertz, Antony A. Stark, J. Stutzki, Christopher K. Walker, and Gregory Wright. "^{13}C I in high-mass star-forming clouds". In: *Astronomy & Astrophysics* 375.2 (Aug. 15, 2001), pp. L23–L26. DOI: 10.1051/0004-6361:20010939.

[31] Shaun M. Powell. "Journal of Brand Management: year end review 2017". In: *Journal of Brand Management* 24.6 (Nov. 10, 2017), pp. 509–515. DOI: 10.1057/s41262-017-0078-9.

[32] Stan Maklan and Simon Knox. "Dynamic capabilities: the missing link in CRM investments". In: *European Journal of Marketing* 43.11/12 (Nov. 13, 2009), pp. 1392–1410. DOI: 10.1108/03090560910989957.

[33] Bashar S. Gammoh, Kevin E. Voss, and Xiang Fang. "Multiple brand alliances: a portfolio diversification perspective". In: *Journal of Product & Brand Management* 19.1 (Mar. 2, 2010), pp. 39–39.

[34] Dhione Oliveira Santana and Raquel Marques Carriço Ferreira. "Licenciamento Autoral e o Filme "Os Vingadores: The Avengers"". In: *Anagrama* 6.4 (Apr. 6, 2013), pp. 1–13. DOI: 10.11606/issn.1982-1689.anagrama.2013.56343.

[35] Janne Parkkila, Kati Järvi, Timo Hynninen, Jouni Ikonen, and Jari Porras. "Gamecloud - A Platform for Connecting Video Games". In: *Journal For Virtual Worlds Research* 9.1 (Apr. 11, 2016). DOI: 10.4101/jvwr.v9i1.7196.

[36] Padmanabh B. "flipkart.com's Strategy on Amazon.com's Entry into India :A Case Study". In: *Ushus - Journal of Business Management* 12.4 (Sept. 9, 2013), pp. 131–143. DOI: 10.12725/ujbm.25.8.

[37] Iwan Setiawan, Sumardjo Sumardjo, Arif Satria, and Prabowo Tjitropranoto. "Strategi Pengembangan Kemandirian Pelaku Muda Agribisnis " Brain Gain Actors" di Jawa Barat". In: *MIMBAR, Jurnal Sosial dan Pembangunan* 31.2 (Dec. 23, 2015), pp. 409–418. DOI: 10.29313/mimbar.v31i2.1491.

[38] Denise Duclaux. "Has the Cobranded Boom Gone Bust". In: *ABA Banking Journal* 88.7 (July 1, 1996), pp. 69–.

[39] Peter H. Farquhar. "The mysteries of branding revealed". In: *Journal of Brand Management* 7.1 (Sept. 1, 1999), pp. 4–7. DOI: 10.1057/bm.1999.34.

[40] Donghun Lee, Michael Cottingham, Demetrius W. Pearson, Soon-Ho Kim, and Jungkun Park. "Collaborative strategy in sports industry: team co-branding". In: *The Service Industries Journal* 36.11-12 (Sept. 9, 2016), pp. 595–613. DOI: 10.1080/02642069.2016.1255727.

[41] John C. Marshall. "More than meets the eye". In: *Nature* 396.6706 (Nov. 5, 1998), pp. 36–36. DOI: 10.1038/23859.

[42] Rubén Huertas-García, Jorge Lengler, and Carolina Consolacion-Segura. "Co-branding strategy in cause-related advertising: the fit between brand and cause". In: *Journal of Product & Brand Management* 26.2 (Apr. 18, 2017), pp. 135–150. DOI: 10.1108/jpbm-07-2015-0939.

[43] Jung Ok Jeon and Sunmee Baeck. "The effect of the valence of word-of-mouth on consumers' attitudes toward co-brands: The moderating roles of brand characteristics". In: *Journal of*

Global Scholars of Marketing Science 26.1 (Jan. 2, 2016), pp. 89–108. DOI: 10.1080/21639159.2015.1116781.

[44] Natasha Zhang Foutz. "Entertainment Marketing". In: *Foundations and Trends® in Marketing* 10.4 (Oct. 16, 2017), pp. 215–333. DOI: 10.1561/1700000049.