Collective Brand Reputation

Volker Nocke (Mannheim University)
Roland Strausz (HU Berlin)

Discussion Paper No. 324

April 19, 2022

Collective Brand Reputation*

Volker Nocke† Roland Strausz‡

April 15, 2022

Abstract

We develop a theory of collective brand reputation for markets in which product quality is jointly determined by local and global players. In a repeated game of imperfect public monitoring, we model collective branding as an aggregation of quality signals generated in different markets. Such aggregation yields a beneficial informativeness effect for incentivizing the global player. It however also induces harmful free-riding by local, market-specific players. The resulting tradeoff yields a theory of optimal brand size and revenue sharing that applies to platform markets, franchising, licensing, umbrella branding, and firms with team production.

Keywords: Collective branding, reputation, free-riding, repeated games, imperfect monitoring.

*We thank Heski Bar-Issac, Helmut Bester, Martin Burkhardt, Jay Pil Choi, Paul Heidhues, Johannes Hörner, Daniel Krähmer, George Mailath, Moritz Meyer-ter-Vehn, Benny Moldovanu, Aniko Öry, Martin Peitz, Markus Reisinger, and Juuso Välimäki for helpful comments. In addition, we thank three anonymous referees and the editor, and also thank seminar audiences at HU Berlin / Bonn, Edinburgh, Paris-Dauphine, NUS, TSE, and the CEPR Virtual IO Seminar as well as participants at the 1st MaCCI-EPoS-CREST Industrial Organization Workshop (Mannheim), the 2019 Meeting of the *Industrieökonomischer Ausschuss* (Bern), the 5th Workshop on Relational Contracts (Madrid), the 12th Workshop on the Economics of Advertising and Marketing (Porto), the 1st Cornell University Applied Theory Conference, the 3rd University of Bergamo Workshop on Advances in Industrial Organization, the 2019 Asia-Pacific Industrial Organization Conference (Tokyo), the 2020 North American Winter Meeting of the Econometric Society (San Diego), the 2021 MaCCI Summer Institute in Competition Policy (Annweiler) and the 2021 Chile Workshop on Industrial Organization and Economic Theory. We gratefully acknowledge financial support from the German Research Foundation (DFG) through CRC TR 224 (Project B03) [Nocke] and CRC TR 190 (Project B02) [Strausz].

†University of Mannheim, University of Surrey, and CEPR. Email: volker.nocke@gmail.com.

‡Humboldt Universität zu Berlin and CEPR. Email: roland.strausz@hu-berlin.de.

1 Introduction

Most products are experience goods in that product quality is not directly observable prior to consumption. Consumers therefore base their purchasing decisions foremost on the reputation of brands, formed through past consumption experiences, word of mouth, internet reviews, and the like, and facilitated by trade-marks and logos. For a firm's long-term success, managing its brand reputation is therefore crucial.[1] Controlling quality at the brand level is, however, challenging because it is a coarse measure that, in general, depends on the collective decisions of different types of agents. For instance, platform brands such as Amazon, Ebay, and Uber, or franchise brands such as McDonald's and Burger King have some control over consumers' experience but, to a large extent, this experience also depends on the individual sellers, local delivery services and drivers, and outlets, who ultimately service the consumer locally.[2] Similarly, the quality of manufacturing goods such as cars or beers depends not only on the decisions taken by the headquarter but also on those by local plant managers.

A crucial feature of brand reputation is, therefore, that it is a collective outcome of actions taken not only by "global" players, who impact the quality of the entire line of products, but also "local" players, each of whom affects quality of only a subset of products. The fact that collective branding involves an aggregation of quality signals about the actions of both global and local players, raises a number of important economic questions. In particular, what are the main economic forces affecting whether a collective brand reputation is sustainable, and when would it be better to sell the different products under different brand names? For instance, is it better to sell four different products under a single, two, or four independent brand names?[3] More generally, what are the economic tradeoffs that affect the optimal size of the collective brand? Moreover, what instruments are key for controlling a collective reputation and how large are the remaining inefficiencies under their optimal use?

To address these and related questions, we develop a framework to study collective brand reputation. More specifically, we analyze an infinitely repeated hidden action game of imper-

[1] Brand Finance, which produces annual rankings of companies based on brand/intangible value, reports that, on average, the total intangible value of a global top ten company accounts for over 85 percent of such a firm's market value (https://www.visualcapitalist.com/intangible-assets-driver-company-value/, last retrieved on 02/20/2022). Similarly, Ocean Tomo, a financial expert and management consulting company, reports that, in the last decades, intangible assets have become ever more important, and are now responsible for 90% of all value of S&P 500 businesses. (https://www.oceantomo.com/INTANGIBLE-ASSET-MARKET-VALUE-STUDY, last retrieved on 02/20/2022).

[2] Nosko and Tadelis (2015) document such problems in platform markets. Blair and Lafontaine (2005) highlight the franchisor's problem of maintaining consistent quality across franchisees, and conclude that the empirical evidence "seems consistent with the concept that free riding, or individual profit maximization by opportunistic franchisees, is an issue in franchised chains" (p.137).

[3] To give a concrete example, Procter & Gamble is a multinational consumer goods corporation that sells its products under many different brand names, whereas Virgin Group Ltd. explicitly engages in umbrella branding, purposely selling their products and services under the same brand name.

fect public monitoring. There are n markets. In each market, two long-lived players jointly produce a good over an infinite number of periods. One of these players – the local player – is active in only one market, whereas the other player – the global player – is active in all n markets. The good is of high quality in a given market only if both players in that market exert effort. Neither effort nor quality are observable but, at the beginning of each period and for each market, a noisy (binary) signal of last period's quality is realized.

Following the key tenet of marketing that consumers identify quality with the reputation of the good's brand,[4] the n goods can either be sold under independent branding or collective branding. Under independent branding, the goods are sold under n different brand names, leading to a market separation in which the market-specific consumer conditions her buying behavior only on the quality signals associated with the products in her market. By contrast, under collective branding, the products are all sold under a single brand and quality signals across all markets are aggregated into a common signal. As a result, behavior in each market depends only on the pooled quality signals from all markets rather than the individual signals.

Comparing the Perfect Public Equilibrium (PPE) outcomes of the repeated games that result under independent and collective branding, we identify the following novel trade-off. On the one hand, collective branding yields an *informativeness effect* that allows to better control the incentives of the global player. On the other hand, collective branding yields a *free-rider effect* that makes it harder to control the incentives of local players. Hence, with only global players, collective branding is optimal, because only the informativeness effect arises. As a result, inefficiencies vanish when both the discount factor approaches 1 and the brand size goes to infinity. In contrast, with only local players, the free-rider effect renders independent branding optimal. Moreover, due to imperfect monitoring the folk theorem fails: a "discountless" inefficiency—bounded away from zero—remains when the discount factor approaches 1. With both global and local players, the optimal branding choice depends on the trade-off between the two effects.

This trade-off yields a theory that allows us to answer our questions concerning optimal brand size, the crucial economic instruments to sustain a collective reputation, and the remaining inefficiencies if frictions such as discounting vanish. To understand this trade-off, recall that monitoring in our setup is imperfect. As a result, high effort in the beginning of the repeated game is sustainable only in an equilibrium in which, on the equilibrium path, the long-lived players switch to low effort with a strictly positive probability after an (erroneously) bad signal. We call such a probabilistic switch to low quality a *market breakdown* of high-quality production. The optimal brand size is then the one that sustains a *high-quality equilibrium* with the smallest on-path probability of market breakdown.

[4]See, for example, the textbook by Kotler (2003, p.420): "A brand is essentially a marketer's promise to deliver a specific set of features, benefits, and services consistently to the buyers."

Under collective branding, the sustainability of a high-quality equilibrium depends crucially on how local and global players share their revenues. Hence, this insight identifies a first crucial instrument for sustaining a collective reputation: a careful *calibration of revenue sharing*. In particular, under independent branding, the smallest on-path probability of market breakdown obtains for "proportional revenue sharing", where each player obtains a share of the revenue that equals his share in the overall effort cost of producing high quality. If such proportional revenue sharing were used under collective branding, independent branding would be superior—due to the free-rider effect. For collective branding to outperform independent branding requires that the local players obtain more than their proportional share.

We characterize the optimal calibration of revenue sharing under collective branding. Moreover, we show how, under such optimal revenue sharing, the minimal on-path probability of market breakdown depends on the firm's choice of brand size and the underlying economic fundamentals. In the above-mentioned example of four products, our results on the optimal brand size imply the following: for intermediate discount factors, it is optimal to sell four goods under two separate brands of size two each; for small discount factors, four separate brands are optimal; and for large discount factors, a single brand is optimal. More generally, collective branding allows a high-quality equilibrium with a lower on-path breakdown probability than independent branding for discount factors that are sufficiently close to 1. For any such discount factor, there is, however, a finite upper bound on the brand size that allows a high-quality equilibrium, and so the optimal brand size is finite. This reveals a second crucial instrument for sustaining a collective reputation: the *size of the collective brand* itself. As the discount factor becomes large, both this maximal sustainable and the optimal finite brand size increase without bound.

However, in the limit as both the discount factor and the optimal brand size become large, the inefficiency from imperfect monitoring does not vanish. The remaining inefficiency is equal to the relative importance of local versus global players, and therefore intimately connected to the trade-off between the informativeness and free-rider effects. In particular, the remaining inefficiency equals the relative importance of the local players, as measured by the share of effort cost they bear, multiplied by the aforementioned discountless inefficiency under independent branding with only local players. Hence, with optimal revenue sharing, it is as if, in this discountless limit, one achieves the best of both worlds: collective branding for the global player without any inefficiencies and independent branding for the local players with its discountless inefficiency due to a failure of the folk theorem.

Related literature. The unique feature of our paper is the analysis of collective reputation in the presence of both global and local players.[5] The key tradeoff underlying our results is between the informativeness effect (which is beneficial in the presence of a global player) and the free-rider effect (which is detrimental in the presence of local players). While the economics literature has studied variants of these two effects in isolation, our paper is the first to analyze their interactions and economic implications.

Studying cooperation in a repeated prisoner's dilemma game with imperfect monitoring, Matsushima (2001) was, to our knowledge, the first to identify the beneficial informativeness effect that underlies our paper. Cabral (2009) shows that this informativeness effect may render umbrella branding optimal, while Cai and Obara (2009) study it as a driver of horizontal integration. These papers effectively only consider global players, thereby abstracting from free riding on a collective reputation.

By contrast, Tirole (1996), Fishman et al. (2018), and Neeman et al. (2019) study free-rider problems associated with a collective reputation in settings with only local players.[6] These papers however do not consider our beneficial informativeness effect of collective reputation, and therefore address different economic forces. Considering a repeated matching environment with overlapping generations, Tirole (1996) shows that a collective reputation may lead to a persistent stigmatization of new generations due to shirking by some earlier generation. In his framework, there are no inherent benefits from a collective reputation. In a two-period model with persistent investment and different types of firms, Fishman et al. (2018) study free-riding on a collective brand reputation, but with the benefit that the collective brand can select its members based on their investment decisions and/or types. Neeman et al. (2019) point out that a collective reputation may serve as a commitment device. Trading off this effect against the free-rider effect, they therefore study a different trade-off from ours.

Focusing on brand management, our study is related to the literature on co-branding, brand extension, and umbrella branding (e.g., Kotler, 2003). With respect to this extensive literature, our contribution is to study potential free-rider problems which, in our view, are endemic to such settings.[7] While most work in this literature focuses on reputation models with hidden information (e.g., Wernerfelt, 1988, Choi, 1998, Cabral, 2000, Miklos-Thal 2012, and Moorthy, 2012), our work is more closely related to studies that analyze umbrella branding in a moral hazard framework. Hence, our modelling of reputation follows the one pioneered in Klein and Leffler (1981) rather than the type-based approach developed

[5]See Bar-Isaac and Tadelis (2008) for a survey of the literature on seller reputation.

[6]Winfree and McCluskey (2005) and Fleckinger (2014) also study collective reputation, but in a fundamentally different framework in which consumers observe the product's (collective) quality at the time of purchase.

[7]See Castriota and Delmastro (2012) for an empirical study of the importance of collective reputation in wine markets.

in Kreps and Wilson (1982) and Milgrom and Roberts (1982).

Building on Klein and Leffler (1981), Andersson (2002) shows that in a repeated game of moral hazard but perfect monitoring, a single brand name that pools the reputation across independent markets is helpful only if markets display asymmetries. Hakenes and Peitz (2008) and Cabral (2009) highlight that, with imperfect monitoring, pooling reputation can be beneficial even when markets are symmetric.

Moreover, our paper contributes to the literature on the management of moral hazard in team production, pioneered by Holmström (1982). In our setup, there are two types of team production: physical team production within a market and reputational team production across markets. Since the focus of our analysis is on the reputational team production problem, we assume that the market-specific effort choices of the local and global players are perfectly complementary so that these players can fully resolve their physical team production problem. That is, the outcome under independent branding is identical to the one that would obtain if the global and local player were to vertically integrate.

Our results also provide insights into the classical quality management problem in franchising and licensing. In franchising (licensing), we can view the franchisor (licensor) as a global player, while the outlets (licensees) are local players. Practitioners and legal scholars have pointed out the importance of free-rider problems in these contexts. For instance, Hadfield (p.949, 1990) notes that the individual "franchisee is inclined to make decisions about how much effort to put into the business based on the profits that will accrue directly to her in her own outlet," whereas "customers make judgments about the quality of the entire franchise system based on their experience at an outlet".[8] Similarly, Klein and Saft (p.349ff, 1985) remark that the "franchise arrangements create an incentive for franchisees to shirk on quality," further pointing out that "the individual franchisee directly benefits from the sales of the low quality product, and the other franchisees share in the losses caused by decreased future demand".[9] In the context of trademark licensing, Calboli (2007) explains that, legally, "trademarks are protected only as conveyors of information about the products which they identify and as symbols of commercial goodwill" (p.357) and points out the free-rider prob-

[8]A concrete example is the Burger King scandal in Germany in 2014. After an undercover report exposed severe problems of poor hygiene in outlets in Cologne, Burger King tried to put the blame on the individual outlets but German consumers associated the negative report with the Burger King brand as a whole rather than its local franchisee in Cologne. Similarly, in Kentucky Fried Chicken Corp. v. Diversified Packaging Corp., 549 F.2d 368, 380 (1977), the Court observed "A customer dissatisfied with one Kentucky Fried outlet is unlikely to limit his or her adverse reaction to the particular outlet; instead, the adverse reaction will likely be directed to all Kentucky Fried stores. The quality of a franchisee's product thus undoubtedly affects Kentucky Fried's reputation and its future success."

[9]In the context of licensing, the court in Siegel v. Chicken Delight, 448 F.2d 43, n.38 (1971) observed that "the licensor owes an affirmative duty to the public to assure that in the hands of his licensees the trade-name continues to represent that which it purports to represent." Klein and Saft (p.349ff, 1985) interpret this view as expressing "a legal obligation for quality maintenance in a system involving many producers operating under a common trade name."

lem that licensees' "lack of direct ownership of the mark could make them less interested in the long-term success of the products" (p.360).

To our knowledge, our paper is the first to formally model, and rigorously analyze, this classical problem.[10] Our result that, without properly calibrating revenue shares, a collective reputation destroys any benefits from pooling reputations confirms that due to free-riding "the value of the trademark will suffer dramatically" (Hadfield, 1980). Yet, our results also show that a franchisor can partially mitigate free-rider problems and thereby maintain the trademark's value by shifting revenue streams from himself to the franchisees. This reinforces the insight of Bhattacharyya and Lafontaine (1995) that revenue sharing is crucial for controlling double moral hazard problems in franchising in that free-rider problems associated with a collective reputation are key factors in the determination of optimal revenue shares.

Plan of the paper. In the next section, we present the model. This is followed, in Section 3, by the equilibrium analysis of independent branding. In Section 4, we first study the polar case of collective branding in which the burden of effort is borne entirely by the global player. We then turn to the other polar case in which all of the burden of effort is borne by the local player. The analysis of these polar cases is instructive for analyzing the generic case in which local and global players share the effort cost. Section 5 addresses the comparative statics in the brand size n and obtains results concerning the maximum implementable brand size, $\overline{n}$, the optimal brand size, $\hat{n}$, and (in)efficiency results for limiting cases. We conclude in Section 6. We collect all proofs in Appendix A.[11]

2 The Model

We consider an infinitely repeated game of imperfect public monitoring in discrete time $t = 0, 1, \ldots$. There are $n \geq 2$ symmetric markets, indexed by $i = \{1, \ldots, n\}$ with one long-lived global player, G, and n long-lived local players, L_i. For each period t, production in a market i requires the market-specific binary input, $e^t_{G,i} \in \{0, 1\}$, of the global player G and the binary input, $e^t_{L,i} \in \{0, 1\}$, of the market-specific local player L_i.[12] The good produced

[10] Extensively discussing the free-rider problem in franchising, Blair and Lafontaine (2005) capture the bare essentials of this problem in a highly stylized model that abstracts from any reputational concerns.

[11] By applying the abstract methods of decomposability and self-generation developed in Abreu, Pearce, and Stacchetti (1990), we study, in Appendix B, asymmetric PPE in the case of $n = 2$ markets, addressing the robustness of our results to asymmetric equilibrium outcomes.

[12] That is, we assume that the global player can choose different effort levels in different local markets. While this "independent effort choice" assumption is the appropriate one for some applications (think of the global player delivering meat to hamburger outlets as discussed in footnote 8), for other applications it makes more sense that the global player has to choose a common effort level in all local markets (think of a global advertising effort). While we focus our formal analysis on the (more stringent) independent effort choice case, we discuss in the conclusion that all of our propositions extend to the common effort level case.

in market i is sold to a (representative) market-specific short-lived consumer C_i.

Production technology. The quality q_i^t of good $i \in \{1, \ldots, n\}$ in period $t \in \{0, 1, 2, ..\}$ is either high, $q_i^t = 1$, or low, $q_i^t = 0$, and depends on the simultaneous effort choices of G and L_i. In particular, it is equal to one if and only if both G and L_i put in effort and zero otherwise: $q_i^t = e_{G,i}^t \cdot e_{L,i}^t$. The aggregate cost of effort for producing high quality in a specific market is $c > 0$, of which G incurs the share λ_G and L_i incurs the remaining share $\lambda_L = 1 - \lambda_G$. That is, λ_G represents the importance of the global player's effort cost relative to that of the local player: G's effort cost is $e_{G,i}^t \lambda_G c$ and L_i's is $e_{L,i}^t \lambda_L c$.

Timing. The infinitely repeated game starts after the long-lived players set, for each good i, the revenue shares π_G and $\pi_L = 1 - \pi_G$ that accrue to G and L_i, respectively. At the beginning of each period $t \geq 1$, before effort choices are made, there is a binary signal $s_i^t \in \{0, 1\}$, providing noisy information about the good's quality in the previous period.[13] As we formalize below, the branding decision determines the (public) observability of these signals. The realization of the market-specific signal s_i^t depends on the previous period as follows: if $q_i^{t-1} = 1$, then $s_i^t = 1$ with probability $1 - \alpha$; similarly, if $q_i^{t-1} = 0$, then $s_i^t = 0$ with probability $1 - \beta$. The parameters $\alpha \in (0, 1)$ and $\beta \in (0, 1 - \alpha)$, measure the noisiness of the signal and represent the probabilities of type II and type I errors, respectively. Before effort choices are made, there is also the realization of an independent public randomization device $r^t \in [0, 1]$, uniformly distributed over the unit interval, and independent over time.[14] After the effort choices have been made, the consumer in market i, C_i, decides whether to buy good i ($b_i^t = 1$) or not ($b_i^t = 0$).

Payoffs. In market i, the consumer's valuation equals the good's quality level q_i^t. As we focus on high-quality equilibria, we fix the price of the good to 1.[15] As agreed upon in an initial stage prior to the repeated game, the global player G and local player L_i receive shares π_G and $\pi_L = 1 - \pi_G$ of the revenue from selling good i. Assuming that the signal s_i^t is non-contractible, these shares are independent of the signal realization and, focusing on symmetric equilibria, uniform across markets. We discuss the feasibility of more elaborate schemes of revenue sharing under different modeling assumptions in the conclusion.

[13] For notational convenience, we set $s_i^0 = 1$.

[14] The public randomization device simplifies the exposition of our results; none of our results require the existence of such a signal. Footnote 19 makes this explicit.

[15] In our formal modelling of the repeated game, we treat this price as fully exogenous. Equivalently, we could have assumed—as is commonly done in the literature on umbrella branding—that the price is equal to consumers' willingness to pay for the good, e.g., because multiple (identical) consumers bid for the good in a (second-price) auction.

A natural division of the revenue is to set a player's reward share equal to his cost share, $(\pi_G, \pi_L) = (\lambda_G, \lambda_L)$. We refer to this sharing rule as *proportional rewards*. For the case in which costly effort is needed from both the global and local players, $\lambda_G, \lambda_L > 0$, we define the *reward-to-cost-share ratio* of player j as $\gamma_j \equiv \pi_j/\lambda_j$. Governed by the accounting identity $\lambda_G\gamma_G + \lambda_L\gamma_L = 1$, there is a one-to-one relationship between the reward shares (π_G, π_L) and the reward-to-cost-share ratios (γ_G, γ_L). In the particular case of proportional rewards, we have $\gamma_G = \gamma_L = 1$.

Summarizing, the period-t profit of a long-lived player $k \in \{G, L_i\}$ in market i is equal to

$$b_i^t \pi_k - e_{k,i}^t \lambda_k c.$$

The long-lived players discount profits with factor $\delta \in (0,1)$. The payoff of the (short-lived) consumer in period t and market i is given by

$$(q_i^t - 1)b_i^t = (e_{G,i}^t e_{L,i}^t - 1)b_i^t.$$

While we view the revenue shares (π_G, π_L) as the outcome of a bargaining process between the long-lived players prior to the repeated game, we do not model this bargaining stage explicitly. As bargaining takes place under full information, we do assume however that it results in an efficient outcome, so that the revenue shares maximize the joint value of the long-lived players. We therefore seek the perfect public equilibrium (PPE) that maximizes the joint value of G and L_i,[16]

$$V_i \equiv \sum_{t=0}^{\infty} \delta^t \left[b_i^t - (e_{G,i}^t \lambda_G + e_{L,i}^t \lambda_L) c \right],$$

and is strongly symmetric in the sense that all n local players use identical strategies after every history. Let $\overline{V}_i$ and $\underline{V}_i$ denote the maximal and minimal values of V_i, respectively, that can be sustained in a PPE.

Public histories. We model the distinction between independent and collective branding purely as differences in public information concerning the signals s. With independent brands, the public signal in market i is the market-specific signal s_i together with the randomization device r. Consequently, the public history, h_i^t, in market i at time t is

$$h_i^t = (s_i^\tau, r^\tau)_{\tau=0,\ldots,t}.$$

[16]As consumer surplus is equal to zero in any equilibrium, this PPE also maximizes the discounted sum of aggregate surplus.

By contrast, the public signal under collective branding consists only of the aggregate signal $\tilde{s}^\tau = \sum_i s_i^\tau$—the number of positive realizations of the n noisy quality signals—together with the randomization device r.[17] Consequently, the public history, h^t, in market i at time t is

$$h^t = (\tilde{s}^\tau, r^\tau)_{\tau=0,\dots,t}.$$

The difference in public histories between independent and collective branding captures the key tenet of marketing: consumers identify the quality of a good through its brand name alone. In particular, under collective branding, consumers cannot discern information about the good's quality that is market specific.[18] The public history can therefore contain only aggregate signals of identically branded products. Our focus on perfect public equilibrium (PPE) then implies that we study only behavior in which players condition their strategies on the coarse brand-specific public history rather than any finer information. This also means that, under independent branding, players' strategies in market i are independent of the quality signals in another market $j \neq i$.

Before proceeding, we formally define our equilibrium concept for both independent and collective branding. Under independent branding, players view the markets as independent, and we therefore consider the PPE of some generic market i. In this generic market, a public strategy for a consumer is a sequence of maps $a_{it} : h_i^t \to \{0,1\}$, a public strategy for the local player i is a sequence of maps $e_{Lit} : h_i^t \to \{0,1\}$, and a public strategy for the global player in market i is a sequence of maps $e_{Git} : h_i^t \to \{0,1\}$. A perfect public equilibrium (PPE) under independent branding is a profile of public strategies $\{a_i, e_{Li}, e_{Gi}\}$ if for each date t and history h^t, the strategies form a Nash equilibrium from that point on.

Under collective branding with n brands, a public strategy for a consumer i is a sequence of maps $a_{it} : h_i^t \to \{0,1\}$, a public strategy for the local player i is a sequence of maps $e_{Lit} : h_i^t \to \{0,1\}$, and a public strategy for the global player is a sequence of maps $e_{Gt} = (e_{G1t}, \dots, e_{Gnt}) : h_i^t \to \{0,1\}^n$. A perfect public equilibrium (PPE) under collective branding is a profile of public strategies $\{a_1, \dots, a_n, e_{L1}, \dots, e_{Ln}, e_G\}$ if for each date t and history h^t, the strategies form a Nash equilibrium from that point on. Moreover, an L-strongly symmetric PPE under collective branding is a PPE in which all n local players use the same strategy after every history. Our analysis of collective branding uses as the equilibrium concept L-strongly symmetric PPE. In Appendix B, we also study asymmetric PPE for the special case of two markets and show that they are strictly suboptimal for discount factors exceeding 1/2.

[17] Rather than its sum, we may take the aggregated signal $\tilde{s}^\tau$ as any symmetric and strictly increasing function of the individual signals $(s_1^\tau, \dots, s_n^\tau)$.

[18] For example, consider a beer tasting website such as *beeradvocate.com*. Even though large beer brands are often produced in several plants, including under license in foreign countries, the tasting notes on such websites do not distinguish between them.

To ensure that our analysis is non-trivial, we assume throughout that $c < \overline{c} \equiv (1-\alpha-\beta)/(1-\beta)$. This assumption is necessary and sufficient for effort to be sustainable for a large enough discount factor under independent branding.

3 Independent Branding

In this section, we analyze equilibrium outcomes when the goods in the different markets are branded independently. Since all markets are symmetric and independent, we fix some market i and drop the market subscript for the remainder of this section. All our payoff results are therefore in terms of "per-market averages". We sometimes use the superscript I to denote optimal solutions in this case of independent branding.

Worst PPE. It is straightforward to see that neither G nor L exerting any effort ($e_G^t = e_L^t = 0$) and short-lived consumers not purchasing the good ($b^t = 0$) in every period t after any history h^t is a PPE. In this PPE, both G and L receive their minmax payoff of zero. This minmax equilibrium outcome represents the worst PPE outcome with the associated payoff of $\underline{V}^I = 0$.

Strategy profiles sustaining high quality. If the best PPE yields a strictly positive payoff, $\overline{V}^I > \underline{V}^I = 0$, then it involves players exerting effort in equilibrium, resulting in high quality. We refer to such an equilibrium as as a high-quality equilibrium. Stated more formally, a *high-quality equilibrium* is a PPE with the outcome that the global and local players choose high effort in period 0 with probability 1.

Because effort is costly, a high-quality equilibrium must provide players with incentives to induce it. From Abreu, Pearce, and Stacchetti (1990), it is without loss to assume that a PPE takes on only extreme points of the equilibrium value set, and this equilibrium value set is a subset of the convex hull of the enforceable payoff vectors. Because of the perfect complementarities in effort, there are, in the case of independent branding, only two enforceable payoff vectors: the payoff vector associated with both long run players exerting effort and the minmax payoffs of zero. As a result, providing incentives for effort involves a probabilistic triggering of these minmax payoffs, intuitively representing a market breakdown. Hence, the best PPE is the high-quality equilibrium with the smallest market breakdown probability, provided high-quality equilibria exist. If a high-quality equilibrium does not exist, the best PPE coincides with the worst PPE so that $\overline{V}^I = \underline{V}^I = 0$.

Under independent branding, there is only a binary public signal s on which players can condition their behavior. As a result, any high-quality equilibrium sustained through the use of extreme points can be characterized by the probability of market breakdown, ρ_0, in the event that the signal s points to shirking ($s = 0$). More formally, a strategy profile $\sigma^I(\rho_0)$ sustaining such a high-quality equilibrium has the following structure: for $\rho_0 \in (0, 1]$, if the period-t history h^t involves $s^\tau = 0$ and $r^\tau \in [0, \rho_0]$ for some $\tau \leq t$, then $e^t_G = e^t_L = 0$ and $b^t = 0$; otherwise, $e^t_G = e^t_L = 1$ and $b^t = 1$.

The strategy profile $\sigma^I(\rho_0)$ implies that, in period 0, both G and L exert effort, and the consumer purchases the good. This continues in all subsequent periods until the public quality signal assumes the value of zero (falsely indicating that the quality in the previous period was zero) *and* the realized value of the public randomization device is not larger than ρ_0; from then on, no effort will ever be exerted and the good will not be purchased. In short, a bad quality signal triggers a reversion to the worst PPE with probability ρ_0.[19]

Payoffs and market breakdown probabilities. Playing the strategy profile $\sigma^I(\rho_0)$ yields a payoff of

$$\tilde{V}_j = \pi_j - \lambda_j c + \delta(1 - p_0)\tilde{V}_j = \lambda_j V(p_0, \gamma_j), \text{ with } V(p_0, \gamma_j) \equiv \frac{\gamma_j - c}{1 - \delta(1 - p_0)}, \tag{1}$$

to long-lived player $j \in \{G, L\}$, where $p_0 \equiv \alpha\rho_0$ represents the expected probability that, in any period after which effort was exerted and the consumer purchased the good, the long-run players stop exerting effort and consumers stop purchasing the good. We refer to p_0 as the on-path *market breakdown probability* of a high-quality equilibrium. Clearly, $V(p_0, \gamma_j) \geq 0$ if and only if $\gamma_j \geq c$. In this case, $V(p_0, \gamma_j)$ is decreasing in p_0 and increasing in γ_j. Using the identity $\lambda_G\gamma_G + \lambda_L\gamma_L = 1$, it follows that the payoffs of both long-lived players exceed the minmax payoff of zero if and only if $\gamma_G \in [c, (1 - \lambda_L c)/\lambda_G]$ and $\gamma_L \in [c, (1 - \lambda_G c)/\lambda_L]$.[20]

Incentive constraints. In equilibrium, every consumer receives a payoff of zero, and it is straightforward to see that no consumer has an incentive to deviate from the above strategy profile (which gives him just his minmax payoff of zero). To see whether any of the long-lived players G or L is better off deviating, note first that the answer is trivially no once the reversion to the worst PPE has been triggered. Consider now a one-shot deviation before such a reversion has been triggered: player j's value from one-time shirking is equal to

$$\tilde{V}^d_j = \pi_j + \delta(1 - p_1)\tilde{V}_j = \lambda_j \times [\gamma_j + \delta(1 - p_1)V(p_0, \gamma_j)]\,,$$

[19]Instead of a probabilistic permanent transition to the worst PPE, an alternative strategy profile would involve a deterministic transition to a finite punishment phase of length T, thus not requiring the existence of a public randomization device. In the absence of integer constraints on T, such deterministic strategies would support the same equilibrium outcome.

[20]The upperbound on γ_j follows from $\gamma_{-j} \geq c \Leftrightarrow 1/\lambda_{-j} - \gamma_j\lambda_j/\lambda_{-j} \geq c \Leftrightarrow \gamma_j \leq (1 - \lambda_{-j}c)/\lambda_j$.

where

$$p_1 \equiv (1-\beta)\rho_0$$

is the market breakdown probability when one of the players shirks.

The incentive constraint of player $j \in \{G, L\}$, $\tilde{V}_j \geq \tilde{V}_j^d$, can be written as

$$\delta(p_1 - p_0)V(p_0, \gamma_j) \geq c. \qquad (IC_j^I)$$

The left-hand side represents the discounted future loss from the one-shot deviation, induced by an increase in the market breakdown probability from p_0 to p_1, whereas the right-hand side represents the short-run gain, which equals the saved effort cost.

Characterizing the best PPE. In the best high-quality equilibrium, the punishment probability ρ_0 maximizes aggregate surplus

$$\tilde{V}_G + \tilde{V}_L = \lambda_G V(\alpha\rho_0, \gamma_G) + \lambda_L V(\alpha\rho_0, \gamma_L) = \frac{1-c}{1-\delta[1-\alpha\rho_0]}$$

subject to (IC_G^I) and (IC_L^I). Because the surplus is decreasing in ρ_0, this amounts to minimizing the punishment probability ρ_0 subject to the two incentive constraints. As (IC_j^I) depends on j only through γ_j, and $V(p_0, \gamma_j)$ is strictly increasing in γ_j, the optimal revenue shares maximize $\min\{\gamma_G, \gamma_L\}$, implying $\gamma_G = \gamma_L = 1$. Moreover, both incentive constraints must be binding: If only one constraint were binding, the revenue shares could be adapted to relax this constraint slightly and thereby lower ρ_0 and increase the objective. Hence, at the optimum, each long-lived player's reward share is proportional to his cost share: $(\pi_G, \pi_L) = (\lambda_G, \lambda_L)$. Defining,

$$\rho_0^I \equiv \frac{(1-\delta)c}{\delta[1-\alpha-\beta-(1-\beta)c]} \text{ and } \overline{\delta}^I \equiv \frac{c}{1-\alpha-\beta+\beta c}, \qquad (2)$$

we obtain the following proposition:

Proposition 1. *If $\delta \geq \overline{\delta}^I$, the best PPE is a high-quality equilibrium with a market breakdown probability of $p_0^I = \alpha\rho_0^I$. In this equilibrium, the sharing rule is proportional (i.e., $\gamma_G = \gamma_L = 1$); and the joint value is equal to $\overline{V}^I = V(p_0^I, 1) > 0 = \underline{V}^I$.*

Otherwise, a high-quality equilibrium does not exist and the best PPE coincides with the worst PPE in that $\overline{V}^I = \underline{V}^I = 0$.

The *average per-period payoff* $\overline{v}^I \equiv (1-\delta)\overline{V}^I$ in the best PPE is therefore given by

$$\overline{v}^I = \begin{cases} 0 & \text{if } 0 < \delta < \overline{\delta}^I, \\ 1 - c - \left(\frac{\alpha}{1-\alpha-\beta}\right) c & \text{if } \overline{\delta}^I \leq \delta < 1. \end{cases} \qquad (3)$$

For $\delta \geq \overline{\delta}^I$, the average payoff $\overline{v}^I$ is independent of the discount factor, and strictly less than the efficient payoff of $(1-c)$, identifying a failure of the folk theorem.[21] In the limit as the probability of a "false negative" (α) becomes small, this inefficiency vanishes: $\lim_{\alpha\to 0} \overline{v}^I = 1-c$. While this inefficiency also decreases as the probability of a "false positive" decreases, it does not vanish in the limit as β becomes small: $\lim_{\beta\to 0} \overline{v}^I = 1 - c/(1-\alpha) < 1-c$. Finally, note that the critical discount factor $\overline{\delta}^I$ is positively related to both α and β, with $\lim_{(\alpha+\beta)\to 0} \overline{\delta}^I = c$.

Indeed, under perfect monitoring ($\alpha = \beta = 0$), high quality provision in every period is sustainable for $\delta \geq c$ and yields a per-period equilibrium value of $1-c$. Hence, imperfect monitoring exacerbates the implementation of high quality in two ways. First, for a discount factor $\delta \in [c, \overline{\delta}^I)$, high quality is not sustainable with imperfect monitoring whereas it would be under perfect monitoring. Second, for $\delta \geq \overline{\delta}^I$, high quality in the initial period is sustainable both with perfect and imperfect monitoring, but the equilibrium value is lower with imperfect monitoring, $\overline{v}^I < 1-c$, as high quality cannot be sustained forever.

Note that expression (3) does not depend on the effort cost structure (λ_G, λ_L). This means that "vertical integration"—where the same agent chooses e_G^t and e_L^t and, in return, gets all of the revenue from selling the good—has no effect on the set of PPE values. As alluded to before, in our model, the physical team production problem can thus be solved costlessly.

4 Collective Branding

In this section, we analyze equilibrium outcomes when the long-lived players sell the goods in the n markets under one collective brand. In this case, the public history h^t contains the aggregated signals $\tilde{s}^\tau = \sum_i s_i^\tau$ of the previous periods $\tau \leq t$ rather than the individual signals s_i^τ. Since the public signal $\tilde{s}^\tau$ has $n+1$ possible realizations rather than only two as in the case of independent branding, the players' strategies in a PPE with collective branding are potentially more complex.

Worst PPE. Similar to the case of independent branding, however, there always exists a straightforward PPE in which, irrespective of the public histories, all long-lived players exert no effort in every period, and consumers do not purchase the good. As this gives each long-lived player his minmax payoff of zero, the joint value in the worst PPE under collective

[21] As first shown by Radner et al. (1986), the folk theorem may not hold in repeated games with imperfect monitoring and double-sided moral hazard. In our model, the signal structure conditional on the global player shirking is identical to the one conditional on the local player shirking. This implies a failure of the pairwise full rank condition in the repeated game, and results in the set of feasible payoffs not having full dimension. We refer to Appendix B for more details.

branding is $\underline{V}^C = 0$, coinciding with the worst PPE outcome under independent branding, $\underline{V}^I = 0$.

Poisson's binomial distribution. Recall that under collective branding, the aggregated public signal in period t, $\tilde{s}^t$, is equal to the number of markets in which the quality signal s_i^t indicated that quality was high in period $t-1$. That is, suppressing the period superscript t for notational convenience, $\tilde{s}$ has $n+1$ possible realizations. The probability distribution of these aggregated signals $\tilde{s}$ depends on the distribution of the underlying market-specific signals s_i.

If players exert effort in all n markets, the aggregated signal $\tilde{s}$ follows the standard binomial distribution of n independent Bernoulli trials, each with the identical success probability $1-\alpha$. However, if shirking occurs in some (but not all) markets, the distribution of $\tilde{s}$ does not correspond to a standard binomial distribution, since the success probability in a market without shirking is $1-\alpha$, whereas it is only β in a market where shirking occurs. In particular, if shirking takes place in k of the n markets, $\tilde{s}$ obtains from n trials of which $n-k$ have a success probability of $1-\alpha$, and k have a success probability of β. Let $\mathbb{P}_n(\tilde{s}|k)$ denote the resulting probability of $\tilde{s}$ successes.

The distribution $\mathbb{P}_n(\cdot|k)$ is thus the convolution of the binomial distribution of k trials with success probability β and the binomial distribution of $n-k$ trials with success probability $1-\alpha$.[22] It is also the convolution of $\mathbb{P}_{n-1}(\cdot|k)$ and a signal from a market without shirking, as well as the convolution of $\mathbb{P}_{n-1}(\cdot|k-1)$ and a signal from a market with shirking. Hence, $\mathbb{P}_n(s|k)$ exhibits the following recursive structure:

$$\mathbb{P}_n(s|k) = (1-\alpha)\mathbb{P}_{n-1}(s-1|k) + \alpha\mathbb{P}_{n-1}(s|k) = \beta\mathbb{P}_{n-1}(s-1|k-1) + (1-\beta)\mathbb{P}_{n-1}(s|k-1). \quad (4)$$

Being a special case of "Poisson's binomial distribution" (Wang, 1993), the probability distribution $\mathbb{P}_n(\cdot|k)$ is unimodal and log concave with expectation $\mathbb{E}_n(s|k) = k(1-\beta) + (n-k)\alpha$. A property, crucial for our analysis, is that Poisson's extension of the binomial distribution retains the monotone likelihood ratio property (MLRP). That is, for all $s, k \in \{1, \ldots, n\}$, the following holds:

$$\frac{\mathbb{P}_n(s|k)}{\mathbb{P}_n(s|k-1)} < \frac{\mathbb{P}_n(s-1|k)}{\mathbb{P}_n(s-1|k-1)}.$$

This also means that the distributions $\mathbb{P}_n(\cdot|k)$ are ordered in the sense of first-order stochastic dominance (FOSD), since MLRP implies FOSD.

[22] Our analysis does not require the use of an explicit formula for $\mathbb{P}_n(s|k)$. However, for completeness, we report here that, following Rukhin et al. (2009), $\mathbb{P}_n(s|k)$ can be written as $\mathbb{P}_n(s|k) = \sum_{i=0}^{k} \binom{k}{i}\binom{n-k}{s-i}\beta^i(1-\beta)^{k-i}(1-\alpha)^{s-i}\alpha^{n-k-s+i}$, using the convention that the binomial coefficient $\binom{k}{i}$ is 0 for a negative integer i.

Indeed, the FOSD-relation reflects the simple intuition that when shirking occurs in one more market, the players are less likely to observe at least the same number of successes. Yet, the recursive structure (4) implies that the probability of observing at least $s-1$ successes is greater than the probability of observing at least s successes without shirking in one more market.[23] Hence, for our Poisson's binomial distribution the magnitude of FOSD is also limited: for all $k = 0, \ldots, n-1$ and $s = 1, \ldots, n$, it holds that

$$\mathbb{P}_n(\tilde{s} \geq s|k+1) \leq \mathbb{P}_n(\tilde{s} \geq s|k) \leq \mathbb{P}_n(\tilde{s} \geq s-1|k+1). \tag{5}$$

The first inequality is FOSD, the second inequality describes the sense in which FOSD is limited. Both this notion of limited FOSD and the recursive structure (4) will be useful in the subsequent analysis.

Collective-branding strategies. Similar to the analysis with independent branding, we consider collective-branding strategy profiles $\sigma^C(\cdot)$ that are characterized by $n+1$ punishment probabilities, $\{\rho_s\}_{s=0}^n$, where s indicates the number of positive quality signals. In particular, if the period-t history is such that $\tilde{s}^\tau = s$ and $r^\tau \in [0, \rho_s]$ for some $\tau \leq t$, then $e^t_{G,1} = e^t_{G,2} = e^t_{L,1} = e^t_{L,2} = 0$ and $b^t_1 = b^t_2 = 0$; otherwise, $e^t_{G,1} = e^t_{G,2} = e^t_{L,1} = e^t_{L,2} = 1$ and $b^t_1 = b^t_2 = 1$.[24]

The strategy profile $\sigma^C(\cdot)$ implies that the repeated game starts, in period 1, with all long-lived players exerting effort, and in each market the consumer purchasing the good. This continues in all subsequent periods until the number of realized positive quality signals in some future period t is $\tilde{s}^t$ and the realization of the public randomization device is less than $\rho_{\tilde{s}^t}$, which then triggers a reversion to the worst PPE.

Market breakdown probabilities. Conditional on all long-lived players having exerted effort in the past, and consumers having purchased the goods, the strategy profile $\sigma^C(\cdot)$ induces breakdown probabilities, both on-path as well as following a deviation. The breakdown probability in the period after shirking in k markets is denoted p_k, and given by

$$p_k = \sum_{s=0}^{n} \mathbb{P}_n(s|k)\rho_s. \tag{6}$$

[23]To see this, note that by (4) it holds, $\mathbb{P}_n(\tilde{s} \geq s-1|k+1) = \beta\mathbb{P}_{n-1}(s-2|k) + \sum_{j=s-1}^{n-1} \mathbb{P}_{n-1}(j|k) + (1-\beta)\mathbb{P}_{n-1}(n|k)$ and $\mathbb{P}_n(\tilde{s} \geq s|k) = (1-\alpha)\mathbb{P}_{n-1}(s-1|k) + \sum_{j=s}^{n-1} \mathbb{P}_{n-1}(j|k) + \alpha\mathbb{P}_{n-1}(n|k)$, where $\mathbb{P}_{n-1}(n|k) = 0$. Subtracting the second from the first yields $\mathbb{P}_n(\tilde{s} \geq s-1|k+1) - \mathbb{P}_n(\tilde{s} \geq s|k) = \beta\mathbb{P}_{n-1}(s-2|k) + \alpha\mathbb{P}_{n-1}(s-1|k) \geq 0$.

[24]Since $\tilde{s} = n$ is indicative of no-shirking, it will always be optimal to have $\rho_n = 0$.

Incentive constraints. When players adopt the strategy profile $\sigma^C(\cdot)$, G's average value across the n markets, $\tilde{V}_G$, equals $\pi_G - \lambda_G c + \delta(1-p_0)\tilde{V}_G$, implying

$$\tilde{V}_G = \frac{\pi_G - \lambda_G c}{1-\delta(1-p_0)} = \lambda_G V(p_0, \gamma_G), \tag{7}$$

where $V(p_0, \gamma_G)$ is as defined in equation (1).

The global player is free to choose different effort levels in different markets.[25] Her (average-per-market) value from shirking in k markets in the current period and subsequently reverting to the collective branding strategy $\sigma^C(\cdot)$ is

$$\tilde{V}_G^{d,k} = \pi_G - \frac{n-k}{n}\lambda_G c + \delta[1-p_k]\tilde{V}_G, \tag{8}$$

where the second term on the right-hand side represents the average-per-market effort cost when shirking in k markets and exerting effort in the other $n-k$ markets. We can rewrite the incentive constraint, $\tilde{V}_G \geq \tilde{V}_G^{d,k}$, as

$$\delta(p_k - p_0)V(p_0, \gamma_G) \geq c \cdot \frac{k}{n}. \tag{$IC_G^{C,k}$}$$

Intuitively, the left-hand side represents the (average) long-term loss — the breakdown probability rising from p_0 to p_k — from the one-shot deviation, whereas the right-hand side represents the average short-run gain — the (per-market-average) reduction in effort costs — from that deviation.

Local player L's value under strategy profile $\sigma^C(\cdot)$, $\tilde{V}_L$, equals $\pi_L - \lambda_L c + \delta(1-p_0)\tilde{V}_L$, implying

$$\tilde{V}_L = \frac{\pi_L - \lambda_L c}{1-\delta(1-p_0)} = \lambda_L V(p_0, \gamma_L). \tag{9}$$

The local player's value from deviating to shirking in the current period and subsequently reverting back to the collective branding strategy $\sigma^C(\cdot)$ is

$$\tilde{V}_L^d = \pi_L + \delta(1-p_1)\tilde{V}_L. \tag{10}$$

We can write the local player's incentive-constraint, $\tilde{V}_L \geq \tilde{V}_L^d$, as

$$\delta(p_1 - p_0)V(p_0, \gamma_L) \geq c. \tag{IC_L^C}$$

Intuitively, the left-hand side represents the long-term loss — the breakdown probability rising from p_0 to p_1 — from the one-shot deviation, whereas the right-hand side represents

[25] In the conclusion, we discuss the case where G is constrained to take the same action in all markets.

the short-run gain — the reduction in effort costs — from that deviation.

The best high-quality equilibrium. The best high-quality equilibrium is characterized by the vector of punishment probabilities $\rho^C = (\rho_0^C, \ldots, \rho_n^C)$ that maximizes aggregate surplus

$$n\tilde{V}_G^d + n\tilde{V}_L^d = \frac{n(1-c)}{1-\delta(1-p_0)} = nV(p_0, 1) \tag{11}$$

subject to the $(n+1)$ incentive constraints $(IC_G^{C,k})$ and (IC_L^C). Note that maximizing aggregate surplus is equivalent to minimizing the on-path breakdown probability p_0, which is linear in the punishment probabilities ρ_s. Because we can also express the incentive constraint as constraints that are linear in the punishment probabilities ρ_s, characterizing the best PPE under collective branding involves solving a linear programming problem with $n+1$ constraints. Fixing the players' reward- to-cost-share ratios $\gamma \equiv (\gamma_G, \gamma_L)$, the vector of punishment probabilities ρ^C is a solution to the linear program

$$\begin{aligned}
\mathcal{P}(\gamma): \min_{(\rho_0,\ldots,\rho_n)} \quad & \sum_{s=0}^{n} \mathbb{P}_n(s|0)\rho_s \\
\text{s.t.} \quad & \sum_{s=0}^{n} \left[\frac{\mathbb{P}_n(s|k)-\mathbb{P}_n(s|0)}{k} - \frac{c\mathbb{P}_n(s|0)}{n(\gamma_G - c)}\right]\rho_s \geq \frac{(1-\delta)c}{n\delta(\gamma_G - c)}, \ \forall\, k \in \{1,\ldots,n\} \\
& \sum_{s=0}^{n} \left[\mathbb{P}_n(s|1)-\mathbb{P}_n(s|0) - \frac{c\mathbb{P}_n(s|0)}{\gamma_L - c}\right]\rho_s \geq \frac{(1-\delta)c}{\delta(\gamma_L - c)},
\end{aligned}$$

where constraint k is equivalent to the incentive constraint $(IC_G^{C,k})$ and the last constraint is equivalent to (IC_L^C).

The next lemma is our main step towards a characterization of solutions to $\mathcal{P}(\gamma)$.

Lemma 1. *Suppose ρ^C is a solution to $\mathcal{P}(\gamma)$. Then at ρ^C at least one constraint is binding and ρ^C exhibits the following cut-off structure: there is an integer $\overline{s} < n$ such that $\rho_s^C = 1$ for all $s < \overline{s}$ and $\rho_s^C = 0$ for all $s > \overline{s}$. The solution implies $p_0 \leq \ldots \leq p_n$.*

The lemma shows that, optimally, the punishment probabilities are concentrated in a bang-bang fashion on the lowest values of the public aggregated signal $\tilde{s}$. In particular, there is a *cutoff signal* $\overline{s}$ such that all realization of $\tilde{s}$ that lie below this threshold imply a market breakdown with certainty, whereas realizations of $\tilde{s}$ above the threshold imply no market breakdown whatsoever. Formally, the result follows from the MLRP property of Poisson's binomial distribution, which reflects the intuitive notion that low values of the public signal $\tilde{s}$ are less likely when agents put in effort. This property implies that concentrating the punishment probabilities on the lowest values of $\tilde{s}$ provides the strongest incentives for effort.[26]

[26] Cai and Obara (2009) establish a similar result in a related model but without any free-rider effects.

The cutoff signal $\overline{s}$, and its relation to the discount factor δ, play a crucial result in the subsequent analysis. Defining

$$\overline{\gamma}_G \equiv c + \frac{c\alpha}{n(1-\alpha-\beta)} \leq \overline{\gamma}_L \equiv c + \frac{c\alpha}{1-\alpha-\beta},$$

the next lemma formalizes the sense in which the optimal cutoff signal $\overline{s}$ deceases with the discount factor δ. It moreover shows that the cutoff signal equals 0 if the discount factor is close to 1, and, in this case, all constraints $k = 2, \dots, n$ in program $\mathcal{P}(\gamma)$ are slack at a solution.

Lemma 2. *A high-quality equilibrium exists only if $\gamma_G > \overline{\gamma}_G$ and $\gamma_L > \overline{\gamma}_L$. In this case, there is a critical discount factor $\overline{\delta}^C \in [0, 1)$ such that a high-quality equilibrium exists if and only if $\delta \geq \overline{\delta}^C$. For $\delta > \overline{\delta}^C$, the cut-off signal $\overline{s}$ is decreasing in δ. In particular, there is a threshold $\overline{\delta}_0^C < 1$ such that for $\delta > \overline{\delta}_0^C$, we have $\overline{s} = 0$ and, moreover, of the first n constraints in program $\mathcal{P}(\gamma)$ at most the constraint with respect to $k = 1$ is binding.*

Identifying the effects of a collective reputation. In order to identify both the informativeness effect and the free-rider effect under collective reputation, it is instructive to begin by considering two polar cases: First, the case in which only the global player, G, has to incur costly effort (i.e., $\lambda_G = 1$) so that only the informativeness effect arises, and, second, the one in which only local players have to incur costly effort for producing high quality (i.e., $\lambda_L = 1$), so that only the free-rider effect obtains. Due to the informativeness effect, collective branding in the first polar case permits sustaining a better reputation and a higher value in the best PPE than independent branding. In the second case, by contrast, collective branding induces only the free-rider effect, which tends to reduce the maximum sustainable value in the best PPE.

The informativeness effect. In order to identify the informativeness effect and show that in the absence of any reputational free-riding, collective branding is optimal, we first study the polar case in which the global player incurs all effort costs for producing high quality, i.e., $\lambda_G = 1$. The local players' effort in this polar case is costless so that they do not need any incentives to exert effort. It is therefore optimal to give the entire revenue share to the global player. Hence, just as under independent branding, the proportional reward scheme, $(\pi_G, \pi_L) = (\lambda_G, \lambda_L)$, is optimal.

Our first step in identifying the informativeness effect of collective branding is to show that when only effort from the global player matters, $\lambda_G = 1$, we can replicate the best PPE outcome under independent branding by collective branding. The replication is trivial if $\rho_0^I > 1$, implying that no high-quality equilibrium exists under independent branding. Suppose therefore instead that under independent branding we have $\rho_0^I \leq 1$. That is, independent branding allows a high-quality equilibrium, so that $\overline{V}^I > 0$, with $\alpha\rho_0^I$ as the minimum on-path market breakdown probability sustaining the best high-quality equilibrium. Turning to collective branding of the n goods, define the vector of punishment probabilities $\rho^C(\rho_0^I) = (\rho_1(\rho_0^I), \ldots, \rho_n(\rho_0^I))$ with

$$\rho_s(\rho_0^I) \equiv \frac{n-s}{n}\rho_0^I, \qquad s = 0, ..., n. \tag{12}$$

The following lemma obtains:

Lemma 3. *Suppose $\lambda_G = 1$ and $\rho_0^I \in [0,1]$ with value $\overline{V}^I > 0$. Then, the vector of punishment probabilities $\rho^C(\rho_0^I)$ sustains a high-quality equilibrium under collective branding with the same value $\overline{V}^I$ and the same on-path breakdown probability $p_0 = \alpha\rho_0^I$. Moreover, for each $k = 0, \ldots, n-1$, the incentive constraint $(IC_G^{C,k})$ coincides with the incentive constraint (IC_G^I) under independent branding.*

The lemma shows that, under collective branding, the vector of punishment probabilities $\rho^C(\rho_0^I)$ replicates the best high-quality equilibrium outcome under independent branding. This result may appear surprising.[27] Formally, it means that there is always a solution to a system of $(n+1)$ linear equations—the requirement $p_0^C = p_0^I$ and the n binding incentive constraints (IC_G^{Ck})—by $(n+1)$ variables—the punishment probabilities $(\rho_0, \ldots, \rho_n)$ each of which has to lie between 0 and 1. That such a solution always exists relies on the properties of Poisson's binomial distribution. To provide an intuition for why the construction in (12) represents such a solution, note that it ensures that p_k^C with collective branding equals the market breakdown probability ρ_0^I multiplied by the expected *ratio of failures* when shirking in a share k/n of n markets. This has two implications. First, in the absence of shirking, the induced probability of market breakdown with n markets equals the probability of market breakdown under independent branding, implying $p_0^C = p_0^I$. This is so because the expected ratio of failures when there are n markets equals this expected ratio when there is only one market: without shirking, they both equal α. As a consequence, collective branding with the punishment probabilities $\rho^C(\rho_0^I)$ induce the same value $\overline{V}^I$. Second, the construction implies that the *difference* in market breakdown probabilities between shirking in k and $l < k$ markets is linear in the additional number of markets in which shirking occurs: $p_k^C - p_l^C = (k-l)A$, where $A = (1-\alpha-\beta)$. Thus, $p_k^C - p_0^C = kA$ so that, under $\rho^C(\rho_0^I)$, all n incentive constraints (IC_G^{Ck}) collapse into a single one. The first implication then implies that this collapsed incentive constraint coincides with the incentive constraint under independent branding.

[27]It also contradicts the claim in Proposition 4 of Cabral (2009) that no-umbrella branding can be strictly optimal.

Note, however, that—except for the knife-edge case $n = 2$ and $\rho_0^I = 1$—the vector $\rho^C(\rho_0^I)$ does not satisfy the cutoff structure that Lemma 1 identifies as necessary for an optimal solution.[28] Consequently, under collective branding we can improve on the outcome associated with $\rho^C(\rho_0^I)$ by raising the breakdown probabilities $\rho_s(\rho^I)$ for smaller s to one and lowering them to zero for higher s.

This improvement identifies the beneficial informativeness effect of collective branding. Indeed, optimal punishments under collective branding concentrate market breakdown on those events in which the number of bad quality signals is large, because these events are more likely when shirking occurs and therefore represent the most efficient way to discourage shirking. Hence, the improvement on the replicated outcome under collective branding, as induced by $\rho^C(\rho_0^I)$, shows that the information structure under collective branding is more effective in preventing the global player to shirk than the information structure under independent branding. This informativeness effect is an implication of the natural MLR property of Poisson's binomial distribution, because Lemma 1's optimality result – showing that only punishment probabilities that display a cutoff structure use the collective signal optimally – is based on that property.

The following proposition confirms the superiority of collective branding when only the effort of the global player matters, even though signals are independent across markets and the global player has the flexibility to shirk in any number of markets.

Proposition 2. *Suppose $\lambda_G = 1$. Then, the optimal rewards exhibit $\hat{\gamma}_G = 1$ and collective branding is superior to independent branding: $\overline{V}^C \geq \overline{V}^I$. This superiority is strict if $\rho_0^I \leq 1$, except in the special case $n = 2$ and $\rho_0^I = 1$. If $\rho_0^I > 1$, then collective branding is strictly superior to independent branding for $n > 2$ and ρ_0^I close to one.*

The free-rider effect. In order to identify the free-rider effect, we next turn to the other polar case in which only local players incur effort costs, i.e., $\lambda_L = 1$. As this implies $\lambda_G = 1-\lambda_L = 0$, the global player's effort can be induced "for free" in that the global player does not need any incentives to exert effort. Hence, it is optimal to give the entire revenue share of each good i to L_i, implying once more that the proportional reward scheme, $(\pi_G, \pi_L) = (\lambda_G, \lambda_L)$, is optimal.

As in the previous polar case, we again first ask the question whether, for $\rho_0^I \leq 1$, collective branding can replicate the best outcome under independent branding. Note that for $\lambda_L = 1$, program $\mathcal{P}(\gamma)$ simplifies to a program with only the last constraint, since for $\lambda_G = 0$, the

[28]Note that in order to express program $\mathcal{P}$ properly for $\lambda_L = 0$, we must first rewrite the final constraint as $\sum_{s=0}^{n} \left[\mathbb{P}_n(s|1) - \mathbb{P}_n(s|0) - \frac{c\mathbb{P}_n(s|0)\lambda_L}{\pi_L - c\lambda_L}\right] \rho_s \geq \frac{(1-\delta)c\lambda_L}{\delta(\pi_L - c\lambda_L)}$, implying that for $\lambda_L = 0$ the constraint is automatically met. Hence, for the polar case $\lambda_G = 1$ Lemma 1 also holds.

first n constraints are automatically satisfied.[29]

Regardless of this simplification, Lemma 1 still applies. Hence, under collective branding, the optimal punishment vector ρ^C has the cut-off structure $\rho^C = (1, \ldots, 1, \rho^C_{\overline{s}}, 0, \ldots, 0)$ with some cutoff $\overline{s}$. Combining this with the observation that for the polar case $\lambda_L = 1$ only the local player's incentive constraint matters, allows us to pinpoint exactly the free-rider effect of collective branding. To do so, note that—under both independent and collective branding—we can rewrite the local player's incentive constraint as

$$\delta\left(\frac{p_1}{p_0} - 1\right) p_0 V(p_0, 1) \geq c,$$

where we used that for $\lambda_L = 1$ the best PPE exhibits $\gamma_L = 1$. Under independent branding, $p_1/p_0 = (1-\beta)/\alpha > 1$. By contrast, under collective branding, we can exploit the recursive structure (4) to rewrite p_0 and p_1, induced by $\rho^C = (1, \ldots, 1, \rho^C_{\overline{s}}, 0, \ldots, 0)$ with cutoff $\overline{s}$, as follows:

$$p_0 = \alpha\Delta + B; \text{ and } p_1 = (1-\beta)\Delta + B,$$

where $\Delta \equiv \mathbb{P}_{n-1}(\overline{s}-1|0)(1-\rho^C_{\overline{s}}) + \mathbb{P}_{n-1}(\overline{s}|0)\rho^C_{\overline{s}} > 0$ and $B \equiv \mathbb{P}_n(\tilde{s} \leq \overline{s}-1|0) + \mathbb{P}_{n-1}(\overline{s}-1|0)\rho^C_{\overline{s}} \geq 0$. Hence, under collective branding and for any ρ^C exhibiting a cutoff structure, the ratio of punishment probabilities satisfies

$$\frac{p_1}{p_0} = \frac{(1-\beta)\Delta + B}{\alpha\Delta + B} \leq \frac{1-\beta}{\alpha}, \tag{13}$$

where the (weak) inequality holds with equality if and only if $B = 0$, which requires $\overline{s} = 0$.

Recalling that under independent branding $p_1/p_0 = (1-\beta)/\alpha > 1$, we can now fully identify the free-rider effect on the basis of inequality (13). In particular, a local player choosing to shirk under collective branding does not increase the punishment probability by as much as he would under independent branding as he correctly anticipates the other local players to put in effort (and thus likely to generate positive signals). Only in the case in which, under collective branding, the transition to the worst PPE occurs only if all signals are bad ($\overline{s} = 0$) is the punishment probability ratio p_1/p_0 the same as under independent branding.

In order to see that this free-rider effect renders collective branding suboptimal, note that the minimum cutoff of ρ^C is $\overline{s} = 0$. This minimum cutoff is indeed the optimal one if and only if the only remaining constraint in $\mathcal{P}(\gamma)$ is satisfied for $\rho = (1, 0, \ldots, 0)$. Using $\mathbb{P}_n(0|1) = \alpha^{n-1}(1-\beta)$ and $\mathbb{P}_n(0|0) = \alpha^n$, we find the value of ρ^C_0 at which the constraint

[29] In order to see that the first n constraints in $\mathcal{P}(\gamma)$ are automatically satisfied for $\lambda_L = 1$, express them first in their proper form for $\lambda_G = 0$ as $\sum_{s=0}^{n}\left[\frac{\mathbb{P}_n(s|k)-\mathbb{P}_n(s|0)}{k} - \frac{c\mathbb{P}_n(s|0)\lambda_G}{n(\pi_G - c\lambda_G)}\right]\rho_s \geq \frac{(1-\delta)c\lambda_G}{n\delta(\pi_G - c\lambda_G)}$.

with $\rho^C = (\rho_0^C, 0, \ldots, 0)$ binds:

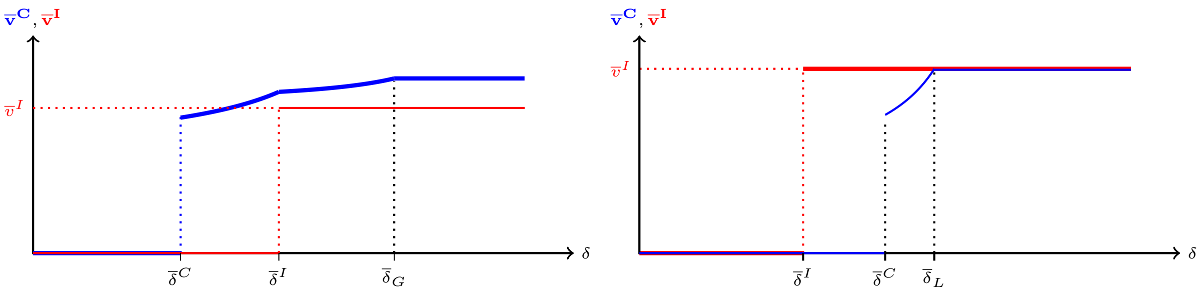
$$\rho_0^C = \overline{\rho}_0^L \equiv \frac{(1-\delta)c}{\delta\alpha^{n-1}[1-\alpha-\beta-(1-\beta)c]} = \frac{\rho_0^I}{\alpha^{n-1}}.$$

Hence, the optimal cutoff equals 0 whenever $\rho_0^I \leq \alpha^{n-1}$. In that case, market breakdown occurs only if all n markets yield a bad signal, which on path occurs with probability α^n. The on-path probability of market breakdown is therefore $\alpha^n \rho_0^C = \alpha\rho_0^I$, which is equal to the minimum on-path probability of market breakdown under independent branding. Consequently, for $\rho_0^I \leq \alpha^{n-1}$, the aggregate surplus associated with the best PPE under collective branding matches the aggregate surplus associated with the best PPE under independent branding so that we have $\overline{V}^C = \overline{V}^I$.

If $\rho_0^I > \alpha^{n-1}$ instead, then the best PPE with collective branding must be strictly worse than with independent branding: either collective branding can implement effort only with punishment probabilities $\rho_0^C = 1$ and $\rho_1^C > 0$, which due to MLRP, yields a larger breakdown probability p_0—implying $\overline{V}^I > \overline{V}^C > 0$, or collective branding cannot implement high effort at all—implying $\overline{V}^I > \overline{V}^C = 0$. In either case, we have $\overline{V}^I > \overline{V}^C$, i.e., collective branding performs strictly worse. This also means that $\overline{V}^I = 0$ implies $\overline{V}^C = 0$ so that the critical discount factor at which effort is sustainable with collective branding cannot be smaller than the corresponding discount factor under independent branding, $\overline{\delta}^I \leq \overline{\delta}^C$. We collect these insights in the following proposition.

Proposition 3. *Suppose $\lambda_L = 1$. Then, independent branding is superior to collective branding, i.e., $\overline{V}^I \geq \overline{V}^C$, and $\overline{\delta}^I \leq \overline{\delta}^C$. This superiority is strict if $\rho_0^I \in (\alpha^{n-1}, 1]$. If $\rho_0^I \leq \alpha^{n-1}$, independent branding and collective branding perform equally well, i.e., $\overline{V}^I = \overline{V}^C$.*

The proposition's superiority result of independent branding is due to the harmful free-rider effect of collective branding. In particular, the local player's continuation payoff under collective branding depends on the signals generated by other players. As he cannot affect those other signals, collective branding can only hurt incentives. Perhaps surprisingly, however, collective branding does as well as independent branding if $\rho_0^I \leq \alpha^{n-1}$. To understand this, note that – under independent branding – the local player is, conditional on generating a bad signal, punished only with probability ρ_0^I. If this conditional punishment probability is small (i.e., $\rho_0^I \leq \alpha^{n-1}$), then the same on-path punishment probability can be generated under collective branding by transiting to the worst PPE only if *all* n signals are bad. That is, only if the local player himself as well as all the other $n-1$ local players generate bad signals, implying that $\overline{s} = 0$ so that (13) holds with equality. From the viewpoint of the local player, the outcome of the other signals is purely random and the probability that all of them are bad (given that the other local players do not shirk) equals α^{n-1}. In other words, the randomness of the other $n-1$ signals under collective branding plays the same role as the



(a) Case $\lambda_G = 1$: $\overline{v}^C \geq \overline{v}^I$.

(b) Case $\lambda_L = 1$: $\overline{v}^I \geq \overline{v}^C$.

Figure 1: Best PPE's per-period values under independent vs. collective branding.

public randomization device under independent branding and therefore collective branding does not distort incentives. By contrast, if $\rho_0^I > \alpha^{n-1}$, then to generate the same on-path punishment probability under collective branding requires that the transition to the worst PPE may have to occur even if at least one of the n signals is positive. That is, if $\rho_0^I > \alpha^{n-1}$, then $\overline{s} > 0$ so that the inequality in (13) is strict. Since this means that, with some probability, a local player is "punished" even after generating a positive signal, collective branding is strictly harmful for incentives.

Contrasting the two effects. Figure 2 contrasts the two polar cases, displaying the comparative statics for the average per period payoff $\overline{v}^C$ and $\overline{v}^I$ with respect to the discount factor δ for those cases. Panel (a) depicts the case where only the global player needs to be incentivized ($\lambda_G = 1$). It illustrates the result of Proposition 2 that collective branding is superior to independent branding in general, and strictly so in two ways. First, in the case $\delta \geq \overline{\delta}^I$, where a high-quality equilibrium is sustainable with independent branding, collective branding can sustain it with a strictly lower on-path market breakdown probability (except in the special case $n = 2$ and $\rho_0^I = 1$). Second, in the case $\delta < \overline{\delta}^I$, where a high-quality equilibrium is not sustainable with independent branding, collective branding can sustain it for δ smaller but close enough to $\overline{\delta}^I$.

By contrast, panel (b) illustrates the implied comparative statics for the other polar case, where only the local players have to be incentivized ($\lambda_L = 1$). Defining $\overline{\delta}_L$ as the value of δ at which $\rho_0^I = \alpha^{n-1}$, the values $\overline{V}^C$ and $\overline{V}^I$ coincide for $\delta \geq \overline{\delta}_L$. As illustrated in panel (b), this implies that, in addition to $\overline{v}^I$, also the maximum average per-period payoff under collective branding, $\overline{v}^C$, is constant. For $\delta < \overline{\delta}_L$, we have $\overline{V}^C < \overline{V}^I$ and, due to continuity of $\overline{v}^C$ for $\delta > \overline{\delta}^C$, the maximum average per-period payoff $\overline{v}^C$ is therefore strictly increasing in δ over the interval $[\overline{\delta}^C, \overline{\delta}_L]$. Moreover, in the interval $[\overline{\delta}^I, \overline{\delta}^C]$ a high-quality equilibrium is only sustainable for independent branding. Consequently, the blue curve $\overline{v}^C$ lies always (weakly) below the red curve $\overline{v}^I$ – in stark contrast to panel (a). In short, panel (a) displays the optimality of collective branding for the case $\lambda_G = 1$, whereas panel (b) shows the optimality of independent branding for the other polar case $\lambda_L = 1$.

The trade-offs of a collective reputation. We now turn to the generic case in which the global and local players share the overall effort cost c according to the proportions $\lambda_G \in (0, 1)$ and $\lambda_L = 1 - \lambda_G$.

In Section 3, we showed that – under independent branding – it is optimal to provide the long-lived agents with proportional rewards: $(\pi_G, \pi_L) = (\lambda_G, \lambda_L)$. Trivially, this was also the case in the two polar cases of collective branding studied above.

It is therefore instructive to start our analysis of collective branding in the generic case assuming such proportional rewards: $\gamma_G = \gamma_L = 1$. It then follows that the incentive constraints coincide with the two polar cases studied above. Under collective branding and proportional rewards, the optimal on-path breakdown probability p_0 is minimized subject to

$$\delta(p_k - p_0)V(p_0, 1)\frac{n}{k} \geq c, \qquad k = 1, ..., n; \qquad (IC_G^{Ck})$$

$$\delta(p_1 - p_0)V(p_0, 1) \geq c. \qquad (IC_L^C)$$

Note however that by Lemma 1, we have $p_k - p_0 \geq p_1 - p_0$, which together with $k \leq n$ implies that (IC_G^{Ck}) follows from (IC_L^C). As a result, all (IC_G^{Ck}) are redundant so that the optimal on-path breakdown probability p_0 is minimized subject only to (IC_L^C). This, however, implies that, for proportional rewards $\gamma_G = \gamma_L = 1$, the intermediate case $\lambda_G \in (0, 1)$ boils down to the polar case with only local effort costs ($\lambda_L = 1$). As a result, Proposition 3 applies, meaning that it extends to all $\lambda_L \in (0, 1]$ and proportional rewards:

Proposition 4. *Suppose $\lambda_L \in (0, 1]$. Then collective branding is suboptimal with proportional rewards $(\pi_G, \pi_L) = (\lambda_G, \lambda_L)$, and strictly so for $\rho_0^I \in (\alpha^{n-1}, 1]$.*

This result demonstrates in an extreme sense the drawback of a collective reputation. As soon as explicit incentives for the local players' effort are needed, independent branding always outperforms collective branding with proportional rewards.

The proposition raises the question whether the long-lived agents can use the reward structure as a tool to mitigate this extreme effect of a collective reputation. We next argue that they can indeed do so: by carefully calibrating the revenue shares γ_G and γ_L, the long-lived players can reduce the local players' free-rider problem.

To see this, recall that, in general, the incentive constraints depend on the revenue shares γ_G and γ_L as follows:

$$\delta(p_k - p_0)V(p_0, \gamma_G)\frac{n}{k} \geq c, \qquad k = 1, ..., n; \qquad (IC_G^{Ck}(\gamma_G))$$

$$\delta(p_1 - p_0)V(p_0, \gamma_L) \geq c. \qquad (IC_L^C(\gamma_L))$$

Since $V(p_0, \gamma_L)$ is increasing in γ_L, an increase in γ_L relaxes the constraint (IC_L^C). Of course, an increase in the local players' revenue share γ_L is accompanied by a decrease in the global player's revenue share γ_G.[30] Yet, for $\gamma_L = \gamma_G = 1$, each (IC_G^{Ck}) holds strictly whenever the constraint (IC_L^C) binds so that by continuity a small increase in γ_L ensures that the corresponding small decrease in γ_G is such that each (IC_G^{Ck}) remains satisfied, while (IC_L^C) is relaxed. This reasoning suggests that, starting with proportional rewards, we can improve the objective under collective branding by relaxing (IC_L^C) through increasing the local players' reward γ_L. Defining

$$\tilde{\gamma}_G \equiv \frac{1 + (n-1)(1-\lambda_G)c}{\lambda_G + (1-\lambda_G)n} < 1 \text{ and } \tilde{\gamma}_L \equiv \frac{n - (n-1)\lambda_G c}{\lambda_G + (1-\lambda_G)n} > 1,$$

the following lemma refines this intuition and determines the bounds on the optimal γ_G and γ_L.

Lemma 4. *Suppose $\lambda_G \in (0,1)$ and $\overline{V}^C > 0$. Then the optimal reward-to-cost-share ratios $\hat{\gamma}_G$ and $\hat{\gamma}_L$ exhibit $\hat{\gamma}_G \in [\tilde{\gamma}_G, 1)$ and $\hat{\gamma}_L \in (1, \tilde{\gamma}_L]$ and ensure that the local and at least one of the global incentive constraints are binding. In particular, for $\delta \in (\overline{\delta}_0^C, 1)$, the global incentive constraint (IC_G^{C1}) is binding and $(\hat{\gamma}_G, \hat{\gamma}_L) = (\tilde{\gamma}_G, \tilde{\gamma}_L)$.*

Lemma 4 shows that, when incentives are needed for inducing local players' effort, proportional rewards are never optimal under collective branding. Hence, the lemma leaves open the possibility that, for the optimal reward structure, the minimum on-path breakdown probability that sustains high quality is actually lower under collective branding than under independent branding, implying $\overline{V}^C > \overline{V}^I$. The next proposition shows that this is indeed the case for large discount factors—no matter how small the global player's share of the effort cost is, provided it is strictly positive.

Proposition 5. *There exists a threshold $\overline{\delta} < 1$ such that $\overline{V}^C > \overline{V}^I$ for all $\delta \geq \overline{\delta}$ and all $\lambda_G > 0$.*

The previous results characterize properties of the best high-quality equilibrium without characterizing its outcome explicitly. Recall, however, that the best PPE induces a value $\overline{V}^C > 0$ if and only if the following program has a solution $\hat{p}_0$:

$$\mathcal{P}^{\mathcal{C}}: \min_{(\gamma_G, \gamma_L, \rho_0, \ldots, \rho_n)} p_0 = \sum_{s=0}^{n} \mathbb{P}_n(s|0)\rho_s$$

$$\text{s.t. } \delta(p_k - p_0)V(p_0, \gamma_G)n/k \geq c \quad \forall k = 1, \ldots, n; \qquad (IC_G^{Ck}(\gamma_G))$$

[30] In particular, $\gamma_G = \pi_G/\lambda_G = (1-\pi_L)/(1-\lambda_L) = (1-\lambda_L\gamma_L)/(1-\lambda_L)$.

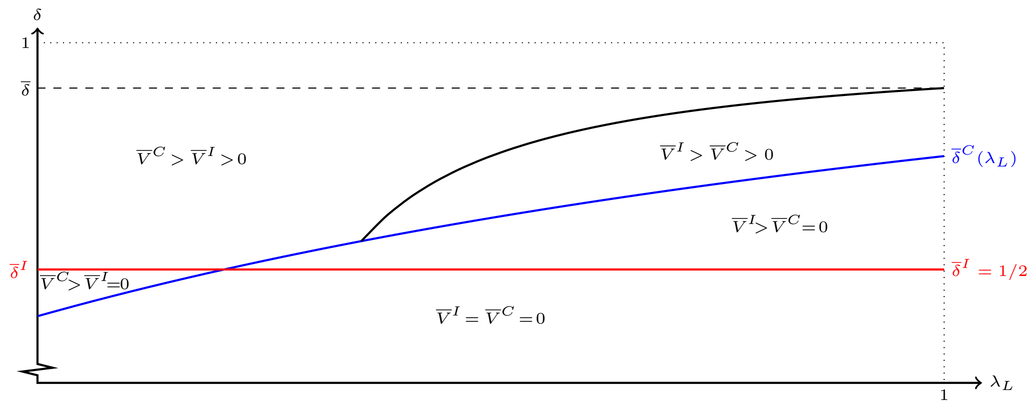


Figure 2: Comparison of $\overline{V}^C$ and $\overline{V}^I$ for $n = 3$ and parameters $\alpha = \beta = 1/3$ and $c = 1/5$.

$$\delta(p_1 - p_0)V(p_0, \gamma_L) = c; \qquad (IC^C_L(\gamma_L))$$

$$\lambda_L\gamma_L + \lambda_G\gamma_G = 1, \qquad (14)$$

where (14) is the accounting identity that links γ_G and γ_L. We denote a solution to $\mathcal{P}^C$ by a triple $(\hat{\gamma}_G, \hat{\gamma}_L, \hat{\rho}) \in \mathbb{R}^{n+1}$. In the case in which $\mathcal{P}^C$ admits a solution $(\hat{\gamma}_G, \hat{\gamma}_L, \hat{\rho})$ with value $\hat{p}_0$, it follows that

$$\overline{V}^C = \frac{1-c}{1-\delta(1-\hat{p}_0)} > 0.$$

Writing out the program that determines the value $\overline{V}^C$ allows us to obtain the following comparative static result:

Proposition 6. *Suppose that collective branding sustains a high-quality equilibrium: $\overline{V}^C > 0$. Then, the optimal reward-to-cost-share ratio $\hat{\gamma}_L$ is strictly increasing, and the induced value $\overline{V}^C$ strictly decreasing, in the cost share λ_L.*

The proposition confirms the intuition that we can interpret λ_L as a measure of the relative magnitude of the free-rider effect. The larger is the share of the effort cost that needs to be borne by the local players, the larger are those players' incentives to take a free ride, and therefore the larger is the reward-to-cost-share ratio $\hat{\gamma}_L$ that the local players optimally receive. While this increase mitigates the free-rider problem, the flip side is that this decreases the reward-to-cost-share ratio $\hat{\gamma}_G$ and therefore exacerbates the global player's incentive problem so that the resulting aggregate payoff is reduced.

For a specific parametrization, Figure 2 illustrates our results on the optimality of collective branding in relation to the discount factor δ and the magnitude of the source of the free-rider effect, λ_L. It shows the general insight that, for low discount factors, effort is neither implementable under independent nor collective branding so that $\overline{V}^I = \overline{V}^C = 0$. For intermediate discount factors, effort is only implementable with collective branding when λ_L is small (i.e., $\overline{V}^C > \overline{V}^I = 0$), whereas it only implementable with independent branding when λ_L is large (i.e., $\overline{V}^I > \overline{V}^C = 0$). For larger discount factors, effort is implementable under both independent and collective branding ($\overline{V}^C > \overline{V}^I > 0$), but collective branding outperforms independent branding only when λ_L is not too large ($\overline{V}^I > \overline{V}^C > 0$). For discount factors exceeding $\overline{\delta}$, collective branding outperforms independent branding, no matter how large $\lambda_L < 1$.

5 Optimal Collective Brand Size

In the previous section, we have shown that, in the presence of local players, a careful calibration of the revenue shares is essential for building a collective brand reputation. In this section, we identify a second essential tool for managing collective branding reputation: the size of the collective brand. In particular, we show that the relative magnitude of the free-rider effect, as measured by the parameter λ_L, determines how the sustainability of high quality varies with brand size.

In this section, we thus analyze the comparative statics in the brand size n, obtaining results on the maximum implementable brand size, $\overline{n}$, and the optimal brand size, $\hat{n}$. Our main result reveals that, even when brand size is chosen optimally (without any additional constraints), there remains an inefficiency in the limit as the discount factor becomes large. Despite the usual intractability of Poisson's binomial distributions (e.g., Biscarri et al., 2018), the recursive structure (4) of the probability distribution $\mathbb{P}_n(\cdot|k)$ enables us to express the limiting inefficiency in closed form as a function only of (i) the corresponding limiting inefficiency in the polar case in which the local players bear all effort costs, and (ii) the effort cost share λ_L of the local players.

To make clear the dependencies on the collective brand size n, we denote variables with a superscript n throughout this section. For instance, $\overline{V}^n$ now denotes the average per-market value in the best equilibrium when the brand size is n.

Informativeness effect only ($\lambda_L = 0$). We first consider the benchmark in which collective reputation does not exhibit a free-rider problem. For this benchmark, we show that, for any discount factor $\delta > c$, we obtain efficiency in the limit as n goes to infinity, implying that both the maximum implementable brand size, $\overline{n}$, and the optimal brand size, $\hat{n}$, are unbounded. An intuition behind this efficiency results follows from the observation that with perfect monitoring ($\alpha = \beta = 0$) efficiency obtains if and only if $\delta \geq c$. Extending the brand size mitigates the inefficiency induced by imperfect monitoring. This inefficiency vanishes in the limit as n becomes large.[31]

Recall that this benchmark corresponds to the polar case $\lambda_G = 1$. For this case, the incentive constraint $(IC_L^n(\gamma_L))$ is redundant so that the optimal pair $(\hat{\rho}^n, \hat{\gamma}_G^n)$ that characterizes the best PPE exhibits

$$\hat{\gamma}_G^n = 1 \text{ and } \hat{\rho}^n = \arg\min_{\rho} p_0^n(\rho) \text{ s.t. } (IC_G^{n;1}(1)), \ldots, (IC_G^{n;n}(1)).$$

Proposition 7. *Suppose $\lambda_G = 1$ and $\delta > c$. Then the maximum implementable brand size, $\overline{n}$, and the optimal brand size, $\hat{n}$, are unbounded, and efficiency obtains in the limit:*

$$\lim_{n\to\infty} \overline{v}^n = 1 - c.$$

In the appendix, we prove the proposition using the following steps. Based on a sandwich argument, we first show that, in the limit, $\hat{\rho}^n$ is of the deterministic form $\hat{r}^n = (1, \ldots, 1, 0, \ldots, 0)$. We next show that minimizing p_0 with respect to such deterministic cutoffs, only the two extreme constraints (IC_G^1) and (IC_G^n) need to be considered.[32] In a final step, we apply central limit arguments and Chebyshev's inequality to show that, in the limit as n goes to infinity, efficiency obtains for these deterministic cutoffs.

Including the free-rider effect ($\lambda_L > 0$). Proposition 7 implies that without a moral hazard problem for the local players, efficiency obtains when expanding the collective reputation over an infinite number of markets. The next proposition shows that this is no longer true as soon as there is a slight moral hazard problem concerning the local player's effort. In this case, there is an upper bound on the extent of collective branding.

Proposition 8. *Suppose $\lambda_L > 0$. Then there is an upper bound $\overline{n} \in \mathbb{N}$ such that effort is implementable under collective branding of size n only if $n < \overline{n}$. Consequently, $\overline{V}^n > 0$ only if $n < \overline{n}$. Moreover, $\overline{n}$ is increasing in δ.*

Our final result returns to the main focus of our paper: identifying the inefficiencies of a collective reputation due to free-riding by local players. Proposition 7 shows that if collective reputation does not suffer from a free-rider problem, then, for any discount factor $\delta > c$, inefficiencies vanish as the collective brand size grows large. Our final proposition shows that if there is a free-rider problem ($\lambda_L > 0$), then, even as the discount factor δ approaches 1, the maximum per-period payoff $\overline{v}^n$ is bounded away from efficiency.

[31] In the literature on repeated games with imperfect monitoring, Matsushima (2001) also notes this effect.

[32] Confirmed by simulations, we conjecture that this also holds with respect to an optimal non-deterministic $\hat{\rho}^n$ for finite n. The intractability of the Poisson's binomial distribution however prevents us from proving this analytically.

Proposition 9. *Suppose* $\lambda_L > 0$. *Then,*

$$\lim_{\delta\to 1} \overline{n} = \infty; \ \lim_{\delta\to 1} \hat{n} = \infty; \ \textit{and } \overline{v}^{\infty} \equiv \lim_{n\to\infty} \lim_{\delta\to 1} \overline{v}^n = \lambda_G(1-c) + \lambda_L \overline{v}^I < 1-c.$$

The proposition shows that, even in the limit, efficiency is not attained: the limiting payoff, $\overline{v}^{\infty}$, lies strictly below $1-c$. More importantly, it reveals the economic insight that we can decompose this limiting payoff by expressing it as a convex combination of the efficient payoff, $1-c$, and the (inefficient) payoff under independent branding, $\overline{v}^I$, with the weights corresponding to the effort cost shares λ_G and λ_L.

This decomposability result indicates, that in the limit, it is as if we obtain the best of both worlds: implementing collective branding for the global player, yielding the efficient payoff $1-c$, and, at the same time, independent branding for the local players, with its optimal but inefficient payoff of $\overline{v}^I$. The optimal calibration of the revenue shares is crucial for this decomposability. This does, however, not mean that the two incentive problems are independent and do not interact; at all times incentive constraints of both types of players are binding.

6 Conclusion

We have developed a theory of collective brand reputation in a repeated game of imperfect public monitoring. The key novelty is the interaction between a global player, who takes costly actions to impact the quality of the entire product line, and local players, each of whom is able to affect the quality of only a single product. This makes the analysis applicable to a large set of economic environments in which such two-sided moral hazard problems are endemic, including platform markets, franchising, licensing, and team production.

While under independent branding, the quality signals relating to different products are received individually by the product's consumer, they are effectively pooled under collective branding. If all of the effort costs are borne by the global player, only a beneficial informativeness effect arises, implying that collective branding is superior to independent branding. In that case, any inefficiency arising from imperfect monitoring vanishes in the limit as the collective brand size becomes large. By contrast, if all of the effort costs are borne by the local players, only a free-rider effect obtains, implying that collective branding is never superior to independent branding, and is strictly inferior unless the discount factor is sufficiently large. In the generic case in which both types of players bear some of the effort costs, a careful calibration of revenues shares mediates the tradeoff between the beneficial informativeness effect and the harmful free-rider effect. Under optimal revenue sharing, collective branding is superior to independent branding as long as the share of the effort costs borne by the local players is sufficiently small or the discount factor sufficiently large. As the discount factor becomes large, the optimal size of the collective brand increases without bound. In the limit as both the discount factor and the collective brand size become large, the remaining inefficiency is equal to the local players' effort cost share multiplied by the inefficiency under independent branding. In that limit, it is thus as if the best of both worlds could be achieved: collective branding for the global player and independent branding for the local players.

Throughout the paper, we have assumed that the global player makes separate effort choices for each product/market. This may be a reasonable assumption for some applications (think of a franchisor's delivery of beef to hamburger outlets or its advertising in local media[33]) but perhaps less so for some others (think of a headquarter's advertising in national media). If, under collective branding, the headquarter had to choose the same effort level in all markets, then only one aspect would need to be changed in our analysis: the global player would no longer have n incentive constraints but only a single one, namely (IC_G^{Cn}). It follows immediately that the resulting value under collective branding is weakly larger – and strictly larger for sufficiently large discount factors, than with separate effort decisions.[34] As we show in the final part of Appendix A, however, all of our propositions would continue to hold under that alternative assumption. In particular, the comparative statics and limiting values remain valid.

We have also assumed that effort choices are private information. An exciting avenue for future research consists in allowing for within-brand monitoring of effort choices. Such an analysis would, however, require a solution concept beyond PPE and therefore a more complex (and less well-understood) analytical framework. For instance, if the global player were to observe signals of local players' effort choices that are more informative than those observed by consumers, the global player's strategy would naturally depend on her private history (at least for the equilibrium to improve upon the PPE outcome in the absence of monitoring).

Because our analysis reveals the crucial role of revenue sharing, a further interesting question is to identify alternative modeling assumptions under which players benefit from more elaborate revenue sharing schemes than the ones we have analyzed. First, note that if the signals were fully contractible so that revenue shares could directly condition on them, then the players would be able to solve completely the moral hazard problem by using budget breakers, along the lines of Holmström (1982). Suppose instead, as we assume in the paper, that such direct conditioning on signal realizations is, due to their non-verifiability, infeasible. Then the players may, following the logic of relational contracting (e.g., Levin,

[33]As discussed in Blair and Lafontaine (2005, ch. 9.6), regional or local advertising effort is often determined not by the local franchisee but the franchisor, leading to frictions between franchisees and franchisor (see, e.g., *Broussard v. Meineke Discount Muffler Shops*).

[34]Recall from Lemma 4 that, in the best PPE, (IC_G^{C1}) is binding, and (IC_G^{Cn}) slack, for δ large.

2003), try to exploit the repeated game structure to implement conditional revenue sharing implicitly through voluntary payments.[35] However, under our assumption that signals are fully uninformative about behavior of specific players (and, under collective branding, in specific markets), such relational contracts cannot help alleviate the moral hazard problem.[36] An open question though is whether this might be different if, under collective branding, all players could not only observe the aggregate signal but also attribute the individual signals to a specific local market. While such an assumption does not reflect our interpretation of collective branding, we expect relational contracts that implement voluntary payments from local markets with a bad signal to local markets with a good signal, to be sustainable and alleviate the moral hazard problem. However, for such outcomes to be attainable in a PPE, consumers must also be able to observe such voluntary transfers between producers in different markets, which seems unlikely to hold in practice.

Moreover, we assumed that effort choices of the global and local players are binary and perfectly complementary within a market. This assumption allowed us to focus on the reputational team production problem across markets and abstract from the physical team production problem within a market. In a richer production structure with continuous effort choices that are imperfect complements, players must then, on the one hand, also solve the physical team production problem, but have, on the other hand, more punishments abilities to control reputational team production. As these complexities are beyond the scope of this paper, we leave such an analysis for future research.

In our analysis, we have assumed that all markets are identical. Allowing for a market-specific signal structure (α_i, β_i) would enable us to study a number of novel questions such as *which* products to group together and sell under a common brand name, or how to optimally design aggregate brand-level quality signals. However, such an extension would have to deal with at least two analytical difficulties. First, this would require giving up the convenient restriction to symmetric equilibria. Second, this would require dealing with more complex Poisson's binomial distributions.[37]

Another interesting topic for future work consists in studying optimal task assignment within a collective brand. Suppose that production requires a continuum of tasks, indexed by $i \in [0, 1]$. Let $c_G(i)$ and $c_L(i)$ denote the effort cost for the global and local player, respectively, in performing task i. Suppose that $\Delta c(i) \equiv c_G(i) - c_L(i)$ is strictly decreasing in i, with $\Delta c(\hat{\imath}) = 0$ for some $\hat{\imath}$. First-best efficiency thus requires that tasks $[0, \hat{\imath})$ are performed by local players, and tasks $(\hat{\imath}, 1]$ by the global player. An implication of our analysis in Section

[35]The relational contracting literature offers the insight that our focus on static revenue shares is without loss.

[36]Technically, the condition of "pairwise identifiability" (Fudenberg et al., 1994) fails in our context.

[37]As Poisson's binomial distributions in general satisfy MLRP, we would expect the main insights of our analysis to carry over to such an extension.

3 is that such "myopic" cost minimization is indeed optimal under independent branding. Under collective branding, however, our results imply that the global player should optimally take on more tasks as the joint value is decreasing in the share of the effort costs borne by the local players.[38] That is, optimal task assignment introduces a productive inefficiency under collective branding to mediate the free-rider problem.

Finally, we point out that in our modeling of collective branding, we have assumed that players only see a global, non-market specific signal about previous qualities. In practice, buyers may however observe both global and local signals about the provided quality, giving rise to the possibility of distinct reputations of global and local players. As a concrete example of this, consider Boeing's problems at their local assembly lines in Charleston.[39] After airlines learned about these local problems, they refused to take delivery of airplanes assembled there. Hence, a further worthwhile extension is to consider a model in which players base their quality perceptions on both types of signal, and study their interplay and effects on buyer behavior.

Appendix A: Proofs

Proof of Proposition 1: To obtain the best PPE, we minimize ρ_0 subject to (IC_G^I) and (IC_L^I). As (IC_j^I) is violated for ρ_0 small (provided $\lambda_j > 0$), the optimal ρ_0 must be such that one of the two incentive constraints holds with equality, and the other with a weak inequality. That is,

$$\rho_0 = \hat{\rho}_0(\gamma_G, \gamma_L) \equiv \frac{(1-\delta)c}{\delta\left[(1-\alpha-\beta)\min_j \gamma_j - (1-\beta)c\right]}. \tag{15}$$

This solution satisfies the domain restriction $\rho_0 \in [0,1]$ if $\hat{\rho}_0(\gamma_G, \gamma_L)$ is positive and not larger than one, which holds if and only if[40]

$$\min_j \gamma_j \geq \overline{\gamma} \equiv \frac{(1-\beta\delta)c}{\delta(1-\alpha-\beta)}. \tag{16}$$

If (16) does not hold, then an equilibrium with $e_G^0 = e_L^0 = 1 = b^0 = 1$ and the given (γ_H, γ_L) does not exist.

From these considerations it then follows that, at an optimum, we must have $\gamma_H = \gamma_L$.

[38] In the context of platform markets, the "Fulfillment by Amazon" (FBA) program may be understood through this lens. The FBA program, established in 2006, amounted to Amazon (as the global player) taking over the tasks of storage, shipping and handling returns from individual merchants (the local players).

[39] See https://www.bizjournals.com/seattle/news/2020/09/15/boeing-consolidation-787-assembly-study-everett.html (last retrieved on 02/20/2022).

[40] It is straightforward to verify that if the r.h.s. of (15) is not larger than 1 it is also positive.

To see this, first suppose $\gamma_G > \gamma_L$ is optimal. We must then have that (16) is satisfied and

$$\rho_0 = \hat{\rho}_0(\gamma_G, \gamma_L) = \frac{(1-\delta)c}{\delta\left[(1-\alpha-\beta)\gamma_L - (1-\beta)c\right]}.$$

But then we can lower ρ_0 further by raising γ_L slightly, since $\hat{\rho}_0(\gamma_G, \gamma_L)$ is decreasing in γ_L whenever $\gamma_G > \gamma_L$ (and this increase is also feasible since it relaxes (16)). Conversely, also $\gamma_G < \gamma_L$ cannot be optimal, since we could then lower ρ_0 further by raising γ_G. Hence, at the optimum it must hold $\gamma_G = \gamma_L$, which, together with the accounting identity $\lambda_G\gamma_G + \lambda_L\gamma_L = 1$ implies that $\gamma_G = \gamma_L = 1$.

Having established that at an optimum $\gamma_G = \gamma_L = 1$, it then follows from (15) that the minimizing ρ_0 equals ρ_0^I as defined in (2). Finally from (16), it follows that an equilibrium sustaining $e_G^0 = e_L^0 = 1 = b^0 = 1$ exists if and only if $\delta \geq \overline{\delta}^I$, where we note that our parameter restriction at the end of Section 2, $c < \overline{c}$, implies $\overline{\delta}^I < 1$. Q.E.D.

Proof of Lemma 1: First note that solving $\mathcal{P}(\gamma)$ disregarding all constraints, yields $\rho_s = 0$ for all s, but this violates all constraints. Because the optimization problem is linear, it follows that at least one of the constraints must be binding at an optimal solution.

Second, suppose to the contrary that ρ^C is optimal but such a $\overline{s}$ does not exist. Then there are $l < h$ such that $\rho_l < 1$ and $\rho_h > 0$, with ρ^C satisfying all the constraints of $\mathcal{P}(\gamma)$. Consider changing ρ^C to $\hat{\rho}^C$ by only lowering ρ_h by $\Delta\rho > 0$ and raising ρ_l by $\Delta\rho \cdot \mathbb{P}_n(s_h|0)/\mathbb{P}_n(s_l|0)$. This change does not affect $p_0 = \sum_s^n \mathbb{P}_n(s|0)\rho_s$. Therefore, the objective and the right-hand side of all the constraints remain unchanged. The left-hand side of the constraints change by

$$\left\{[\mathbb{P}_n(s_l|k) - \mathbb{P}_n(s_l|0)]\frac{\mathbb{P}_n(s_h|0)}{\mathbb{P}_n(s_l|0)} - [\mathbb{P}_n(s_h|k) - \mathbb{P}_n(s_h|0)]\right\}\frac{\Delta\rho}{k}.$$

After rewriting the term in curly brackets as

$$\left\{\frac{\mathbb{P}_n(s_l|k) - \mathbb{P}_n(s_l|0)}{\mathbb{P}_n(s_l|0)} - \frac{\mathbb{P}_n(s_h|k) - \mathbb{P}_n(s_h|0)}{\mathbb{P}_n(s_h|0)}\right\}\mathbb{P}_n(s_h|0) = \left\{\frac{\mathbb{P}_n(s_l|k)}{\mathbb{P}_n(s_l|0)} - \frac{\mathbb{P}_n(s_h|k)}{\mathbb{P}_n(s_h|0)}\right\}\mathbb{P}_n(s_h|0),$$

the MLRP of Poisson's binomial distribution implies that the term is strictly positive so that the left-hand side of each constraint strictly increases. As a result, $\hat{\rho}^C$ must also be optimal, since it attains the same objective value and all constraints are strictly satisfied. The latter however contradicts the first observation that for any solution to $\mathcal{P}(\gamma)$ at least one constraint is binding.

To see the final claim of the lemma, note that given the cutoff structure, it follows that $p_k = \sum_{s=0}^n \mathbb{P}_n(s|k)\rho_s = \mathbb{P}_n(\tilde{s} < \overline{s}|k) + \mathbb{P}_n(\overline{s}|k)\rho_{\overline{s}} = (1-\rho_{\overline{s}})\mathbb{P}_n(\tilde{s} < \overline{s}|k) + \rho_{\overline{s}}\mathbb{P}_n(\tilde{s} \leq \overline{s}|k) \leq (1-\rho_{\overline{s}})\mathbb{P}_n(\tilde{s} < \overline{s}|k+1) + \rho_{\overline{s}}\mathbb{P}_n(\tilde{s} \leq \overline{s}|k+1) = \mathbb{P}_n(\tilde{s} < \overline{s}|k+1) + \mathbb{P}_n(\overline{s}|k+1)\rho_{\overline{s}} = p_{k+1}$ for all $k = 0, \ldots, n-1$, where the inequality follows from first-order stochastic dominance (i.e., equation (5)). Q.E.D.

Proof of Lemma 2: In order to prove the lemma, we first construct an algorithm for finding a solution to $\mathcal{P}(\gamma)$ on the basis of Lemma 1. First, define $\overline{s}$ as the smallest $\tilde{s} \in \{0, \ldots, n-1\}$ such that

$$\sum_{s=0}^{\tilde{s}} \Delta_G(k,s) \geq \frac{(1-\delta)c}{n\delta(\gamma_G - c)}, \qquad \forall k = 1, ..., n, \tag{17}$$

and

$$\sum_{s=0}^{\tilde{s}} \Delta_L(1,s) \geq \frac{(1-\delta)c}{\delta(\gamma_L - c)}, \tag{18}$$

where

$$\Delta_G(k,s) \equiv \frac{\mathbb{P}_n(s|k) - \mathbb{P}_n(s|0)}{k} - \frac{c\mathbb{P}_n(s|0)}{n(\gamma_G - c)} \text{ and } \Delta_L(k,s) \equiv \frac{\mathbb{P}_n(s|k) - \mathbb{P}_n(s|0)}{k} - \frac{c\mathbb{P}_n(s|0)}{\gamma_L - c}.$$

The variable $\overline{s}$ is found algorithmically by starting with $\tilde{s} = 0$ and increasing it successively until either all of the $n+1$ inequalities associated with (17) and (18) hold, or $\tilde{s} = n$. If this procedure ends with $\tilde{s} = n$, then $\overline{s}$ does not exist, implying that the feasible set of $\mathcal{P}(\gamma)$ is empty so that, for the given γ, there is no high-quality equilibrium. If the procedure ends with $\tilde{s} < n$, then $\overline{s} = \tilde{s}$. In a next step, compute for $k = 1, \ldots, n$,

$$\rho_G^k \equiv \frac{1}{\Delta_G(k,\overline{s})} \left\{ \frac{(1-\delta)c}{n\delta(\gamma_G - c)} - \sum_{s=0}^{\overline{s}-1} \Delta_G(k,s) \right\}; \quad \rho_L \equiv \frac{1}{\Delta_L(1,\overline{s})} \left\{ \frac{(1-\delta)c}{\delta(\gamma_L - c)} - \sum_{s=0}^{\overline{s}-1} \Delta_L(1,s) \right\}.$$

By construction, each ρ_G^k and ρ_L lies in $[0,1]$. Taking $\overline{\rho}$ as the maximum over all ρ_G^k and ρ_L, it then follows from Lemma 1 that the solution ρ^C of program $\mathcal{P}(\gamma)$ exhibits $\rho_s^C = 1$ for $s < \overline{s}$, $\rho_{\overline{s}}^C = \overline{\rho}$, and $\rho_s^C = 0$ for $s > \overline{s}$, with an attained objective of $p_0 = \sum_{s=0}^{\overline{s}} \mathbb{P}_n(s|0) + \mathbb{P}_n(\overline{s}|0)\overline{\rho}$.

From this algorithm, we next identify the comparative statics of the threshold signal $\overline{s}$ with respect to the discount factor δ. To do so, first note that the left-hand sides of the $n+1$ inequalities associated with (17) and (18) are independent of δ, whereas the right-hand sides decrease with δ over $[0,1]$ and grow arbitrarily large as δ approaches 0. This implies that the threshold $\overline{s}$ fails to exist when δ is small and when it does exist for some $\hat{\delta}$, it exists for all $\delta > \hat{\delta}$ and, moreover, $\overline{s}$ is decreasing in δ. This establishes

In order to derive the critical discount factor $\overline{\delta}^C$, we first compute for each of the constraints in (17) the minimum $\overline{\delta}_G^k$ such that there is an $\tilde{s}$ that fulfills it. In particular,

$$\overline{\delta}_G^k \equiv \frac{c}{\Delta_G^k n(\gamma_G - c) + c}, \text{ where } \Delta_G^k \equiv \max_{\tilde{s}} \left\{ \sum_{s=0}^{\tilde{s}} \Delta_G(k,s) \right\}. \tag{19}$$

In order to see that $\gamma_G > \overline{\gamma}_G$ is a sufficient and necessary condition for $\Delta_G^k > 0$ for all

$k = 1, \ldots, n$, note that $\Delta_G(k, s) \geq 0$ only if $\Delta_G(k, 0) \geq 0$, since

$$\Delta_G(k,0) < 0 \;\Leftrightarrow \frac{\mathbb{P}_n(0|k)}{\mathbb{P}_n(0|0)} < 1 + \frac{kc}{n(\gamma_G - c)} \Rightarrow \frac{\mathbb{P}_n(s|k)}{\mathbb{P}_n(s|0)} < 1 + \frac{kc}{n(\gamma_G - c)} \Leftrightarrow \Delta_G(k,s) < 0,$$

where "$\Rightarrow$" follows from MLRP. Hence, $\Delta_G^k > 0$ if and only if $\Delta_G(k, 0) > 0$, where the latter is equivalent to

$$\gamma_G > c + \frac{kc}{n[((1-\beta)/\alpha)^k - 1]}.$$

Since the right hand side is decreasing in k,[41] a sufficient and necessary condition for $\Delta_G^k > 0$ for all $k = 1, \ldots, n$ is $\gamma_G > \overline{\gamma}_G$.

Likewise, compute for the constraint (18) the maximum $\overline{\delta}_L$ such that there is an $\tilde{s}$ that fulfills it. That is,

$$\overline{\delta}_L \equiv \frac{c}{\Delta_L(\gamma_L - c) + c}, \text{ where } \Delta_L \equiv \max_{\tilde{s}} \left\{ \sum_{s=0}^{\tilde{s}} \Delta_L(1, s) \right\} \tag{20}$$

and a sufficient and necessary condition for $\Delta_L > 0$ is $\gamma_L > \overline{\gamma}_L$.

Hence, we have established the first statement of the lemma that a high-quality equilibrium exists only if $\gamma_G > \overline{\gamma}_G$ and $\gamma_L > \overline{\gamma}_L$.

By taking $\overline{\delta}^C$ as the maximum over all $\overline{\delta}_G^k$ and $\overline{\delta}_L$, we also establish the the second and third statement of the lemma, concerning the threshold discount factor $\overline{\delta}^C$ and that the optimal cut-off signal $\overline{s}$ is decreasing in δ.

We finally show the last part of the lemma, proving the statements about the threshold level $\overline{\delta}_0^C$.

As shown, $\gamma_G > \overline{\gamma}_G$ and $\gamma_L > \overline{\gamma}_L$ imply $\Delta_G(k, 0) > 0$ for all $k = 1, \ldots, n$, and $\Delta_L(1, 0) > 0$. Noting that the right-hand sides of (17) and (18) vanish when δ approaches 1, the algorithm stops for $\tilde{s} = 0$, for δ close enough to 1, namely for $\delta \geq \overline{\delta}_0^C$, where

$$\overline{\delta}_0^C \equiv \max \left\{ \max_k \left\{ \frac{c}{c + n(\gamma_G - c)\Delta_G(k,0)} \right\}, \frac{c}{c + (\gamma_L - c)\Delta_L(1,0)} \right\}.$$

To see the last statement of the lemma, note that with $\overline{s} = 0$, (IC_G^{Ck}) reduces to

$$n\delta(\mathbb{P}_n(0|k) - \mathbb{P}_n(0|0))\rho_0 \frac{V(p_0, 1)}{k} \geq c.$$

Because $\mathbb{P}_n(0|k) = [(1-\beta)/\alpha]^k \alpha^n$ it follows from $(1-\beta)/\alpha > 1$ that $\mathbb{P}_n(0|k)$ is convex in k. As a result, the left hand side of (IC_G^{Ck}) is increasing in k while the right hand side is

[41] Its derivative is of the same sign as $\psi(r, k) \equiv r^k - 1 - kr^k \log r$, where $r \equiv (1-\beta)/\alpha > 1$. As $\psi(1, k) = 0$ and $\partial\psi(r, k)/\partial r = -k^2 r^{k-1} \log r < 0$ for $r > 1$, it follows that $\psi(r, k) < 0$.

independent of k. Hence, if the constraint holds for $k = 1$, it holds for all $k > 1$. Q.E.D.

Proof of Lemma 3: Suppose $\rho_0^I \leq 1$, implying $\delta \leq \overline{\delta}^I$, and consider the associated vector of punishment probabilities $\rho^C(\rho_0^I) = (\rho_0(\rho_0^I), \ldots, \rho_n(\rho_0^I))$ with $\rho_s(\rho_0^I)$ as defined by (12). Since $\rho_s(\rho_0^I) \leq \rho_0^I \leq 1$ and the signs of $\rho_s(\rho_0^I)$ and ρ_0^I coincide, it follows $\rho_s(\rho_0^I) \in [0, 1]$ for all $s = 0, ..., n$. Moreover, the punishment probabilities in $\rho^C(\rho_0^I)$ lead to a market breakdown after shirking in k markets of

$$p_k = \sum_{s=0}^{n} \mathbb{P}_n(s|k)\rho_s(\rho_0^I) = \sum_{s=0}^{n} \mathbb{P}_n(s|k)\frac{n-s}{n}\rho_0^I = \frac{(n - \mathbb{E}[s|k])}{n}\rho_0^I = \frac{(n-k)\alpha + k(1-\beta)}{n}\rho_0^I.$$

In particular, the on-path breakdown probability p_0 coincides with the one under independent branding: $p_0 = \alpha\rho_0^I$. Moreover, for these punishment probabilities each $(IC_G^{C,k})$ is equivalent to the incentive constraint (IC_G^I):

$$\frac{n}{k} \cdot \delta(p_k - p_0)V(p_0, 1) \geq c \;\Leftrightarrow\; \delta\rho_0^I(1 - \alpha - \beta)V(p_0, 1) \geq c.$$

Hence, $\rho^C(\rho_0^I)$ replicates the best PPE outcome under independent branding. Q.E.D.

Proof of Proposition 2: As defined in (2), recall from Proposition 1 that ρ_0^I represents the punishment probability associated with the best PPE under independent branding. If $\rho_0^I > 1$ then $\overline{V}^I = 0$ so that a (weak) superiority of collective branding holds trivially since $\overline{V}^C \geq 0$. Hence, suppose $\rho_0^I \leq 1$. Lemma 3 shows that the $\rho^C(\rho_0^I)$ is feasible in the sense that is satisfies all incentive constraints $(IC_G^{C,k})$, and attains the value $\overline{V}^I$. But because $\rho_s(\rho_0^I)$ violates the optimal cutoff structure (except for the special case of $n = 2$ and $\rho_0^I = 1$), Lemma 1 then implies that we can strictly improve upon this outcome. It follows that the best PPE yields a strictly higher payoff under collective branding. This then also implies that $\overline{\delta}^I \geq \overline{\delta}^C$, with a strict inequality for $n > 2$. Hence, for $n > 2$, it holds $\overline{\delta}^I > \overline{\delta}^C$ so that for any $\delta \in [\overline{\delta}^C, \overline{\delta}^C)$, implying that $\rho_0^I > 1$ but close to 1, we have $\overline{V}^C > \overline{V}^I = 0$, which proves the last statement of the proposition. Q.E.D.

Proof of Proposition 3: Follows directly from the text. Q.E.D.

Proof of Proposition 4: Follows directly from the text. Q.E.D.

Proof of Lemma 4: Suppose to the contrary that $\overline{V}^C > 0$ but $\hat{\gamma}_L \leq 1$, implying $\hat{\gamma}_G \geq 1$. Then, as argued, the program of minimizing p_0 w.r.t. ρ^C subject to all (IC_G^{Ck}) and (IC_L^C) has a solution $\rho^C = (1, \ldots, 1, \rho_{\overline{s}}, 0, \ldots, 0)$ such that (IC_L^C) binds, while all constraints (IC_G^{Ck}) are slack. Since $\partial V(p_0, \gamma_L)/\partial\gamma_L > 0$, raising γ_L slightly results in all constraints being slack, allowing to lower p_0 by reducing $\rho_{\overline{s}}$ (or if $\rho_{\overline{s}} = 0$ lowering $\overline{s}$). This contradicts that $\hat{\gamma}_L \leq 1$

is optimal. If $\hat{\gamma}_L \geq 1$ but no (IC_G^{Ck}) is binding, one can lower p_0 by the same procedure, whereas if (IC_L^C) is slack one can lower p_0 through a similar procedure by raising γ_G. We conclude that $\hat{\gamma}_G < 1 < \hat{\gamma}_L$ and are such that (IC_L^C) and at least one (IC_G^{Ck}) is binding.

Lemma 2 shows that for $\delta \in (\overline{\delta}_0^C, 1)$, the binding constraint must be (IC_G^{C1}). In this case, $(\hat{\gamma}_L, \hat{\gamma}_G)$ are such that both (IC_G^{C1}) and (IC_L^C) hold with equality, implying

$$\delta(p_1 - p_0)V(p_0, \hat{\gamma}_G)n = c = \delta(p_1 - p_0)V(p_0, \hat{\gamma}_L) \;\Leftrightarrow\; n(\hat{\gamma}_G - c) = \hat{\gamma}_L - c.$$

Combining the latter equation with the identity $\lambda_G\gamma_G + \lambda_L\gamma_L = 1$ yields $\hat{\gamma}_G = \tilde{\gamma}_G$ and $\hat{\gamma}_L = \tilde{\gamma}_L$. Note that for $\delta < \overline{\delta}_0^C$ the constraint (IC_G^{C1}) may be slack at the optimum. Hence, in general, we have

$$\delta(p_1 - p_0)V(p_0, \hat{\gamma}_G)n \geq c = \delta(p_1 - p_0)V(p_0, \hat{\gamma}_L) \;\Leftrightarrow\; n(\hat{\gamma}_G - c) \geq \hat{\gamma}_L - c.$$

Combining this latter inequality with our previous finding that $\hat{\gamma}_G < 1 < \hat{\gamma}_L$ yields the result $\hat{\gamma}_G \in [\tilde{\gamma}_G, 1)$ and $\hat{\gamma}_L \in (1, \tilde{\gamma}_L]$. Q.E.D.

Proof of Proposition 5: By Proposition 3, collective branding with proportional rewards can attain the value $\overline{V}^I$, whenever $\rho_0^I \leq \alpha^{n-1}$. Lemma 4 then implies that $\overline{V}^C > \overline{V}^I$, since it shows that proportional rewards are strictly suboptimal, meaning that collective branding can attain a value exceeding $\overline{V}^I$. The result then follows from the observation that $\rho_0^I \leq \alpha^{n-1}$ if and only if $\delta \geq \overline{\delta}$, where

$$\overline{\delta} = \frac{c}{c + \alpha^{n-1}(1 - \alpha - \beta - (1-\beta)c)}.$$

Q.E.D.

Proof of Proposition 6: Suppose the model's parameters are such that $\overline{V}^C > 0$, implying that there is an optimal triple $(\hat{\gamma}_G, \hat{\gamma}_L, \hat{\rho})$ to $\mathcal{P}^C$. In particular, for this triple $(\hat{\gamma}_G, \hat{\gamma}_L, \hat{\rho})$, the constraint $(IC_L^C(\hat{\gamma}_L))$ as well as at least one constraint $(IC_G^{Ck}(\hat{\gamma}_G))$ are binding. Now consider an increase in the parameter λ_G, while keeping γ_G constant at $\hat{\gamma}_G$. These changes imply a decrease in λ_L together with an increase in π_G. From $\gamma_L = (1 - \lambda_G\gamma_G)/(1 - \lambda_G)$, it however follows that $\partial\gamma_L/\partial\lambda_G > 0$ so that the overall effect on γ_L is positive. Hence, (IC_L^C) is relaxed, implying that we can, in fact, also increase γ_G slightly above $\hat{\gamma}_G$, thereby relaxing all constraints so that $\hat{\rho}$ together with the raised γ_G and γ_L lead to the same p_0 but with all constraints satisfied with strict inequality. By Lemma 4, the triple is suboptimal, meaning there is a different triple leading to a strictly lower p_0, implying a strictly larger $\overline{V}^C$. Q.E.D.

Proof of Proposition 7: Let $r^s \in \mathbb{R}^{n+1}$ denote deterministic cutoffs ρ of the form $(1, \ldots, 1, 0, \ldots, 0)$. That is, r^s is an $(n+1)$-dimensional vector with the first s entries being 1 and the remaining

$n+1-s$ entries being 0. Denote by $R^n = \{r^0, r^1, \dots, r^{n+1}\} \subset \mathbb{R}^{n+1}$ the set of all r^s for a given n.

Given n and $\gamma_G^n = 1$, recall that we have

$$\hat{\rho}^n = \arg\min_{\rho} p_0(\rho) \text{ s.t. } (IC_G^{n;1}(1)), \dots, (IC_G^{n;n}(1)).$$

In addition to this minimization problem, consider for a given n and $\gamma_G^n = 1$ the problems

$$\tilde{\rho}^n = \arg\min_{\rho} p_0(\rho) \text{ s.t. } (IC_G^{n;1}(1)) \text{ and } (IC_G^{n;n}(1));$$

$$\hat{r}^n = \arg\min_{r^s \in R^n} p_0(r^s) \text{ s.t. } (IC_G^{n;1}(1)), \dots, (IC_G^{n;n}(1));$$

$$\tilde{r}^n = \arg\min_{r^s \in R^n} p_0(r^s) \text{ s.t. } (IC_G^{n;1}(1)) \text{ and } (IC_G^{n;n}(1)).$$

Because the first problem is less stringent than the minimization problem underlying $\hat{\rho}^n$, whereas the second is more stringent, it follows

$$p_0(\tilde{\rho}^n) \leq p_0(\hat{\rho}^n) \leq p_0(\hat{r}_n).$$

Moreover, if $\tilde{\rho}^n$ has the cutoff at $\bar{s}$, then $\tilde{r}^n = r^{\bar{s}+1}$ so that

$$p_0(\tilde{r}^n) - p_0(\tilde{\rho}^n) = \mathbb{P}_n\{\bar{s}|0\}(1 - \tilde{\rho}_{\bar{s}}^n).$$

Since $\lim_{n\to\infty} \mathbb{P}_n\{\bar{s}|0\} = 0$, it follows

$$\lim_{n\to\infty} p_0(\tilde{\rho}^n) = \lim_{n\to\infty} p_0(\tilde{r}^n). \tag{21}$$

Because the minimization problem associated with $\tilde{r}^n$ is a relaxed version of the minimization problem associated with $\hat{r}_n$, it holds $p_0(\hat{r}_n) \geq p_0(\tilde{r}^n)$. The next lemma shows that, in fact, $p_0(\hat{r}_n) = p_0(\tilde{r}^n)$.

Lemma 5. *The minimizer $\hat{r}^n$ minimizes p_0 subject to only the incentive constraints $(IC_G^{n;1}(1))$ and $(IC_G^{n;n}(1))$.*

Proof. The statement is trivially satisfied for $n = 2$, so suppose $n > 2$. Rewriting $(IC_G^{n;k}(1))$ as

$$\frac{p_k^n - p_0^n}{k} \geq \frac{c}{\delta n V(p_0, 1)} \tag{22}$$

shows that if p_k is convex in k, in the sense that the incremental differences

$$p_{k+2}^n - p_{k+1}^n - (p_{k+1}^n - p_k^n) \tag{23}$$

are positive, then the left hand side of (22) is increasing in k so that $(IC_G^{n;1}(1))$ implies all other constraints and, hence, at most $(IC_G^{n;1}(1))$ can be binding. If, in contrast, p_k is concave in k, then the left hand side is decreasing in k so that $(IC_G^{n;n}(1))$ implies all other constraints and, hence, at most $(IC_G^{n;n}(1))$ can be binding.

We next argue that, for any $r^{\overline{s}} \in R^n$, the curvature of $p_k(r^{\overline{s}})$ w.r.t. k, i.e., the sign of (23), depends on the cutoff $\overline{s}$. To see this, note first that the recursive structure (4) of $\mathbb{P}_n(s|k)$ implies that the incremental differences (23) rewrites as:

$$p^n_{k+2} - p^n_{k+1} - (p^n_{k+1} - p^n_k) = \sum_{s=0}^{\overline{s}} [\mathbb{P}_n(s|k+2) + \mathbb{P}_n(s|k) - 2\mathbb{P}_n(s|k+1)]$$
$$= (1-\alpha-\beta)^2\{[\mathbb{P}_{n-2}(\overline{s}|k) - \mathbb{P}_{n-2}(\overline{s}-1|k)].\} \qquad (24)$$

The single peakedness of $\mathbb{P}_{n-2}(.|k)$ implies that $\mathbb{P}_{n-2}(s|k)$ is increasing in s for all s smaller than the mode, m_k, of $\mathbb{P}_{n-2}(.|k)$ and decreasing in s for all s larger than m_k. Since m_k is decreasing in k, m_k lies in between the modes of the binomial distributions $B(n-2,\beta)$ and $B(n-2,(1-\alpha))$, i.e., $m_k \in [m_0, m_{n-2}]$ with $m_0 = \lfloor (n-1)\beta \rfloor$ and $m_{n-2} = \lfloor (n-1)(1-\alpha) \rfloor$, where $\lfloor x \rfloor$ is the greatest integer less than or equal to x. Therefore, if $\overline{s} \leq m_0$, then (24) is positive for all k and, hence, the incremental differences are positive for all k, implying p_k is convex. Similarly, if $\overline{s}-1 > m_{n-2}$, then (24) is negative for all k and, hence, the incremental differences are negative for all k, implying p_k is concave. For $\overline{s} \in [m_0, m_{n-2}+1]$, the sign of (24) may depend on k. The single peakedness of $\mathbb{P}_{n-2}(.|k)$ together with the decreasing mode m_k implies however that (24) can switch sign at most once as k increases and only from positive to negative. If so, there is a $\overline{k}$ such that p_k is convex for all $k \leq \overline{k}$ and concave for all $k \geq \overline{k}$. This means that for $k \leq \overline{k}$, $(IC_G^{n;1}(1))$ implies $(IC_G^{n;k}(1))$, and for $k \geq \overline{k}$, $(IC_G^{n;n}(1))$ implies $(IC_G^{n;k}(1))$. □

A direct corollary of Lemma 5 is that

$$p^n_0(\tilde{\rho}^n) \leq p^n_0(\hat{\rho}^n) \leq p^n_0(\hat{r}_n) = p^n_0(\tilde{r}^n).$$

Since this string of inequalities holds for all n, it holds also in the limit, implying that

$$\lim_{n\to\infty} p^n_0(\tilde{\rho}^n) \leq \lim_{n\to\infty} p^n_0(\hat{\rho}^n) \leq \lim_{n\to\infty} p^n_0(\hat{r}_n) = \lim_{n\to\infty} p^n_0(\tilde{r}^n).$$

By (21) and a sandwich theorem, it then follows that

$$\lim_{n\to\infty} p^n_0(\tilde{\rho}^n) = \lim_{n\to\infty} p^n_0(\hat{\rho}^n) = \lim_{n\to\infty} p^n_0(\hat{r}_n) = \lim_{n\to\infty} p^n_0(\tilde{r}^n).$$

The next lemma shows that, for large enough n, we can pick deterministic cutoffs such that p_0 is arbitrarily close to zero, while the constraints are all slack, provided that $\delta > c$.

Lemma 6. *Suppose $\delta > c$ so that $\overline{\varepsilon} \equiv (\delta - c)/\delta > 0$. Then, for all $\varepsilon \in (0, \overline{\varepsilon})$*

$$\exists \overline{n} \in \mathbb{N} : \forall n > \overline{n},\ \exists \overline{s}(n) \leq n : p_0^n(r^{\overline{s}(n)}) < \varepsilon \text{ and } (IC_G^{n;1}(1)) \text{ and } (IC_G^{n;n}(1)) \text{ are slack.}$$

Proof. Fix $\varepsilon > 0$ and let

$$K = \frac{c}{\delta V(\varepsilon, 1)(1 - \alpha - \beta)}.$$

Denote the pdf of the normal distribution $\mathcal{N}(\mu, \sigma^2)$ with mean $\mu = n(1-\alpha)$ and variance $\sigma^2 = n\alpha(1-\alpha)$ by

$$\varphi_n(s) = \frac{1}{\sqrt{2\pi n\alpha(1-\alpha)}} e^{-\frac{1}{2}\frac{(s-n(1-\alpha))^2}{n\alpha(1-\alpha)}}.$$

Then there is an $\tilde{n}_1$ such that for all $n > \tilde{n}_1$, the equation

$$\varphi_n(\tilde{s}(n)) = \frac{K}{(1-\varepsilon)n}$$

has the solution $\tilde{s}(n) = \mu - \sigma h(n) < (1-\alpha)n - 1$ with

$$h(n) \equiv \sqrt{\ln\left(\frac{n}{2\pi\alpha(1-\alpha)K^2/(1-\varepsilon)^2}\right)}.$$

By noting that $\lim_{n\to\infty} h(n) = \infty$, Chebyshev's inequality implies that there is an $\tilde{n}_2 (> \tilde{n}_1)$ such that for all $n > \tilde{n}_2$ it holds

$$\Phi_n(\tilde{s}(n)) \equiv \Pr\{s \leq \tilde{s}(n)\} = \int_{-\infty}^{\tilde{s}(n)} \varphi_n(s) ds < \varepsilon/2.$$

Since the binomial distribution $B(n, (1-\alpha))$ converges in distribution to $\mathcal{N}(\mu, \sigma^2)$, there is an $\tilde{n}_3 (> \tilde{n}_2)$ such that for any $n > \tilde{n}_3$

$$|\mathbb{P}_n\{s \leq \overline{s}(n)|0\} - \Phi_n(\tilde{s}(n))| < \varepsilon/2,$$

where $\overline{s}(n) = \lceil \tilde{s}(n) \rceil$ is the smallest integer greater than $\tilde{s}(n)$. Hence, for any $n > \tilde{n}_3$

$$p_0^n(r^{\overline{s}(n)}) = \mathbb{P}_n\{s \leq \overline{s}(n)|0\} < \Phi_n(\tilde{s}(n)) + \varepsilon/2 < \varepsilon.$$

It remains to be shown that $r^{\overline{s}(n)}$ satisfies $(IC_G^{n;1}(1))$ and $(IC_G^{n;n}(1))$ with slackness.

The constraint $(IC_G^{n;1}(1))$ for $r^{\overline{s}(n)}$ rewrites as

$$\mathbb{P}_{n-1}(\overline{s}(n)|0) \geq \frac{1}{n} \frac{c}{\delta V(p_0(r^{\overline{s}(n)}), 1)(1-\alpha-\beta)}.$$

Since $V(p_0^n, 1)$ is strictly decreasing in p_0^n and $p_0^n(r^{\overline{s}(n)}) < \varepsilon$, the inequality holds strictly if

$$\mathbb{P}_{n-1}(\overline{s}(n)|0) \geq K/n.$$

The de Moivre-Laplace theorem implies there is an $\tilde{n}_4(> \tilde{n}_3)$ such that for all $n > \tilde{n}_4$

$$1 - \varepsilon < \frac{\mathbb{P}_{n-1}(\overline{s}(n)|0)}{\varphi_n(\overline{s}(n))} < 1 + \varepsilon.$$

Hence, for $n > \tilde{n}_4$ it holds

$$\mathbb{P}_{n-1}(\overline{s}(n)|0) > (1-\varepsilon)\varphi_n(\overline{s}(n)) \geq (1-\varepsilon)\varphi_n(\tilde{s}(n)) = K/n,$$

where the second inequality follows since, $\overline{s}(n) \leq \lceil \tilde{s}(n) \rceil \leq \lceil (1-\alpha)n - 1 \rceil \leq (1-\alpha)n$. This shows that $r^{\overline{s}(n)}$ satisfies $(IC_G^{n;1})$ with slackness for all $n > \tilde{n}_4$.

Recall $(IC_G^{n;n})$ for $r^{\overline{s}(n)}$ rewrites as

$$p_n^n(r^{\overline{s}(n)}) - p_0^n(r^{\overline{s}(n)}) \geq \frac{c}{\delta V(p_0(r^{\overline{s}(n)}), 1)}.$$

Since $\lim_{n\to\infty} \tilde{s}(n)/n = (1-\alpha) > \beta$, it follows by the law of large numbers that for any $\tilde{\varepsilon} > 0$, there is an $\tilde{n}_5(> \tilde{n}_4)$ so that for all $n > \tilde{n}_5$, we have

$$p_n^n(r^{\overline{s}(n)}) = \sum_{s=0}^{\overline{s}(n)} \mathbb{P}_n(s|n) > 1 - \tilde{\varepsilon}.$$

In particular, for $\tilde{\varepsilon} \in (0, (\delta - c - \delta\varepsilon)/(\delta(1-c)))$, it follows

$$p_n^n(r^{\overline{s}(n)}) - p_0^n(r^{\overline{s}(n)}) > 1 - \tilde{\varepsilon} - \varepsilon > \frac{c(1-\delta+\delta\varepsilon)}{\delta(1-c)} = \frac{c}{\delta V(\varepsilon, 1)} > \frac{c}{\delta V(p_0^n(r^{\overline{s}(n)}), 1)}.$$

This confirms that $(IC_G^{n;n}(1))$ for $r^{\overline{s}(n)}$ holds with slackness for all $n > \tilde{n}_5$. Taking $\overline{n} = \tilde{n}_5$ completes the lemma. □

Noting that p_0^n arbitrarily close to 0 means that the average per period payoffs, v^n, is arbitrarily close to $1-c$, then yields the Proposition. Q.E.D.

Proof of Proposition 8: If $\lambda_L > 0$ and $\overline{V}^n > 0$, then there is a triple $(\hat{\gamma}_G^n, \hat{\gamma}_L^n, \hat{\rho}^n)$ where

$\hat{\rho}^n = (1, \ldots, 1, \rho_{\overline{s}^n}, 0, \ldots, 0) \in \mathbb{R}^{n+1}$ with $\overline{s}^n \in \{0, \ldots, n-1\}$ such that incentive constraint $(IC_L^C(\hat{\gamma}_L^n))$ is satisfied. That is,

$$p_1^n(\hat{\rho}^n) - p_0^n(\hat{\rho}^n) \geq \frac{c}{\delta V(p_0^n(\hat{\rho}^n), \hat{\gamma}_L^n)}. \tag{25}$$

Because $p_0^n(\hat{\rho}^n) \geq 0$ and $\hat{\gamma}_L^n \leq 1/\lambda_L$, we have

$$V(p_0^n(\hat{\rho}^n), \hat{\gamma}_L^n) = \frac{\hat{\gamma}_L^n - c}{1 - \delta(1 - p_0^n(\hat{\rho}^n))} \leq \frac{1/\lambda_L - c}{1 - \delta},$$

so that the right-hand side of (25) is larger than some lower bound that is strictly larger than zero and independent of n.

We next argue that the left-hand side, $p_1^n - p_0^n$, of (25) goes to zero as n grows arbitrarily large. To see this, apply the recursive structure (4) to obtain

$$p_1^n(\hat{\rho}^n) - p_0^n(\hat{\rho}^n) = (1 - \alpha - \beta)[\mathbb{P}_{n-1}(\overline{s}^n - 1|0)(1 - \hat{\rho}_{\overline{s}^n}) + \mathbb{P}_{n-1}(\overline{s}^n|0)\hat{\rho}_{\overline{s}^n}]. \tag{26}$$

Since $\mathbb{P}_{n-1}(.|0)$ is a binomial distribution of $n-1$ trials with success probability $1-\alpha$, the individual probability $\mathbb{P}_{n-1}(s|0)$ goes to zero for any s as n grows arbitrarily large. Equation (26) then implies that $p_1^n(\hat{\rho}^n) - p_0^n(\hat{\rho}^n)$ goes to zero as n grows arbitrarily large, which implies that there exists an $\hat{n}$ such that (25) is violated for all $n > \hat{n}$. Consequently, there is an upper bound on n such that $\mathcal{P}^n$ has a solution. Since n is an integer, there is actually a lower upper bound $\overline{n}$ such that $\mathcal{P}^n$ has a solution if and only if $n < \overline{n}$. Note that if for parameters $(\alpha, \beta, c, \delta, \lambda_L, n)$ the program $\mathcal{P}^n$ has a solution then it also has a solution when δ increases, implying that $\overline{n}$ is increasing in δ. Q.E.D.

Proof of Proposition 9: We first prove the second statement (which implies the first statement) by showing that for any n, there is a $\overline{\delta}(n)$ such that for all δ that exceed $\overline{\delta}(n)$, it holds $\hat{n} > n$. To see this, fix n and take $\gamma_L = \tilde{\gamma}_L^n$. Lemma 2 implies that for $\delta > \overline{\delta}_0^n$, we have: $\hat{\rho}_0^n < 1$, $\hat{\rho}_i^n = 0$ for all $i = 1, \ldots, n$, $(IC_G^{n;1}(\tilde{\gamma}_G^n))$ holds with equality while all other $(IC_G^{n;k}(\tilde{\gamma}_G^n))$ are slack. Moreover, by Lemma 4, $\hat{\gamma}_L = \tilde{\gamma}_L^n$ so that $(IC_L^n(\tilde{\gamma}_L^n))$ holds with equality for $\hat{\rho}^n$. Using this latter equality yields that for any $\delta > \overline{\delta}_0^n$ we have

$$\hat{\rho}_0^n = \frac{c/\delta - c}{(\mathbb{P}_n(0|1) - \mathbb{P}_n(0|0))(\tilde{\gamma}_L^n - c) - \mathbb{P}_n(0|0)c} = \frac{c/\delta - c}{\alpha^{n-1}[(1 - \alpha - \beta)(\tilde{\gamma}_L^n - c) - \alpha c]}. \tag{27}$$

Note that the right-hand side converges to zero as δ goes to one. Hence, given n, we can find a $\overline{\delta}(n) < 1$ so that for all $\delta > \overline{\delta}(n)$, we have $\hat{\rho}_0^n < \alpha$.

We next argue that for all $\delta > \overline{\delta}(n)$, we must have $n \neq \hat{n}$, because, already for brand size $n+1$, we can, given δ, find a ρ^{n+1} that yields a strictly lower p_0^{n+1}. To see this, consider

$\rho^{n+1} = (\hat{\rho}_0^n/\alpha, 0, \ldots, 0)$ and $\gamma_L = \tilde{\gamma}_L^{n+1}$. It follows that

$$p_0^{n+1}(\rho^{n+1}) = \mathbb{P}_{n+1}(0|0)\hat{\rho}_0^n/\alpha = \alpha^{n+1}\hat{\rho}_0^n/\alpha = \alpha^n\hat{\rho}_0^n = \mathbb{P}_n(0|0)\hat{\rho}_0^n = p_0^n(\hat{\rho}^n)$$

and

$$p_1^{n+1}(\rho^{n+1}) = \mathbb{P}_{n+1}(0|1)\hat{\rho}_0^n/\alpha = \alpha^n(1-\beta)\hat{\rho}_0^n/\alpha = \alpha^{n-1}(1-\beta)\hat{\rho}_0^n = \mathbb{P}_n(0|1)\hat{\rho}_0^n = p_1^n(\hat{\rho}^n).$$

Hence,

$$\begin{aligned}[p_1^{n+1}(\rho^{n+1}) - p_0^{n+1}(\rho^{n+1})]\tilde{\gamma}_L^{n+1} - p_0^{n+1}(\rho^{n+1})c &= [p_1^n(\hat{\rho}^n) - p_0^n(\hat{\rho}^n)]\tilde{\gamma}_L^{n+1} - p_0^n(\hat{\rho}^n)c \\ &> [p_1^n(\hat{\rho}^n) - p_0^n(\hat{\rho}^n)]\tilde{\gamma}_L^n - p_0^n(\hat{\rho}^n)c = c/\delta - c,\end{aligned}$$

where the inequality holds because $\tilde{\gamma}_L^{n+1} > \tilde{\gamma}_L^n$. This implies that $(IC_L^{n+1}(\tilde{\gamma}_L^{n+1}))$ is slack. As $\gamma_L = \tilde{\gamma}_L^{n+1}$, the left-hand sides of $(IC_L^{n+1}(\tilde{\gamma}_L^{n+1}))$ and $(IC_G^{n+1;1}(\tilde{\gamma}_G^n))$ are equal to each other, implying that $(IC_G^{n+1;1}(\tilde{\gamma}_G^n))$ and, thus, all the other $(IC_G^{n+1;k}(\tilde{\gamma}_G^n))$ are slack as well. For the optimal $\hat{\rho}^{n+1}$ it therefore holds $p_0^{n+1}(\hat{\rho}^{n+1}) < p_0^n(\hat{\rho}^n)$. For $\delta > \overline{\delta}(n)$, brand size $n+1$ is therefore superior to brand size n. Since this argument holds for any n, it follows that $\hat{n}$ cannot be finite and $\lim_{\delta\to 1}\hat{n} = \infty$.

To see the final statement of the proposition, first fix n and $\gamma_L = \tilde{\gamma}_L^n$, and pick a δ arbitrarily close to 1, then, as discussed above, $\hat{\rho}^n$ is such that $\hat{\rho}_0^n$ equals (27) and $\hat{\rho}_i^n = 0$ for all $i = 1, \ldots, n$. The average per-period value can therefore be written as

$$\overline{v}^n = \frac{(1-\delta)(1-c)}{1-\delta+\delta\alpha^n\hat{\rho}_0^n} = \frac{(1-c)}{1+\frac{c\alpha}{(1-\alpha-\beta)(\tilde{\gamma}_L^n - c)-\alpha c}} = (1-c) - \frac{\alpha c(1-\lambda_L+\lambda_L n)}{(1-\alpha-\beta)n}.$$

Hence,

$$\lim_{n\to\infty}\overline{v}^n = 1 - c - \lambda_L\frac{\alpha c}{1-\alpha-\beta}.$$

Q.E.D.

Collective branding when the global player has to choose the same effort level in each market. In the following, we show that all of our propositions continue to hold if (under collective branding) the global player must choose, in each period t, the same effort level in all n markets so that $e_{G,i}^t = e_G^t$ for all i and t. As we argue below, for Proposition 1-8, this follows straightforwardly, and only for Proposition 9 do we have to adapt its proof slightly.

Let $\widehat{V}^C$ denote the resulting value in the best equilibrium. As discussed in the Conclusion, if the global player has to choose the same effort level in each market, then—under collective branding—the global player would no longer have n incentive constraints (IC_G^{Ck}), $k = 1, ..., n$, but only one, namely (IC_G^{Cn}). Because $\widehat{V}^C$ is the value of a more relaxed problem than $\overline{V}^C$, it immediately follows that $\widehat{V}^C \geq \overline{V}^C$.

Proposition 1 is concerned with equilibrium under independent branding and thus remains unaffected.

Proposition 2, which considers collective branding in the polar case $\lambda_G = 1$, also continues to hold: As the local player does not have to bear any effort cost, it is still optimal to give all of the rewards to the global player (so that $\gamma_G = 1$). Since $\widehat{V}^C \geq \overline{V}^C$, collective branding continues to be superior to independent branding, and strictly so if $\rho_0^I \leq 1$ (except possibly in the special case $n = 2$ and $\rho_0^I = 1$) or if $n > 2$ and ρ_0^I is larger than, but sufficiently close to, one.

Proposition 3 is concerned with equilibrium in the polar case in which the local player bears all of the effort cost ($\lambda_L = 1$) and the global player does not need to be incentivized; the proposition therefore remains unchanged.

Turning to Proposition 4, recall from the arguments in the main text that—under proportional rewards (i.e., $(\pi_G, \pi_L) = (\gamma_G, \gamma_L)$)—all of the global player's incentive constraints are slack whereas the local player's constraint is binding at the optimum. Hence, independently of whether or not the global player has to choose the same effort level in all markets, collective branding is suboptimal.

Proposition 5 also carries over: Since $\widehat{V}^C \geq \overline{V}^C$, collective branding still yields a strictly larger value in the best equilibrium than independent branding for all $\delta \geq \overline{\delta}$ (with the same value of $\overline{\delta}$ as before) and $\lambda_G > 0$.

To see that Proposition 6 goes through, note that its proof only relies on the observation that (IC_L^C) and at least one (IC_G^k), $k = 1, ..., n$, must be binding at the optimal reward-to-cost share ratio $\hat{\gamma}_L$. This still holds true when the global player has to choose the same effort level in all markets: in this case, (IC_L^C) and (IC_G^n) must be binding at the optimal $\hat{\gamma}_L$. (Of course, the optimal value $\hat{\gamma}_L$ may be different but that is irrelevant for the argument.)

Proposition 7 focuses on the case $\lambda_G = 1$ and $\delta > c$, stating that both the maximum implementable and the optimal brand size are unbounded, and efficiency obtains in the limit as the number of brands grows large. As the minimum sustainable on-path breakdown probability p_0 is weakly lower when (IC_G^k), $k < n$, can be ignored, and since the efficient payoff $1 - c$ is an upper bound on the per-period equilibrium value the proposition remains unchanged.

Next, Proposition 8 states that, if $\lambda_L > 0$, then the maximum sustainable brand size is finite and increasing in the discount factor. For the finiteness result, the proof only uses the incentive constraint of the local player and shows that it is violated for n large, even if all of the rewards are given to the local player ($\gamma_L = 1/\lambda_L$). The assertion on the maximum implementable brand size being increasing in the discount factor follows from the observation that any solution to the program of minimizing the on-path breakdown probability is still a solution when δ increases, which continues to hold true when (IC_G^k), $k < n$, can be ignored. Hence, the proposition carries over to the case where the global player has to choose the same effort level in all markets.

Finally, we turn to Proposition 9. When the global player has to choose the same effort in all markets, then in the best equilibrium, the reward-to-cost share ratios (for a given brand size n), $(\hat{\gamma}_L^n, \hat{\gamma}_G^n)$, are such that both (IC_G^{Cn}) and (IC_L^C) hold with equality. Hence,

$$\delta(p_n^n - p_0^n)V(p_0^n, \hat{\gamma}_G^n) = c = \delta(p_1^n - p_0^n)V(p_0^n, \hat{\gamma}_L^n) \;\Leftrightarrow\; (p_n^n - p_0^n)(\hat{\gamma}_G^n - c) = (p_1^n - p_0^n)(\hat{\gamma}_L^n - c).$$

Combining this with the identity $\lambda_G \gamma_G^n + \lambda_L \gamma_L^n = 1$ and $\lambda_G = 1 - \lambda_L$ yields

$$\hat{\gamma}_L^n = \frac{(p_n^n - p_0^n) - (p_n^n - p_1^n)(1 - \lambda_L)c}{p_1^n - p_0^n + (p_n^n - p_1^n)\lambda_L}$$

For δ large, we have $\bar{s}^n = 0$.[42] Hence,

$$p_0^n = \alpha^n \rho_0^n; p_1^n = \alpha^{n-1}(1-\beta)\rho_0^n; p_n^n = (1-\beta)^n \rho_0^n,$$

implying

$$\hat{\gamma}_L^n = \frac{((1-\beta)^n - \alpha^n) - ((1-\beta)^n - \alpha^{n-1}(1-\beta))(1-\lambda_L)c}{\alpha^{n-1}(1-\beta) - \alpha^n + ((1-\beta)^n - \alpha^{n-1}(1-\beta))\lambda_L} \tag{28}$$

$$= \frac{(1-x^n) - (1-x^{n-1})(1-\lambda_L)c}{x^{n-1} - x^n + (1-x^{n-1})\lambda_L}, \tag{29}$$

where $x^n \equiv \alpha^n/(1-\beta)^n$. Since $\alpha < 1 - \beta$, it follows that x^n goes to zero as $n \to \infty$. We thus obtain

$$\lim_{n\to\infty} \lim_{\delta\to 1} \hat{\gamma}_L^n = \frac{1 - (1-\lambda_L)c}{\lambda_L}$$

That is, as n grows unbounded and δ goes to one, $\hat{\gamma}_L^n$ and $\tilde{\gamma}_L^n$ converge to the same limit. From the proof of Proposition 9, it follows that the limiting per-period value is the same whether or not the global player has to choose the same effort in all markets.

[42] This follows from the proof of Lemma 2. When the global player has to choose the same effort in all markets, the critical discount factor above which $\bar{s}^n = 0$ is given by

$$\hat{\delta}_0^C = \max\left\{\frac{c}{c + n(\hat{\gamma}_G^n - c)\Delta_G(n,0)}, \frac{c}{c + (\hat{\gamma}_L^n - c)\Delta_L(n,0)}\right\}$$

.

Differentiating (29) w.r.t. n yields

$$\frac{\partial \hat{\gamma}_L^n}{\partial n} = \frac{-(1-c)(1-\lambda_L)(1-x)x^{n+1}ln(x)}{(\lambda_L x + x^n(1-\lambda_L - x))^2},$$

which is positive since $ln(x) < 0$ so that $\hat{\gamma}_L^n$ is strictly increasing in n. Using the same arguments as in the proof of Proposition 9, this implies that the optimal, as well as the maximum sustainable, brand size becomes unbounded as the discount factor goes to one.

Appendix B

In this appendix, we apply the abstract methods of decomposability and self-generation developed in Abreu, Pearce, and Stacchetti (1990) for the case of two markets, $n = 2$, under collective branding. This more formal analysis confims that the failure of the folk theorem is due to the dimensionality problem that the enforceable payoff vectors lie on a two dimensional plane within $\mathbb{R}^3$. Focusing on $n = 2$, allows us to study also asymmetric equilibrium outcomes and show that they are necessarily suboptimal for lower discount factors. In particular, we focus on the implementation of equilibrium outcomes in which players choose the cooperative actions $b_1 = e_{G,1} = e_{L,1} = b_2 = e_{G,2} = e_{L,2} = 1$ in the first period of the repeated game, both for symmetric and asymmetric perfect public equilibrium outcomes.

For fixed revenue shares $\pi_G = 1 - \pi_L$, the infinitely repeated game with imperfect monitoring has 5 players – the two local players, the global player, and the two consumers. We will refer to local player 1 as player 1, local player 2 as player 2, the global player as player 3, consumer 1 as player 4, and consumer 2 as player 5. Players 1, 2, and 3 are long-lived players, whereas the consumers, player 4 and 5, are short-lived. Except for the global player, all players have a binary action set, $A_1 = A_2 = A_4 = A_5 = \{0, 1\}$. The global player, as player 3, has an action set containing four actions that we can express as binary numbers, denoting in which of the two markets the global player picks effort: i.e., $A_3 = \{e_{G1}e_{G2}\}_{e_{G1},e_{G2}\in\{0,1\}} = \{00, 10, 01, 11\}$. Expressing player 3's action as a binary number, we can represent a pure action profile a as an element from $\{0, 1\}^6$ and the set of pure action profiles contains $2^6 = 64$ elements.

Because in equilibrium the short-lived players play myopic best replies, the set of *feasible* pure action profiles in the stage game of the overall repeated game is smaller. As explained in the main text, a consumer in market i buys if and only if the local and global player exert effort in market i. As a result, the set of feasible pure action profiles, $\mathbf{B}$, contains $2^4 = 16$ elements and for any feasible pure action profile consumers obtain a payoff of zero. Restricting attention to the set of feasible pure action profiles allows us to focus on the long-lived players, 1, 2, and 3, while ensuring equilibrium behavior of the short-run players.

Concerning the long-lived players, the feasible action profile $a = (e_{L1}, e_{L2}, e_{G1}, e_{G2}, b_1, b_2) \in \mathbf{B}$ yields the following stage payoffs to the three (long-lived) players:

$$u_1(a) = b_1\pi_L - \lambda_L e_{L1}c;\ u_2(a) = b_2\pi_L - \lambda_L e_{L2}c;\ u_3(a) = (b_1 + b_2)\pi_G - \lambda_G(e_{G1} + e_{G2})c.$$

Note that, restricted to the feasible pure action profiles in $\mathbf{B}$, each player can guarantee himself at least a zero payoff by not exerting any effort. Moreover, by not exerting any effort, any pair of players can ensure that the other player gets at most a zero payoff. Hence, the minmax-payoff of each player is zero.

The observable signals are the aggregated quality reports $s = s_1 + s_2 \in \{0, 1, 2\}$ and the uniformly distributed public correlation device $r \in [0, 1]$. Given the action profile a, the perfect complementarity of efforts imply that the probability of signal $s_i \in \{0, 1\}$ in market i is

$$\mathbb{P}\{s_i = 1|a\} = (1 - \alpha)a^i a^{i+2} + \beta(1 - a^i a^{i+2}),$$

where a^i is the $i-$th element of the action profile $a = (e_{L1}, e_{L2}, e_{G1}, e_{G2}, b_1, b_2) \in \mathbf{B}$. Following Mailath and Samuelson (2006,p.253), we combine the signal s and r into one continuous signal y by defining

$$y = s_1 + s_2 + r \in Y \equiv [0, 3].$$

We denote the density of this continuous signal over the support Y by $\rho(y|a)$. Since the distribution of r is uniform, the density $\rho(y|a)$ is the step function:

$$\rho(y|a) = \begin{cases} \mathbb{P}\{s = 0|a\} & \text{, if } y \in [0, 1) \\ \mathbb{P}\{s = 1|a\} & \text{, if } y \in [1, 2) \\ \mathbb{P}\{s = 2|a\} & \text{, if } y \in [2, 3]. \end{cases}$$

Following Abreu, Pearce, and Stacchetti (1990), we define an action profile $a \in \mathbf{B}$ as enforceable on $\mathcal{W} \subset \mathbb{R}^3_+$ if there exists a (Lebesgue measurable) mapping $\gamma : Y \to \mathcal{W}$ such that for any $i = 1, 2, 3$,

$$\begin{aligned} V_i(a, \gamma) &\equiv (1 - \delta)u_i(a) + \delta \int_0^3 \gamma_i(y)\rho(y|a)dy \\ &\geq (1 - \delta)u_i(a_i', a_{-i}) + \delta \int_0^3 \gamma_i(y)\rho(y|a_i', a_{-i})dy \text{ for all } a_i' \in A_i. \end{aligned}$$

Note that, given the density $\rho(y|a)$, it holds for any $a \in \mathbf{B}$ which is enforceable on $\mathcal{W}$ that

$$\int_0^3 \gamma_i(y)\rho(y|a)dy = \sum_{j=0}^{2} \int_j^{j+1} \gamma_i(y)\rho(y|a)dy = \sum_{j=0}^{2} \mathbb{P}\{s = j|a\} \int_j^{j+1} \gamma_i(y)dy = \sum_{j=0}^{2} \mathbb{P}\{s = j|a\} w_{j+1},$$

where $w_{j+1} = \int_j^{j+1} \gamma_i(y)dy$ lies in the convex hull of $\mathcal{W}$. Consequently, the definition of enforceability in our framework is equivalent to saying that an action profile $a \in \mathbf{B}$ is enforceable on $\mathcal{W} \in \mathbb{R}^3_+$ if there exists a triple w_1, w_2, w_3 in the convex hull of $\mathcal{W}$ such that

$$\begin{aligned} V_i(a,\gamma) &\equiv (1-\delta)u_i(a) + \delta \sum_{j=0}^{2} \mathbb{P}\{s=j|a\}w_{j+1} \\ &\geq (1-\delta)u_i(a_i', a_{-i}) + \delta \sum_{j=0}^{2} \mathbb{P}\{s=j|a_i', a_{-i}\}w_{j+1} \text{ for all } a_i' \in A_i. \end{aligned} \tag{30}$$

The perfect complementarity in the effort levels implies that any action profile for which, in some market, effort is only supplied by one player is not enforceable. As a result, only the following 4 action profiles of the 16 feasible action profile in $\mathbf{B}$ are enforceable:

$$(0,0,00,0,0); (1,0,10,1,0); (0,1,01,0,1); (1,1,11,1,1);$$

with associated enforceable payoff vectors

$$\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}; \begin{pmatrix} u_L \\ 0 \\ u_G \end{pmatrix}; \begin{pmatrix} 0 \\ u_L \\ u_G \end{pmatrix}; \begin{pmatrix} u_L \\ u_L \\ 2u_G \end{pmatrix},$$

where $u_L = \pi_L - \lambda_L c$ and $u_G = \pi_G - \lambda_G c$.

Abreu, Pearce, and Stacchetti (1990) show that the set of equilibrium payoffs, $\mathcal{E}(\delta)$, is a subset of the convex hull of the enforceable payoff vectors. Because these payoff vectors lie on a two dimensional plane within $\mathbb{R}^3$, we can express any point $w \in \mathbb{R}^3$ in the convex hull of these four points by a unique pair of scalars $(\mu_1, \mu_2) \in [0,1]^2$ such that

$$w = \hat{w}(\mu_1, \mu_2) \equiv \mu_1 \begin{pmatrix} u_L \\ 0 \\ u_G \end{pmatrix} + \mu_2 \begin{pmatrix} 0 \\ u_L \\ u_G \end{pmatrix}.$$

Hence, the convex hull is the set $\hat{W} \equiv \{\hat{w}(\mu_1, \mu_2) | \mu_1, \mu_2 \in [0,1]\}$ and the aggregate payoff associated with any point $\hat{w}(\mu_1, \mu_2)$ is $V(\mu_1, \mu_2) = (\mu_1 + \mu_2)(u_L + u_G)$. As a result, the equilibrium payoff that maximizes aggregate payoffs is attained by a solution (μ_1^*, μ_2^*) of the following program:

$$\mathcal{P} : \max_{\mu_1, \mu_2} V(\mu_1, \mu_2) \text{ s.t. } \hat{w}(\mu_1, \mu_2) \in \mathcal{E}(\delta).$$

Apart from the fact that $\mathcal{E}(\delta) \subset \hat{W}$, the presence of the correlation device r implies that the set $\mathcal{E}(\delta)$ is convex. Moreover, since the two markets are symmetric, $\mathcal{E}(\delta)$ exhibits the symmetry that $w(\mu_1, \mu_2) \in \mathcal{E}(\delta)$ implies $w(\mu_2, \mu_1) \in \mathcal{E}(\delta)$. From the convexity and this symmetry of $\mathcal{E}(\delta)$, it follows that there is a symmetric solution $\mu_1^* = \mu_2^* = \mu^*$ to problem $\mathcal{P}$ with an associated payoff vector w^* that is symmetric in the sense that $w^* = \hat{w}(\mu, \mu)$ for some $\mu \in [0, 1]$. Note also that for any $w = (w_1, w_2, w_3) \in \mathcal{E}(\delta)$ we have $w = \hat{w}(w_1/u_L, w_2/u_L)$ so that be defining

$$M(\delta) \equiv \{(w_1/u_L, w_2/u_L)|(w_1, w_2, w_3) \in \mathcal{E}(\delta)\},$$

we obtain an equivalent representation of $\mathcal{E}(\delta)$ in terms of pairs (μ_1, μ_2).

Implementation of $a^{11} \equiv (1, 1, 1, 1, 1, 1)$ in the first period. We next study an equilibrium in which a^{11} is the aggregate-payoff-maximizing strategy profile. We first argue that, in this case, the action profile $a^{11} \equiv (1, 1, 1, 1, 1, 1)$ in the first period must also lead to a symmetric equilibrium payoff $w \in \mathcal{E}(\delta)$. To see this, note first that for the action profile a^{11} to be implementable in equilibrium it has to be enforceable on $\mathcal{E}(\delta)$. The convexity of $\mathcal{E}(\delta)$ implies it is equal to its convex hull so that the requirement is that we have to find three equilibrium payoffs $w_1, w_2, w_3 \in \mathcal{E}(\delta)$ such that (30) holds for each long-lived player. Since any equilibrium value $w \in \mathcal{E}(\delta)$ corresponds to a (unique) pair $(\mu_1, \mu_2) \in [0, 1]$ such that $w = \hat{w}(\mu_1, \mu_2)$, finding three equilibrium values is equivalent to finding three pairs $\mu^0 = (\mu_1^0, \mu_2^0)$, $\mu^1 = (\mu_1^1, \mu_2^1)$, $\mu^2 = (\mu_1^2, \mu_2^2)$ in $M(\delta)$ such that (30) holds for each long-lived player with $w_j = \hat{w}(\mu^j)$.

With respect to player 1, (30) is

$$(1-\delta)u_L + \delta \sum_{j=0}^{2} \mathbb{P}\{s = j|a^{11}\}\hat{w}_1(\mu^j) \geq (1-\delta)(u_L + \lambda_L c) + \delta \sum_{j=0}^{2} \mathbb{P}\{s = j|0, a_{-1}^{11}\}\hat{w}_1(\mu^j). \quad (31)$$

With respect to player 2, (30) is

$$(1-\delta)u_L + \delta \sum_{j=0}^{2} \mathbb{P}\{s = j|a^{11}\}\hat{w}_2(\mu^j) \geq (1-\delta)(u_L + \lambda_L c) + \delta \sum_{j=0}^{2} \mathbb{P}\{s = j|0, a_{-2}^{11}\}\hat{w}_2(\mu^j).$$

With respect to player 3, (30) leads to three conditions of which, due to $\mathbb{P}\{s = j|01, a_{-3}^{11}\} = \mathbb{P}\{s = j|10, a_{-3}^{11}\}$, the latter two coincide:

$$(1-\delta)2u_G + \delta \sum_{j=0}^{2} \mathbb{P}\{s = j|a^{11}\}\hat{w}_3(\mu^j) \geq (1-\delta)2(u_G + \lambda_G c) + \delta \sum_{j=0}^{2} \mathbb{P}\{s = j|00, a_{-3}^{11}\}\hat{w}_3(\mu^j),$$

$$(1-\delta)2u_G + \delta \sum_{j=0}^{2} \mathbb{P}\{s = j|a^{11}\}\hat{w}_3(\mu^j) \geq (1-\delta)(2u_G + \lambda_G c) + \delta \sum_{j=0}^{2} \mathbb{P}\{s = j|10, a_{-3}\}\hat{w}_3(\mu^j),$$

$$(1-\delta)2u_G+\delta\sum_{j=0}^{2}\mathbb{P}\{s=j|a^{11}\}\hat{w}_3(\mu^j)\geq(1-\delta)(2u_G+\lambda_G c)+\delta\sum_{j=0}^{2}\mathbb{P}\{s=j|01,a^{11}_{-3}\}\hat{w}_3(\mu^j).$$

Combining and rewriting these conditions, we get local player 1's incentive constraint

$$\frac{1-\delta}{\delta}\frac{\lambda_L c}{(1-\alpha-\beta)u_L}\leq-\alpha\mu_1^0+(2\alpha-1)\mu_1^1+(1-\alpha)\mu_1^2; \tag{32}$$

local player 2's incentive constraint

$$\frac{1-\delta}{\delta}\frac{\lambda_L c}{(1-\alpha-\beta)u_L}\leq-\alpha\mu_2^0+(2\alpha-1)\mu_2^1+(1-\alpha)\mu_2^2; \tag{33}$$

the global player's incentive constraint not to shirk in both markets:

$$\frac{2(1-\delta)}{\delta}\frac{\lambda_G c}{(1-\alpha-\beta)u_G}\leq-(1-\beta+\alpha)(\mu_1^0+\mu_2^0)+2(\alpha-\beta)(\mu_1^1+\mu_2^1)+(1-\alpha+\beta)(\mu_1^2+\mu_2^2); \tag{34}$$

and the global player's incentive constraint not to shirk in only one market:

$$\frac{1-\delta}{\delta}\frac{\lambda_G c}{(1-\alpha-\beta)u_G}\leq-\alpha(\mu_1^0+\mu_2^0)+(2\alpha-1)(\mu_1^1+\mu_2^1)+(1-\alpha)(\mu_1^2+\mu_2^2). \tag{35}$$

The action profile a^{11} is implementable if a triple of pairs $\overline{\mu}=(\overline{\mu}^0,\overline{\mu}^1,\overline{\mu}^2)$ in $M(\delta)$ exist that together satisfy the constraints (32), (33), (34), (35). In this case, implementing a^{11} in the first period with continuation payoffs $\hat{w}^0(\mu^0)$, $\hat{w}^1(\mu^1)$, $\hat{w}^2(\mu^2)$ yields an aggregate payoff of

$$V(\mu^0,\mu^1,\mu^2)=(u_G+u_L)[2(1-\delta)+\delta\{\alpha^2(\mu_1^0+\mu_2^0)+2\alpha(1-\alpha)(\mu_1^1+\mu_2^1)+(1-\alpha)^2(\mu_1^2+\mu_2^2)\}] \tag{36}$$

Hence, the triple of pairs, $\overline{\mu}=(\overline{\mu}^0,\overline{\mu}^1,\overline{\mu}^2)$, that maximizes aggregate payoffs from an equilibrium strategy that implements a^{11} in the first period is a solution to the following program

$$\mathcal{P}(a^{11}):\max_{\mu=(\mu^0,\mu^1,\mu^2)\in M(\delta)}V(\mu)\text{ s.t. }(32),(33),(34),(35).$$

If a solution to $\mathcal{P}(a^{11})$ exists, then there is one that is symmetric in the sense that $(\overline{\mu}_1^0,\overline{\mu}_1^1,\overline{\mu}_1^2)=(\overline{\mu}_2^0,\overline{\mu}_2^1,\overline{\mu}_2^2)$, because for any asymmetric solution $\overline{\mu}$ its symmetric average $\tilde{\mu}$ with $\tilde{\mu}_1^i=\tilde{\mu}_1^i=(\overline{\mu}_1^i+\overline{\mu}_2^i)/2$ has the same objective value V, lies in $M(\delta)$ (due to the convexity of $\mathcal{E}(\delta)$), and also satisfies all constraints (since the original $\overline{\mu}$ does so).

Using this observation, program $\mathcal{P}(a^{11})$ simplifies to finding three scalars (μ_a,μ_b,μ_c) with $(\mu_a,\mu_a),(\mu_b,\mu_b),(\mu_c,\mu_c)\in M(\delta)$ that maximize

$$W\;=\;(u_G+u_L)[2(1-\delta)+2\delta\{\alpha^2\mu_a+2\alpha(1-\alpha)\mu_b+(1-\alpha)^2\mu_c\}]\text{ s.t.} \tag{37}$$

$$\frac{1-\delta}{\delta}\frac{\lambda_L c}{(1-\alpha-\beta)u_L} \leq -\alpha\mu_a + (2\alpha-1)\mu_b + (1-\alpha)\mu_c; \tag{38}$$

$$\frac{1-\delta}{\delta}\frac{\lambda_G c}{(1-\alpha-\beta)u_G} \leq -(1-\beta+\alpha)\mu_a + 2(\alpha-\beta)\mu_b + (1-\alpha+\beta)\mu_c; \tag{39}$$

$$\frac{1-\delta}{\delta}\frac{\lambda_G c}{2(1-\alpha-\beta)u_G} \leq -\alpha\mu_a + (2\alpha-1)\mu_b + (1-\alpha)\mu_c. \tag{40}$$

Implementation of $a^{10} \equiv (1,0,1,0,1,0)$ in the first period We next study the implementability and optimality of the asymmetric action profile $a^{10} \equiv (1,0,1,0,1,0)$ in the first period. If the equilibrium that maximizes aggregate payoffs is such that it implements the asymmetric action $a^{10} \equiv (1,0,1,0,1,0)$ in the first period, then the equilibrium attains the value $V^* = V(\mu_1^* + \mu_2^*) = (u_G + u_L)(\mu_1^* + \mu_2^*)$. Moreover, it requires that the action a^{10} is enforceable in $\mathcal{E}(\delta)$. Hence, there must be three pairs $\mu^0 = (\mu_1^0, \mu_2^0)$, $\mu^1 = (\mu_1^1, \mu_2^1)$, $\mu^2 = (\mu_1^2, \mu_2^2)$ in $M(\delta)$ such that (30) holds for $a = a^{10}$ for each long-lived player with $w_j = \hat{w}(\mu^j)$. That is for player 1, we have

$$(1-\delta)u_L + \delta\sum_{j=0}^{2}\mathbb{P}\{s=j|a^{10}\}\hat{w}(\mu^j) \geq (1-\delta)(u_L + \lambda_L c) + \delta\sum_{j=0}^{2}\mathbb{P}\{s=j|0,a_{-1}^{10}\}\hat{w}(\mu^j). \tag{41}$$

For player 2, we have

$$(1-\delta)0 + \delta\sum_{j=0}^{2}\mathbb{P}\{s=j|a^{10}\}\hat{w}(\mu^j) \geq -(1-\delta)\lambda_L c + \delta\sum_{j=0}^{2}\mathbb{P}\{s=j|1,a_{-2}^{10}\}\hat{w}(\mu^j). \tag{42}$$

Since $\mathbb{P}\{s=j|a^{10}\} = \mathbb{P}\{s=j|1,a_{-1}^{10}\}$ for all $j=0,1,2$, constraint (42) is satisfied for any triple μ^0,μ^1,μ^2. For player 3, we have

$$(1-\delta)u_G + \delta\sum_{j=0}^{2}\mathbb{P}\{s=j|a^{10}\}\hat{w}(\mu^j) \geq (1-\delta)(u_G + \lambda_G c) + \delta\sum_{j=0}^{2}\mathbb{P}\{s=j|00,a_{-3}^{10}\}\hat{w}(\mu^j), \tag{43}$$

$$(1-\delta)u_G + \delta\sum_{j=0}^{2}\mathbb{P}\{s=j|a^{10}\}\hat{w}(\mu^j) \geq (1-\delta)(u_G - \lambda_G c) + \delta\sum_{j=0}^{2}\mathbb{P}\{s=j|11,a_{-3}^{10}\}\hat{w}(\mu^j), \tag{44}$$

$$(1-\delta)u_G + \delta\sum_{j=0}^{2}\mathbb{P}\{s=j|a^{10}\}\hat{w}(\mu^j) \geq (1-\delta)u_G + \delta\sum_{j=0}^{2}\mathbb{P}\{s=j|01,a_{-3}^{10}\}\hat{w}(\mu^j). \tag{45}$$

The second inequality (44) holds for any triple (μ^0,μ^1,μ^2), since $\mathbb{P}\{s=j|a^{10}\} = \mathbb{P}\{s=j|11,a_{-3}^{10}\}$ for each $j=0,1,2$. Moreover, (43) implies (45), since $\mathbb{P}\{s=j|00,a_{-3}^{10}\} = \mathbb{P}\{s=j|01,a_{-3}^{10}\}$ for each $j=0,1,2$. Hence, a^{10} is enforceable on $\mathcal{E}(\delta)$ if and only if we find a triple (μ^0,μ^1,μ^2) in $M(\delta)$ such that (41) and (43) hold.

Rewriting (41) yields the retailer's incentive constraint

$$\frac{1-\delta}{\delta}\frac{\lambda_L c}{(1-\alpha-\beta)u_L} \leq -(1-\beta)\mu_1^0 + (1-2\beta)\mu_1^1 + \beta\mu_1^2. \tag{46}$$

Rewriting (43) yields the global player's incentive constraint

$$\frac{1-\delta}{\delta}\frac{\lambda_G c}{(1-\alpha-\beta)u_G} \leq -(1-\beta)(\mu_1^0+\mu_2^0) + (1-2\beta)(\mu_1^1+\mu_2^1) + \beta(\mu_1^2+\mu_2^2). \tag{47}$$

The aggregate payoff associated with the strategy profile a^{10} that is enforceable on $\mathcal{E}(\delta)$ by (μ^0, μ^1, μ^2) is

$$W(\mu^0,\mu^1,\mu^2) = (u_G+u_L)[(1-\delta)+\delta\{(1-\beta)\alpha(\mu_1^0+\mu_2^0)+[(1-\beta)(1-\alpha)+\alpha\beta](\mu_1^1+\mu_2^1)+(1-\alpha)\beta(\mu_1^2+\mu_2^2)\}]$$

It follows that the equilibrium with the maximum aggregate payoffs that implements the asymmetric action a^{10} in the first period is a solution $(\hat{\mu}^0, \hat{\mu}^1, \hat{\mu}^2)$ to the following linear program:

$$\mathcal{P}^{10}: \max_{\mu^0,\mu^1,\mu^2 \in M(\delta)} \quad W(\mu^0,\mu^1,\mu^2) \text{ s.t. } (46), (47).$$

Denote the value of this program as $\hat{W}^{10} = W(\hat{\mu}^0, \hat{\mu}^1, \hat{\mu}^2)$.

If the equilibrium that maximizes aggregate payoffs is such that it implements the asymmetric action a^{10} in the first period then it holds $\hat{W}^{10} = V^*$.

Comparison of a^{11} and a^{10}. We first show that for δ small, the action profile a^{11} is optimal whenever it is implementable.

Lemma 7. *Suppose $\delta \leq 1/2$ and a^{11} is implementable. Then implementing a^{11} is optimal.*

Proof. If a^{10} is not implementable in that no combination $(\hat{\mu}^0, \hat{\mu}^1, \hat{\mu}^2)$ in $M(\delta)$ exists that satisfies (46) and (47), then the result follows trivially, since the only other implementable action profile $a^{00} = (0,0,0,0,0,0)$, which yields aggregate payoffs of 0, which is weakly less than any aggregates payoffs from a a triple of pairs $\overline{\mu} = (\overline{\mu}^0, \overline{\mu}^1, \overline{\mu}^2)$ in $M(\delta)$ that implements profile a^{11}.

So suppose a^{10} is implementable. We next demonstrate that for $\delta \leq 1/2$, the aggregate payoffs associated with any triple $(\hat{\mu}^0, \hat{\mu}^1, \hat{\mu}^2)$ in $M(\delta)$ that implements a^{10}, the aggregate payoffs are less than $2(u_G+u_L)(1-\delta)$, a lower bound on the payoffs of implementing a^{11} when it is implementable. To show this, consider the relaxed version of program $\mathcal{P}^{10}$ in which we disregard (46). Denoting this relaxed program as $\tilde{\mathcal{P}}^{10}$ and its value as $\tilde{W}^{10}$, it follows $\tilde{W}^{10} \geq \hat{W}^{10}$. The relaxed program has constraint (47) binding, since disregarding this constraint yields a solution with $\mu_1^0+\mu_2^0 = \mu_1^1+\mu_2^1 = \mu_1^2+\mu_2^2$ which violates (47).

A binding (47) implies

$$(1-\beta)(\mu_1^0+\mu_2^0) = (1-2\beta)(\mu_1^1+\mu_2^1) + \beta(\mu_1^2+\mu_2^2) - \frac{1-\delta}{\delta}\frac{\lambda_G c}{(1-\alpha-\beta)u_G}. \tag{48}$$

Substituting out the expression $(\mu_1^0+\mu_2^0)$, program $\tilde{\mathcal{P}}^{10}$ rewrites as

$$\max_{\mu^1,\mu^2\in M(\delta)} (u_G+u_L)\left\{1-\delta+\delta\left[(1-\beta)(\mu_1^1+\mu_2^1)+\beta(\mu_1^2+\mu_2^2)-\alpha\frac{1-\delta}{\delta}\frac{\lambda_G c}{(1-\alpha-\beta)u_G}\right]\right\}.$$

The solution to this exhibits $\mu_1^1+\mu_2^1=\mu_1^1+\mu_2^1=\mu_1^*+\mu_2^*$ and, hence,[43]

$$\tilde{W}^{10} \leq (u_G+u_L)\left\{1-\delta+\delta\left[(\mu_1^*+\mu_2^*)-\alpha\frac{1-\delta}{\delta}\frac{\lambda_G c}{(1-\alpha-\beta)u_G}\right]\right\}$$

Hence, $\tilde{W}^{10}\geq V^*$ implies

$$(u_G+u_L)\left\{1-\delta+\delta\left[(\mu_1^*+\mu_2^*)-\alpha\frac{1-\delta}{\delta}\frac{\lambda_G c}{(1-\alpha-\beta)u_G}\right]\right\} \geq (u_G+u_L)(\mu_1^*+\mu_2^*)$$

so that

$$\mu_1^*+\mu_2^* \leq 1-\alpha\frac{\lambda_G c}{(1-\alpha-\beta)u_G}$$

Hence, if the equilibrium that maximizes aggregate payoffs is such that it implements the asymmetric action $a^{10}\equiv(1,0,1,0,1,0)$ in the first period, then we have $\mu_1^*+\mu_2^*\leq 1$. As a consequence, the maximum aggregate payoffs in $\mathcal{E}(\delta)$ is smaller than (u_G+u_L), which for $\delta\leq 1/2$ is smaller than $2(u_G+u_L)(1-\delta)$, a lower bound on the payoffs of implementing a^{11} when it is implementable. □

References

Abreu, D., D. Pearce, and E. Stacchetti (1990). "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring." Econometrica 58: 1041-1063.

Andersson, F. (2002). "Pooling Reputations." International Journal of Industrial Organization 20(5): 715–730.

Bar-Isaac, H., and S. Tadelis (2008). "Seller Reputation." Foundations and Trends in Microeconomics 4: 273–351.

Bhattacharyya, S. and F. Lafontaine (1995). "Double-Sided Moral Hazard and the Nature of Share Contracts." RAND Journal of Economics 26:761–781.

[43] The inequality is due to the fact that the right hand value is only attained if after substituting $\mu_1^1+\mu_2^1=\mu_1^1+\mu_2^1=\mu_1^*+\mu_2^*$ into (48) yields a $\mu^0\in\mathcal{E}(\delta)$, otherwise the value $\tilde{W}^{10}$ is smaller than the RHS.

Blair, R., and Lafontaine (2005). The Economics of Franchising. Cambridge University Press.

Biscarri, W., S. Zhao, and R. Brunner (2018) “A simple and fast method for computing the Poisson binomial distribution function.” Computational Statistics & Data Analysis 122: 92–100.

Cabral, L. (2000). “Stretching Firm and Brand Reputation.” RAND Journal of Economics 31(4): 658–673.

Cabral, L. (2009). “Umbrella Branding with Imperfect Observability and Moral Hazard.” International Journal of Industrial Organization 27: 206–213.

Cai, H., and I. Obara (2009). “Firm Reputation and Horizontal Integration.” RAND Journal of Economics 40(2): 340-363.

Calboli, I. (2007). “The Sunset of ”Quality Control” in Modern Trademark Licensing.” American University Law Review 57: 341–407.

Castriota S., and M. Delmastro (2012). “Seller Reputation: Individual, Collective, and Institutional Factors.” Journal of Wine Economics 7: 49–69.

Choi, J. (1998). “Brand Extension and Information Leverage.” Review of Economic Studies 65: 655–669.

Fishman, A., I. Finkelstein, A. Simhon, and N. Yacouel (2018). “Collective brands.” International Journal of Industrial Organization 59: 316–339.

Fleckinger, P. (2014). “Regulating Collective Reputation.” Mimeo CERNA, MINES ParisTech.

Fudenberg, D., D. Levine, E. Maskin (1994). “The Folk Theorem with Imperfect Public Information.” Econometrica 62 (5), 997-1039.

Hadfield, G. (1990). “Problematic Relations: Franchising and the Law of Incomplete Contracts.” Stanford Law Review 42: 927–992.

Hakenes, H., and M. Peitz (2008). “Umbrella branding and the provision of quality.” International Journal of Industrial Organization 26: 546–556.

Holmström, B. (1982). “Moral Hazard in Teams.” The Bell Journal of Economics, 13: 324–340.

Klein, B., and K. Leffler, (1981). “The role of market forces in assuring contractual performance.” Journal of Political Economy 89: 615–641.

Klein, B., and L. Saft. (1985). “The Law and Economics of Franchise Tying Contracts.” The Journal of Law & Economics, 28: 345–361.

Kotler, P. (2003). Marketing Management. 11th Edition, Prentice-Hall, Upper Saddle River.

Kreps, D., and R. Wilson (1982). “Reputation and Imperfect Information.” Journal of Economic Theory 27: 253-279.

Levin, J. (2003). “Relational Incentive Contracts.” American Economic Review 93 (3), 835-857.

Mailath, G., and L. Samuelson (2006). Repeated Games and Reputations: Long-Run Relationships. Oxford University Press.

Matsushima, H. (2001). “Multimarket Contact, Imperfect Monitoring, and Implicit Collusion.” Journal of Economic Theory 98: 158–178.

Miklós-Thal, J. (2012). “Linking reputations through collective branding.” Quantitative Marketing and Economics 10: 335–374.

Milgrom, P., and J. Roberts (1982). “Predation, Reputation, and Entry Deterrence.” Journal of Economic Theory 27, 280-312.

Moorthy, S. (2012). “Can brand extension signal product quality?” Marketing science 31: 756-770.

Neeman, Z, A. Öry, and J. Yu (2019). “The benefit of collective reputation.” RAND Journal of Economics 50: 787–821.

Nosko, C., and S. Tadelis (2015). “The Limits of Reputation in Platform Markets: An Empirical Analysis and Field Experiment.” Mimeo UC Berkeley.

Radner, R., R. Myerson, and E. Maskin (1986). “An Example of a Repeated Partnership Game with Discounting and with Uniformly Inefficient Equilibria.” The Review of Economic Studies 53: 59–69.

Tirole, J. (1996). “A Theory of Collective Reputation (with Application to Corruption and Firm Quality).” Review of Economic Studies 63: 1–22.

Rukhin, A., C. Priebe, and D. Healy Jr. (2009). “On the monotone likelihood ratio property for the convolution of independent binomial random variables.” Discrete Applied Mathematics 157: 2562–2564.

Wang, Y. (1993). “On the number of successes in independent trials.” Statistica Sinica, 3: 295–312.

Wernerfelt, B. (1988). “Umbrella branding as a signal of new product quality: An example of signalling by posting a bond.” The RAND Journal of Economics 30: 458–466.

Winfree, J., and J. McCluskey (2005). “Collective Reputation and Quality.” American Journal of Agricultural Economics 87: 206–213.