Submitted to *Organization Science*
manuscript OS-MS-19-12480.R2

When Reinforcing Processes Generate an Outcome-Quality Dip

When does market success indicate superior merit? We show that when consumer choices between products with equal prices depend on quality but also on past popularity, more popular products are not necessarily of higher quality. Rather, a medium level of popularity may be associated with lower quality than lower levels of popularity. Using a formal model we show that this kind of non-monotonic association occurs when reinforcing processes are strong. More generally, a dip can occur when outcomes depend on both quality and resources and the latter are allocated bimodally, with some being given a lot of resources and most receiving little. Empirically, we illustrate that such a dip occurs in the association between movie theater sales and ratings. The presence of a dip in the outcome-quality association complicates learning from market outcomes and evaluation of individuals and new ventures, challenges the legitimacy of stratification systems, and creates opportunities for sophisticated evaluators who understand the dip.

Key words: reinforcing processes, quality, performance evaluation, movie industry, luck

1. Introduction

What can we learn about management and organizations from market outcomes? Are firms with a high market share more likely to have superior management and effective routines than firms with a low market share? In particular, consider a market in which the prices of competing products are identical, for example where competing apps are free or all music singles cost the same. Organizations compete by trying to develop the best and most appealing product. Suppose consumers care not only about product quality, but also about popularity. They prefer to select a higher-quality product, but may select a lower-quality one if it happens to be very popular. Should we expect products with higher shares of such markets to be of higher quality? The issue is not whether a low-quality product may occasionally get lucky and end up with a high market share, as it is well known that this can happen if choices are stochastic and subject to reinforcing processes (Arthur 1989, Lynn et al. 2009). Rather, the question is, should an informed observer expect products with a higher market share to be, on average, of *lower* quality?

The answer to this question has implications for evaluating and learning from the success of new products, ventures and individuals. If a higher market share indicates higher quality, market outcomes can be used as a proxy for quality in evaluation and selection processes. It makes sense to reward individuals and firms with more successful market outcomes. However, if products with a higher market share cannot be assumed to be of higher quality, learning from observed outcomes and evaluating individuals and firms become more complicated. A "more-successful-is-better" heuristic will be systematically misleading.

In this paper, we show that a higher market share may indicate lower expected quality when consumers care about quality but also about popularity. Using a simple model of consumer choice inspired by Arthur et al. (1987)'s non-linear Pólya urn model, we show theoretically that expected quality and market share may vary non-monotonically. When the market share is low, expected quality is an increasing function of the market share: a higher market share indicates higher expected quality. This association becomes negative at a medium-level market share, when a higher market share indicates lower expected quality. However, when the market share is high, the association becomes positive again, as a higher market share indicates higher expected quality.

A dip in the outcome-quality association occurs when strong reinforcing processes imply that a product's market share is converging to a value that is either high (when the product happened to be chosen early on) or low (when it was not chosen early on). Strong reinforcement implies that low-quality products chosen early on may obtain a higher market share than superior products that were not chosen early on. The dip occurs because low-quality products either reach a very low market share (when they were not chosen early on) or a medium-level market share (when they were chosen early on). High-quality products in contrast reach either a below median-level market share (when they were not chosen early on) or a high market share (when they were chosen early on). The intuition behind the dip is that reaching a medium-level market share indicates that the product was chosen early on and could do well. If the product only reached a medium-level market share, despite having this initial advantage, it is likely a low-quality product.

More generally, we show that a dip can occur when resources are allocated bimodally. Reinforcing processes are one way of generating bimodal allocations: when there is a strong reinforcing process, a product may end up with a high or low installed base. In many settings, evaluators allocate most resources to a few promising alternatives, generating a similar bimodal allocation of resources. When the outcome depends on resources as well as quality, a dip may occur.

A dip is not only a theoretical possibility. We show that a dip occurs in the association between movie sales and movie ratings in the action/adventure genres (ratings are based on movies' average scores on the Internet Movie Database website, www.imdb.com). Being associated with best-selling movies is increasingly important for many actors, directors and writers (Baker and Faulkner 1991, Elberse 2013), but movie sales are difficult to predict (De Vany 2004). Sales are also subject to reinforcing processes, both before a movie is launched (e.g., movies perceived as likely hits receive substantially more investment and are launched at many more theaters) and after its opening week (e.g., movies that do well during the opening week continue to be screened, while those that do poorly are not). We find that the association between sales and ratings is non-monotonic, as average ratings first increase with sales, then decrease with sales, and finally increase again at high sales levels. Consistent with our theoretical mechanism, the data also show evidence of a separation into two categories: "advantaged/lucky" movies that receive substantial investment and are screened in many theaters, lifting their minimum sales; and "disadvantaged/unlucky" movies that receive less investment and are screened in few theaters, limiting their sales potential. In contrast, we do not find a dip in the drama genre, where reinforcing processes are weaker because these movies are launched in considerably fewer theaters.

Our findings demonstrate how a bimodal allocation of resources, typically generated by a reinforcement process, may introduce a fundamental decoupling of outcomes and quality. In contrast to previous studies, which have shown that the correlation between quality and outcomes may become close to zero as a result of reinforcing processes (Arthur 1989, Carroll and Harrison 1994, Chase 1980, Lynn et al. 2009), our model shows that the correlation may become negative for some outcome intervals. The implication is that naïve reliance on a more-successful-is-better heuristic may lead to unmeritocratic resource allocations. Nevertheless, such inefficiencies create opportunities for more informed evaluators who understand the mechanism that generates a dip (Cattani et al. 2018, Denrell et al. 2019). Our finding may also allow more nuanced interpretations of income inequality and the role of luck (Frank 2016). In contrast to arguments for why the association between income and merit should be strong overall (Herrnstein and Murray 1994) or weak overall (Blau and Duncan 1967, Fischer et al. 1996), our model suggests a third possibility: the relationship between income and individual traits may be non-monotonic with a negative association in the middle.

In the remainder of this paper, Section 2 reviews related literature on performance metrics, evaluation, and reinforcing processes. Section 3 shows how a non-linear Pólya urn process generates a non-monotonic association between choice and expected quality. Section 4 explains how a dip may emerge when resources are allocated bimodally. In Section 5, we demonstrate that a dip occurs in the association between movie sales and ratings in the action/adventure genres. Section 6 discusses the scope conditions for a dip and the implications for the status-quality association, performance evaluations and social inequality in interpreting market outcomes.

2. Performance Evaluation and Reinforcing Processes

Using observed performance to make inferences about the potential of individuals, products and firms is central to organizational resource allocation processes, experiential learning and incentive systems. When information on inputs and effort is unavailable, mistrusted or not understood, managers rely on observed outputs to infer effort and merit (Ouchi 1979). However, drawing accurate inferences about future potential based on performance metrics is not easy (Meyer 2002, Starbuck 2005). It poses challenges similar to the well-known difficulties of accurately learning from experience (March 1994, March and Sutton 1997).

2.1. How Performance Metrics Should Be and Are Used

Several challenges make interpreting performance difficult (March and Sutton 1997). Observed performance is often a noisy indicator of future potential. Initial sales may be based on a small sample that is unrepresentative of future customers. Exceptional performance may be due to confounding events, making it unclear whether observed performance should be attributed to the person or to the situation (Ross and Nisbett 1991). For example, low sales during a recession should not necessarily be attributed to poor management or inferior product quality. Externalities also complicate attribution, as performance may depend on what others are doing. A new type of organization may not perform well unless many others have adopted similar forms (Elster and Moene 1989, Carroll and Harrison 1994). High performance in the short run may be associated with poor performance in the long run, and vice versa. Cost-effective manufacturing techniques may be profitable only after production has reached sufficient scale (Nelson and Winter 1982).

A normative theory of rational inferences from noisy performance indicators has been developed in engineering, statistics and economic contract theory (Harris and Raviv 1978, Holmstrom 1979, Milgrom 1981). According to this theory, performance evaluation should be based on informative indicators, and the weight given to any indicator should reflect its relevance and precision (Holmstrom 1979). The confounding effects of profit shocks common to all can be eliminated by measuring performance relative to the average of others operating in similar settings (Holmstrom 1979, Lazear and Rosen 1981). Subjective evaluations by boards or managers (Bushman et al. 1996) and non-financial measures (Gan et al. 2020) should be used to capture information that is not reflected in market- and accounting-based metrics. Long-term performance indicators can be used to avoid short-termism. It may even be rational for evaluators to reward initial failure in order to encourage exploration (Manso 2011).

Given that evaluators are boundedly rational, organizations also face the challenge of structuring the evaluation process to minimize bias and increase the precision of overall evaluations (Simon 1947). Noise can be reduced by averaging individual evaluations (Galton 1907, Surowiecki 2004). Biased and imperfect evaluators can be combined, structuring the decision process to achieve an overall evaluation with low bias and variance that approximates to a preferred trade-off between errors of commission and omission (Sah and Stiglitz 1986, Christensen and Knudsen 2010, Csaszar 2013, Keum and See 2017). To improve accuracy, organizations also try to de-bias individual evaluators or rely on statistical algorithms rather than human judgment (Dawes et al. 1989).

These normative recommendations are not always understood or adhered to. Observations of how performance metrics are actually used in organizations show that measures are misinterpreted or taken too literally, contextual information is ignored, subjective evaluations introduce bias, and long-term consequences are not salient (Hopwood 1972, Townley et al. 2003, Mazmanian and Beckman 2018). For example, Townley et al. (2003) quote a division manager as saying "we were told to improve the indicators and measures without a concern with underlying reality" (2003, p. 1061).

Several strands of organizational research explain why performance metrics are used inappropriately. Research in organizational sociology shows that formal evaluation metrics may become institutionalized and taken for granted (Meyer and Rowan 1977, Espeland and Sauder 2007, Colyvas 2012). When a performance metric is commonly used, managers may feel compelled to use it even if the context is inappropriate. This is partly because such metrics are taken for granted, and partly because conventional metrics provide a focal point when there is uncertainty about the metrics on which others will evaluate us (Correll et al. 2017). Universities know that they will be evaluated by students based on how well they perform in rankings, and therefore care about rankings even though many faculty members believe they are based on faulty methodologies (Espeland and Sauder 2007). Performance metrics also become embedded in organizational structures and routines, become valued in themselves, and are difficult to change when they are no longer appropriate (Colyvas 2012, Mazmanian and Beckman 2018). Behavioral research shows that people do not always understand performance metrics, and do not take mitigating circumstances sufficiently into account. They tend to over-attribute outcomes to the person in charge and discount situational influences (Ross and Nisbett 1991). They focus too much on the outcome achieved—particularly when it is perceived as a success or failure—and take insufficient account of the setting in which this outcome was achieved (Baron and Hershey 1988, Marshall and Mowen 1993). As a result, professional sports teams pay too much for players with high past performance that does not last (Massey and Thaler 2013), and banks pay too much for security analysts whose performance does not transfer to their new organization (Groysberg et al. 2008). Experimental studies show that people focus on activities that improve short-term performance without understanding their negative long-term consequences (Sterman 1989, Herrnstein et al. 1993). They also misinterpret common performance metrics, such as returns on investment (Shapira and Shaver 2014), and do not fully understand the advantages of averaging forecasts (Larrick and Soll 2006) and using structured interviews (Dana et al. 2013).

2.2. Reinforcing Processes

Interpreting business performance is complicated by the fact that current performance is influenced by past performance (March and Sutton 1997). Success in the past may increase the chances of success in the future. This reinforcing process may amplify luck (Arthur 1989, Salganik et al. 2006) and lead to a low correlation between quality and performance (Lynn et al. 2009).

Product and market success are subject to several types of reinforcing process (Arthur 1989, Lieberman and Montgomery 1988). The utility of a product may depend on the number or proportion of others using the same product, partly because a commonly used product attracts more complementary products (e.g., games for game consoles). Such network externalities imply that adopters are concerned not only about attributes, but also about whether many other adopters have or will adopt the same product (Katz and Shapiro 1986). The implication is that a lower-quality product may end up dominating the market as a result of gaining an early lead (Arthur 1989). Empirical studies demonstrate that network externalities impacts adoption (Nair et al. 2004) and explains market dominance in the video cassette recording (Park 2004) and video game console industries (Dubé et al. 2010), although network effects are not always strong enough to stop later entrants with high-quality products from eventually dominating (Liebowitz and Margolis 1995, Tellis et al. 2009, Zhu and Iansiti 2012, Gretz and Basuroy 2013).

Having many users may also increase future adoption, because people assume that a commonly-used product is better (Bikhchandani et al. 1992), feel pressured to conform (Asch 1955), or are more likely to have heard of it. Salganik et al.'s (2006) web-based experiment on music downloads illustrates the strength of social influences on adoption. They set up a website with 48 songs that participants could listen to and download. The experiment was repeated with eight different participant pools, generating data from eight worlds. The results show that when popular songs were displayed prominently (at the top of the website), early popularity strongly impacted on the eventual outcome. For example, a song of medium quality (Song no 31, quality assessed by independent ratings) became the most popular song in one of the eight worlds when it happened to be popular early on. Early popularity does not always have such a large impact (Azoulay et al. 2014, van de Rijt et al. 2014, 2016, van de Rijt 2019). Reinforcing processes will be stronger when consumers are aware only of popular products, when initially popular products or individuals that do well are given additional resources, and when consumers have strong incentives to select apparently superior products (Vincenz and van de Rijt 2016, van de Rijt 2019).

2.3. Our Contribution

It is well known that reinforcing processes may lead to a low correlation between quality and realized performance, because even low-quality products may occasionally achieve success if they happen to gain early popularity (Arthur 1989, Carroll and Harrison 1994, Chase 1980, Lynn et al. 2009, Merton 1968, Skvoretz et al. 1996). However, previous literature has not questioned whether higher performance indicates higher quality. The correlation between quality and performance in past models can be close to zero yet remain positive (Lynn et al. 2009). If the correlation is positive, it still makes sense to reward and imitate the highest performers, because on average they are more skilled than lower performers. Our interest lies in when reinforcing processes imply that the correlation between performance and quality may be *negative* for some levels of performance. Specifically, when is expected quality based on observed performance, $E[q_i|p_i]$, a decreasing function of observed performance for some ranges of p_i? We are interested in this because if $E[q_i|p_i]$ is a decreasing function of p_i, naïve reliance on performance metrics—assuming that better performance indicates higher quality—will lead to biased and unmeritocratic resource allocations.

There is very little research on the social and economic *mechanisms* that imply that expected quality given performance is a decreasing function of performance. It is well-known in statistics and economics that such non-monotonicity may occur. For example, when the error distribution is fat-tailed (Denrell and Liu 2012, Weibull et al. 2007), such a set-up violates the "monotone-likelihood-ratio" condition required for $E[q_i|p_i]$ to be an increasing function of p_i (Karlin and Rubin 1956, Milgrom 1981). Economic research on "bandwagons" also shows that in some settings imitating the minority may be rational (Callander and Hörner 2009, Eyster and Rabin 2014). The paper most closely related to ours is Denrell and Liu's (2012) model of how rich-get-rich dynamics imply that top performance is an unreliable indicator of quality. They assume that the strength of reinforcing processes differs among markets/settings and is unknown to evaluators. In their model, an extreme result (e.g., a very high market share) indicates that reinforcing processes are strong, which in turn implies that evaluators should regard the outcome as an unreliable indicator of quality. This mechanism is compelling in some settings, but does not always apply (Liu 2020). For example, it would not be applicable if it was known that reinforcing processes in an industry are strong due to network externalities. Our contribution is to show more generally that higher market share can be associated with lower quality due to a bimodal allocation of resources.

3. Model

Inspired by Arthur (1989), we analyze a simple model of how consumers choose between two competing products. These consumers care about both quality and the proportion of other consumers who have selected each product. We describe a scenario in which they want not only to choose the best product with the highest level of quality, but also to coordinate their choices with others, perhaps due to network externalities. Consumers may also prefer popular products owing to social influences, because they believe that popular products are superior or because they are more likely to have heard of them.

3.1. Set-up

In each period, $t = 1, 2, ...$, a consumer arrives and selects between two competing products, $j = 1, 2$. The utility of selecting product $j = 1, 2$ in period t is $u_{j,t} = aq_j + bm_{j,t-1}$, where q_j is the quality of product j, $m_{j,t-1}$ is the market share of product j in period $t-1$, $a > 0$ is the importance of quality, and b is the importance of market share. Because market shares sum to one, $m_{2,t-1} = 1 - m_{1,t-1}$. In period one, when $m_{1,t-1}$ is not defined, we set $m_{1,0} = m_{2,0} = 0.5$. We assume that the prices of the two products are identical (e.g., two apps are both free), so the customer's choice does not depend on price. We model the decision process as probabilistic, whereby consumers usually, but not always, select the product with the highest level of utility.[1] Specifically, the probability that the consumer in in period $t+1$ selects product 1 is:

$$P_{1,t+1} = \frac{e^{aq_1+bm_{1,t}}}{e^{aq_1+bm_{1,t}} + e^{aq_2+bm_{2,t}}} = \frac{1}{1 + e^{aq_2+b(1-m_{1,t})-aq_1-bm_{1,t}}} = \frac{1}{1 + e^{-a(q_1-q_2)-2b(m_{1,t}-0.5)}} \quad (1)$$

which is the well-known logit choice model. The probability of selecting product 1 is an increasing function of its relative quality, $q_1 - q_2$, and its market share, $m_{1,t}$. We are interested in whether a higher market share necessarily signals higher quality. To address this question, we fix the quality of the second product to 0.5 (we obtain similar results if we allow q_2 to vary). We allow the quality of product 1 to vary across simulations of the model. Specifically, in each simulation q_1 is drawn from a uniform distribution between zero and one at the start ($t = 0$) and remains the same for all periods. If we simulate the model k times, we thus simulate k pairs of products competing for t periods, where q_1 is redrawn in each of k simulations from a uniform distribution, and q_2 is fixed at 0.5. Hence, each simulation models a sequence of competitions between a firm with product 1 of unknown quality (q_1), and a firm with a product of known medium quality ($q_2 = 0.5$).

3.2. Results

How does the market share achieved by product 1 after 10 or 100 periods reflect its quality? One might imagine that a higher market share indicates a higher expected value of q_1; in other words, $E[q_1|m_{1,t}]$ is an increasing function of $m_{1,t}$. When $b = 0$ and utility depends only on quality, this is true; but it is not true if $b > 0$ and consumers care about the product's past market share when making a selection. As Figure 1 shows, when b is sufficiently large, $E[q_1|m_{1,t}]$ is a decreasing function of $m_{1,t}$ for intermediate values of $m_{1,t}$. In such cases, there is a middle dip in $E[q_1|m_{1,t}]$: expected quality initially increases, then decreases, but increases again with higher market share levels.

[1] Arthur (1989) assumed that consumers always choose their preferred alternative, but that consumers differ in their evaluation of alternatives. Such heterogeneity in preferences is one way of interpreting a probabilistic choice model such as ours.

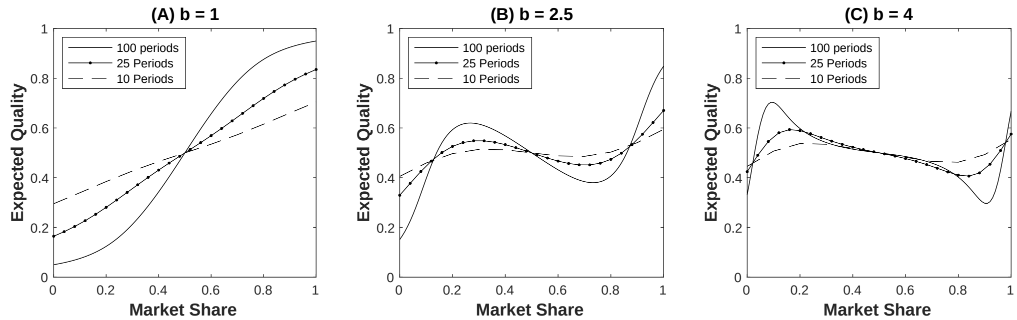

Figure 1 **Expected quality of firm 1 given the market share firm 1 obtained in the last period when $a = 1$ and A) $b = 1$, B) $b = 2.5$ and C) $b = 4$. These results are based on numerical integration, see Appendix D.**

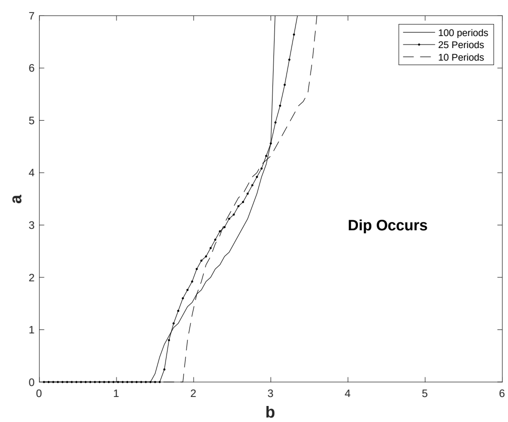

Figure 2 **Combinations of a and b for which there is a "dip" in expected quality given the observed market share. These results are based on numerical integration, see Appendix D.**

Figure 2 shows the combinations of a and b for which a dip in $E[q_1|m_{1,t}]$ occurs. A dip requires that the importance of past market share, b, is sufficiently large compared to the importance of quality, a. In fact, in the long term, as $t \to \infty$, it can be demonstrated formally that a dip occurs whenever $b > 2$, regardless of the value of a and for any distribution of quality (see Theorem EC.2 in Appendix A).

3.3. Intuition

Why does a dip occur, and why does the occurrence of a dip depend on the value of b? Figure 3 shows that a dip occurs when there is a bimodal distribution of market share. If b is sufficiently large, the reinforcing process is strong enough for product 1 to secure a market share of either close to one or close to zero, depending on whether it is lucky and initially frequently chosen, or unlucky

(Figure 3B). However, the market share that product 1 obtains also depends on its quality. A high-quality product ($q = 0.9$) with good initial fortune will nearly always reach a market share above 90%, while a low-quality product ($q = 0.1$) with good initial fortune will reach a market share close to 80%. Similarly, a high-quality product without initial luck will tend to reach a market share close to 15%, while a low-quality product without initial luck will nearly always reach a market share lower than 10%. Reaching a market share of around 80% or below 10% is thus typical of a low-quality product, while a market share above 90% or around 15% is typical of a high-quality product (Figure 3B).

It follows that high-quality products are more likely than low-quality products to achieve either a very high market share (above 90%) or a market share of around 15%. To illustrate this more clearly, Figure 3D plots the proportion of high- versus low-quality products for different market share levels; more specifically, it plots $P(m_{1,100}|q_1 = 0.9)$ divided by $P(m_{1,100}|q_1 = 0.9) + P(m_{1,100}|q_1 = 0.1)$ for different values of $m_{1,100}$. As shown, high-quality products are overrepresented both at very high levels of market share and at market shares close to 20%. Low-quality products are overrepresented at very low levels of market share but also at market shares close to 80%. This provides an explanation for the dip: both a very low market share (below 10%) and a quite high market share (at 80%) are indications of low quality.

No such dip occurs when b is low enough, because the distribution of market shares is then unimodal rather than bimodal (Figure 3A). In this scenario, low-quality products are overrepresented among products with low market shares, medium-quality products are in the middle region, and high-quality products at high levels of market share (Figure 3C), implying that a higher market share indicates higher quality.

To explain why there is a bimodal distribution of market shares only when b is large enough, Figures 4A and 4B plot how the choice probability varies with past market shares for low and high levels of b. In the long term, the market shares to which product 1 will converge are the equilibrium points at which the choice probability equals past market share: $P_{1,t+1} = m_{i,t}$. Graphically, these equilibrium points, or so-called "fixpoints", are the points at which the choice probability function, $P_{1,t+1}$, crosses the "45-degree" line. Only fixpoints at which the choice probability function, $P_{1,t+1}$, crosses the "45-degree" line from below are stable, in the sense that if the market share deviates slightly from this point, it will tend to revert back to this point (for further technical details, see Appendix A). Past work on Pólya urn models demonstrates that in the long run, the process will only converge to these stable fixpoints (Arthur et al. 1987, Hill et al. 1980).

When b is large enough, there is a possibility of multiple stable fixpoints (Figure 4B). A product of intermediate quality (i.e., between 0.35 and 0.65) may converge to either a high or low level of market share. Very low- or very high-quality products, on the other hand, can only converge to

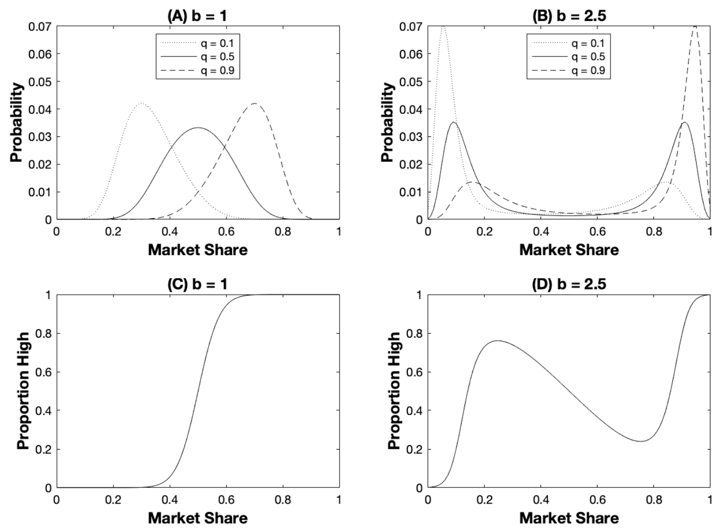

Figure 3 (A) and (B): distribution of market share achieved by firm 1 after 100 periods ($m_{1,100}$) for three levels of quality, when $a = 1$ and $b = 1$ or $b = 2.5$. (C) and (D): proportion of high- versus low-quality individuals at different market share levels. These results are based on numerical integration, see Appendix D.

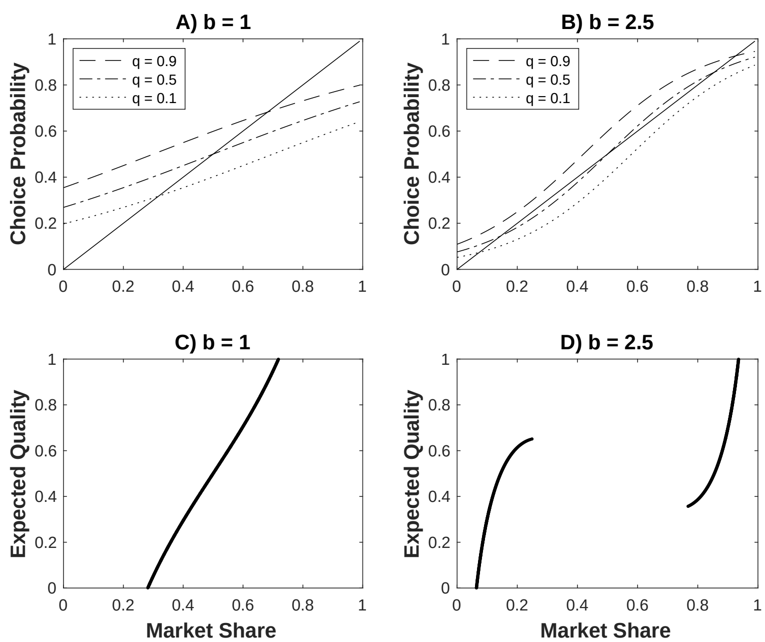

Figure 4 (A) and (B): how the probability that the consumer will select product 1, $P_{1,t+1}$ in period $t+1$ varies with the market share, $m_{i,t}$, at three levels of quality, when $a = 1$ and $b = 1$ or $b = 2.5$. (C) and (D): quality as a function of the market share reached as $t \to \infty$.

one market share level. The possibility of multiple equilibria generates a dip (Figure 4D): expected quality initially increases with market share, up to the level to which high-quality products (about 0.65) without early luck can converge (a market share around 30%). There is then a gap in the market share distribution. At about 70%, expected quality again increases in market share. However, a market share of around 70% is associated with a quality of about 0.35, which is lower than the quality associated with a market share of 30%. When b is lower, there is only one stable fixpoint (Figure 4A), which implies that a higher market share indicates higher quality (Figure 4C).

3.4. Scope Conditions

Whether or not a dip occurs depends on the parameters of the model. In particular, a dip does not occur if past choices only weakly impact future choices, i.e., a dip requires a sufficiently high value of b. An even larger value of b is required for a dip if a product's utility is a strongly concave or convex function of its market share. Consider a model identical to the basic model introduced above, except that $u_{j,t} = aq_j + bm^k_{j,t-1}$. Previously, we assumed that $k = 1$. If $k < 1$, utility is a concave function of the market share, and if $k > 1$, utility is a convex function of market share. Computations show that a dip occurs even when $k > 1$ or $k < 1$, but if k differs significantly from one, a dip only occurs if b is high. The intuition is that when k differs significantly from one, changes in market share have little impact on a product's utility at high (concave function) or low (convex function) values. For example, when $k = 1$, a dip occurs after 100 periods when $a = 1$ and b is larger than approximately 1.8. When $k = 5$, b needs to be larger than 2.8. Similarly, if $k = 0.5$, b must be larger than 2.2, and when $k = 0.2$, it must be larger than 3.0.

Moreover, a dip does not occur in the long term if quality does not influence the choice probability for sufficiently popular products. This may happen when utility depends on the *number* rather than the *proportion* of past consumers who have selected product j. Suppose $u_{j,t} = aq_j + bn_{j,t-1}$, where $n_{j,t-1}$ is the number of consumers who have selected product j during periods 1 to $t-1$. If b is large enough, this model also implies that product 1 will secure either a very high or a very low market share. Yet this model does not generate a dip. Rather, average quality is an increasing function of the obtained market share.[2] To understand why, note that when t is large, $n_{1,t-1} - n_{2,t-1}$ will become a very large positive or negative number. Hence, the choice probability will be influenced mainly by $n_{1,t-1} - n_{2,t-1}$, and the quality of product 1 will be inconsequential. In the long term, the choice probability will become one or zero, regardless of the quality of product 1. Thus, when utility depends on the difference in numbers (which may be unbounded), quality determines the probability that product 1 will win, but does not determine the market share if

[2] When b is large, the market share of product 1 converges to either zero or one. Still, average quality is higher for products with a market share equal to one than for those with a market share equal to zero.

product 1 wins, because all winning firms will achieve a market share equals to one. In contrast, if utility depends on the proportion (which is bounded), quality determines both the probability of winning (i.e., obtaining a market share larger than 50%) and the market share that product 1 will secure if it wins. Whether quality influences the choice probability of popular products depends on the setting. If people assume that the most popular product is superior, and this overrides their own information, all may choose it irrespective of actual quality (Bikhchandani et al. 1992). However, if having few others selecting the same option is sufficient to convince people to choose a minority option they prefer (Asch 1955), only products that most believe are high-quality will become widespread.

If quality does influence the choice probability of popular products, as is true in our basic set-up when the utility of a product depends on the proportion of past consumers, then the existence of a dip (when b is sufficiently large) is robust to several modifications of the model. First, we assumed that quality has a uniform distribution, but a dip occurs for many other distributions (see Appendix B for details). Indeed, Theorem EC.2 in Appendix A shows that a dip occurs in the long term whenever $b > 2$ for any (continuous) distribution of quality on $[0, 1]$. Second, a dip also occurs in a model with more than two competing products (see Appendix B for details).

Third, a dip also occurs for probabilistic choice models other than the logit model. For example, Theorem EC.3 in Appendix A shows that a dip occurs in the long term for the probit choice model (i.e., product 1 is selected in period t if $u_{1,t} + \varepsilon_{1,t} > u_{2,t} + \varepsilon_{2,t}$, where $\varepsilon_{j,t}$ are independently drawn from a standard normal distribution) whenever $b > \sqrt{\pi}$. The crucial feature of both of these probabilistic choice models is that the probability of choosing product 1, $P(u_{1,t} - u_{2,t})$, is a convex function for $u_{1,t} - u_{2,t} < 0$ and a concave function for $u_{1,t} - u_{2,t} > 0$. These properties imply that two stable fixpoints exist in the mapping of $P(z) - z$. Lemma EC.1 in Appendix A shows that a large class of probabilistic choice models share these features: any random utility model in which individual i selects product 1 if $u_{i,1} + \varepsilon_{1,i} > u_{i,2} + \varepsilon_{2,i}$, where $\varepsilon_{j,i}$ are drawn independently from a *unimodal* distribution, has the same essential features (i.e., convex for $m < 0.5$ and concave for $m > 0.5$) and generates a dip if b is sufficiently large (Theorem EC.1). The assumption of a unimodal error term makes sense when large errors are less frequent than small ones. In contrast, if large positive and negative errors are both more frequent than small errors, a dip cannot occur (in the long run). However, dip can occur in the short term even for choice models that do not imply a dip in the long term. In particular, a dip occurs in the short term for a linear probability model: $P_{1,t+1} = wq_1 + (1-w)m_{1,t}$. For example, a dip occurs after 100 periods as long as $0 < w < 0.242$. However, in this linear model, the dip disappears as $t \to \infty$ because there can only be one fixpoint for this linear function, but if $w = 0.1$, a dip occurs even after 20,000 periods.

4. A Bimodal Allocation of Resources Generates the Dip

In the model in Section 3, market share is subject to a reinforcing process which, if sufficiently strong, leads to a bimodal distribution and a dip in the market share-quality association. However, a reinforcing process, operating sequentially over several periods, is not necessary to generate a dip. Reinforcing processes are just one of many ways of generating a bimodal allocation of resources. Whenever there is a bimodal allocation, a dip may occur. The following simple resource allocation model illustrates this more general mechanism.

We consider a setting where an evaluator has to decide how much resource to allocate to a set of individuals or products. For example, investors, funding agencies or movie company executives must decide which firms, researchers or movies to sponsor. We imagine that an outcome (financial success, publications or box office takings) depends on quality as well as on resources: $o_i = q_i + r_i$. In the model presented in Section 3, the outcome was the market share, and the additional resources were the installed base (i.e., the past market share). In the resource allocation model, the outcome might be the product's sales levels, the number of citations an academic paper receives, or the movie's box office takings. Quality is some fixed trait that impacts on the outcome (e.g., skill, business acumen or appeal). For simplicity, we assume that there are only four levels of quality: 1, 2, 3 and 4. Resources include financial and social capital that increase the chances of achieving a better outcome.

When does a dip occur in the resource allocation model? If all products receive the same level of resources (i.e., r_i is the same for all i), a higher outcome (higher o_i) will indicate higher quality (Figure 5A) and there will be no dip. However, a dip may emerge when there are two levels of resources: few ($r_i = 0$) or many ($r_i = h$) (see Figure 5B). Suppose that the lowest-quality products (equal to 1) never receive many resources and the highest-quality products (4) always receive many resources, whereas medium-quality products (2 and 3) may receive either few or many resources. In this case, a dip may occur. Specifically, a dip occurs whenever $h > 1$. When $h > 1$, a product with a quality level of 2 that happens to receive many resources will achieve a higher outcome $(2 + h)$ than a product with a quality level of 3 that receives no resources and obtains an outcome of 3. This is true for any distribution of quality, with some probability at each level. It is also true for any process where resources are allocated to products with different quality levels, which allows for "allocation mistakes" in that a product with a quality of 2 may receive more resources than a product with a quality of 3.

The general mechanism behind the dip in this resource allocation model is the same as the reason for the dip in the model in Section 3. First, there is a bimodal distribution of resources (i.e., products receive either few or many resources); Second, an allocation mistake is possible; Third, both quality and resources impact on the final outcome. These three features imply that

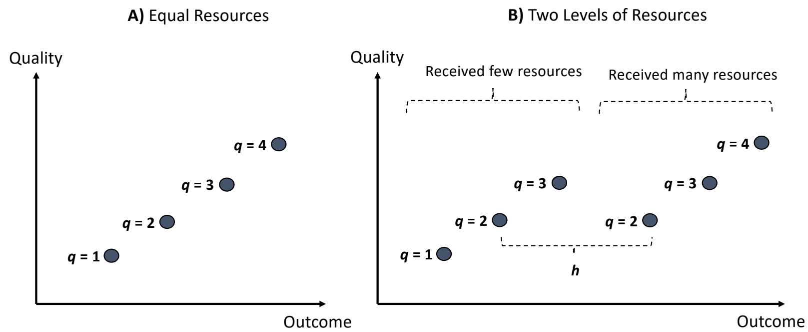

Figure 5 **Quality as a function of achieved outcomes when (A) all products receive equal resources and (B) products receive either a few or many resources.**

the distribution of outcomes will be bimodal, and a moderately high outcome will be associated with relatively low quality.

The dip that occurs in the resource allocation model hinges on the fact that the distribution of resources invested is bimodal. When would this investment pattern make sense? The answer is that focusing resources on a single or a few investments makes sense when there are increasing returns to investments (Wernerfelt and Karnani 1987). Suppose, for example, that the management of a movie company believes that there is a winner-takes-all dynamic in movie sales, where one or at most a few movies attract the bulk of attention. If so, it would make sense to invest substantial resources in a few movies that seem promising, with limited investments in most others (to spread the risk). More generally, if past success increases the chance of future success, investors may rationally allocate more resources to early winners, anticipating that these are more likely to succeed in the future (Katz and Shapiro 1986, Arthur 1989). Such a winner-takes-all investment policy will further increase early winners' chances of success (Merton 1968).

5. Empirical Illustration: Movie Sales and Ratings

Our theoretical argument shows that a dip may occur when the final outcome depends on quality but is also significantly impacted by resources allocated to early winners, through a reinforcing process or a resource allocation decision based on anticipated success, and when decisions are made under considerable uncertainty about quality. The movie industry fits these criteria. An important outcome in the movie industry is box office sales. Previous studies have shown that these sales depend not only on the movie's characteristics (i.e., its quality or appeal to movie-goers, often measured by review scores), but also on resource allocation decisions about investments in production and marketing budgets, and the number of theaters in which the movie will be released. Investments increasingly follow a bimodal pattern, whereby movie companies invest considerable resources in a few movies that are anticipated to be popular, with the aim of these movies achieving "blockbuster" status (Baker and Faulkner 1991, Elberse 2013, Eliashberg et al. 2000).

Unlike in our theoretical models, significant resource allocation decisions in the movie industry do not occur in every period, as in the model in Section 3, or only once, as in the resource allocation model, but at a few important stages in the movie's production. At these stages, decision makers rely on different types of information to estimate future popularity. First, decisions about financing and the production budget are based on the initial pitch and the quality of the screenplay. At this point, no sales data are available. The second crucial resource allocation decision is made at the distribution stage. Based on pre-screenings with test audiences and intuition regarding the size of the market for this type of movie, distributors make crucial decisions about marketing budgets, as well as the number of theaters in which the movie will be released. With a few exceptions, movies initially launched in only a few theaters are unlikely to attract many movie-goers (De Vany 2004, Zuckerman and Kim 2003). Conversely, launching a movie in many theaters typically ensures a higher minimum level of box office sales. A third resource allocation decision is made after the opening weekend, and at this point sales data are available. Movies that do well during their opening week are allocated additional theater screens or continue to be shown on a large number of screens, which enables them to continue to perform well (De Vany 2004, Elberse and Eliashberg 2003). However, movies that have poor ticket sales during their opening weekend are allocated fewer theater screens, limiting their future sales potential.

Although resource allocation decisions are crucial, and movie studios focus increasingly on investing in movies that are most likely to become blockbusters, the association between the amount of resource allocated and movie ratings (by critics, as well as the general public) is weak, owing in part to the difficulty of predicting movie-goers' reactions accurately (De Vany 2004). To illustrate this decoupling, consider the film "Speed 2: Cruise Control" starring Sandra Bullock and Willem Dafoe, which was released on June 15, 1997 as a sequel to the financially successful film "Speed" released in 1994. This sequel garnered domestic box office sales of 48 million USD, but earned a disastrous IMDB rating of 3.7 (on a scale of 0 to 10), ranking in the 3rd percentile of all movies rated on the website. It was nominated for eight Golden Raspberry Awards, which are tongue-in-cheek awards given annually to movies widely considered to be awful, and won the award for "Worst Remake or Sequel". Despite the overwhelmingly negative reviews, the success of "Speed" enabled this sequel to receive considerable financial investment and broad distribution, creating an early advantage as it was screened at most theaters in the United States ($n = 2,615$). This advantage was sufficient to boost box-office sales, making it the number one movie during its opening weekend. However, its poor quality quickly led distributors and theater owners to replace it with other movies (it was screened for six weeks in total, much shorter than other opening-week box office winners).

Contrast the fate of "Speed 2" with that of "The Boondock Saints", released on January 21, 2000, which also included Willem Dafoe, as well as Sean Patrick Flanery. The story centers around two brothers from Boston who try to "clean up" their city by massacring local criminals. Its total domestic box office sales reached only 30,471 USD, but it was highly-rated on IMDB (with an average rating of 7.8, approximately the 93rd percentile). This film, with a budget of 6 million USD, was a box office failure because it was only distributed to five theaters for one week. The movie script had been regarded as promising and had been sought after by major studios, including New Line Cinema, Paramount Pictures and Miramax. However, after the movie was completed, every major distributor refused to be associated with it owing to the recent Columbine High School massacre, which occurred on April 20, 1999. This unfortunate coincidence eliminated this movie's chances of gaining high box office sales, even though it was highly regarded by those who watched it later on DVD/Blu-Ray or through streaming services.

The importance of a movie's budget and the number of theaters in which it is released, the bimodal pattern of resource allocation, and the weak association between quality and resource allocation suggest that a dip may exist in the association between movie quality (as measured by ratings) and outcomes (as measured by box office sales). In particular, there appears to be two "playing fields" in this industry. Movies believed to have low potential are allocated limited resources and open in a few theaters, limiting their sales potential, while movies anticipated to do well are allocated substantial budgets and open in many theaters, enabling much higher sales potential. In particular, opening in a large number of theaters helps to place a floor under box office revenues, so the lower decile increases substantially (De Vany, 2004: 135). In both fields, sales also depend on quality (De Vany, 2004). High-quality movies that open at few theaters do better than low-quality movies that also open at few theaters. High-quality movies that open at many theaters do better than low-quality movies that also open at many theaters. This suggests that a dip may exist for movies that have moderate box office sales, which may be low-quality movies that happen to have received substantial resources.

An important caveat is that not all movies are produced and launched with the intention of maximizing sales. Some, especially those produced by independent moviemakers, focus on a niche market and are deliberately released in a limited number of theaters. We alleviate this problem by focusing on the action/adventure genres, in which financial and distribution support matter most, and there are few independent moviemakers (Zuckerman and Kim 2003). We also compare the results with movies in the drama genre, in which a large proportion of movies are produced by independent moviemakers.

5.1. Data

To examine whether a dip occurs in the movie sales-rating association, we collected data on movie sales and movie ratings from multiple sources. Using cumulative box office sales as the outcome measure, we obtained data from "The Numbers" (www.the-numbers.com), which is based on theater box office earnings in the United States and Canada, for 28,103 movies released between 1948 and 2016.[3] We collected average movie ticket prices from 1910 to 2017 from the National Association of Theatre Owners (NATO), which has conducted annual surveys of its members since 1989 and uses the consumer price index (CPI) to estimate the average movie ticket price for each year prior to 1989. We adjusted box office revenues against movie ticket prices, using 2010 as the base year.

We used movie review scores to measure the "quality" of a movie. It is, of course, very difficult to define and measure the true quality of a movie, and this is not our intention. In our models, we use the term "quality" to denote a stable trait (fixed over time) that impacts the probability that a consumer will select a product. In the same spirit, we rely on review scores to measure attributes of movies that impact whether the relevant audience will view them favorably.[4] IMDB scores are based on reviews submitted by general audiences and reflect their tastes rather than those of experts, such as movie critics. Nevertheless, IMDB scores are significantly and positively associated with critical evaluations. Brown et al. (2012) show that the correlation between Metacritic.com ratings, which are based on several movie critics' reviews, and IMDB scores is 0.75 for movies released between January 1, 2000, and December 31, 2009. IMDB data for all the movies were collected in February 2017. Because most IMDB movie ratings stabilize about three months after the opening weekend (Holbrook and Addis 2007), we expected the review scores to have stabilized for all movies in our dataset.

We obtained 709,488 review scores from IMDB (www.imdb.com). We excluded all non-movie data. We compared the review data with our box office data and only included movies for which data on the title, production year and production company matched. This resulted in 8,738 movies for which information was available on both review scores and inflation-adjusted cumulative box office sales. To examine how early outcomes influenced final box office sales, we collected information

[3] To cross-validate the accuracy of box office sales data, we also downloaded data from Box Office Mojo (www.boxofficemojo.com), which were largely consistent, although we excluded 197 movies where inconsistencies were larger than 100,000 USD.

[4] Review scores are not necessarily stable because they can be impacted by box office sales. Nevertheless, if review scores were (positively) influenced by box office sales, it would presumably generate a monotonic pattern, working against a dip. Endogeneity could only create a dip if the impact of sales on review scores were negative. This negative impact is more likely in the drama genre where box office success attracts audiences with diverse tastes (Kovacs and Sharkey 2014), and is less likely in the action/adventure genres where audience taste is more homogeneous (De Vany 2004).

about movies' opening-week box office sales and production costs, as well as the number of theaters in which they opened.[5] We obtained data on production costs from The Numbers and data on opening-week box office sales and number of theaters from Box Office Mojo, resulting in 2,621 movies with reliable information for all three variables and 1,746 movies with reliable information for these three variables as well as eventual box office sales and IMDB ratings. Opening-week box office sales were also adjusted for inflation using movie ticket prices from NATO, with 2010 as the base year. Table 5.1 provides descriptive statistics for the main variables.

Table 1. Descriptive statistics for movie data

Variable	Obs.	Mean	Std. Dev.	Min	Max
Total box office sales	8,738	17.7m	40.4m	45	936.7m
IMDB rating	8,738	6.4	1.0	1.1	10
Number of theaters	2,621	2,018	954.1	600	4,468
Production costs	2,621	32.9m	40.8m	5000	275m
Opening-week box office sales	2,621	13.3m	18.3m	96,076	208.8m

"m" represents millions of US Dollars

5.2. The Existence of a Dip

Our theory predicts that a middle dip will occur in the association between movies' box office sales and average reviews when some movies are provided with substantially more resources than others, as such investments may substantially increase sales. As demonstrated below, this investment pattern occurs in the action/adventure genres, for which commercial success is crucial, but not in the drama genre where maximizing artistic impact is more important and launching movies in numerous theaters is less common. Hence, in the following paragraphs, we contrast the action/adventure and drama genres, which are three of the largest movie categories.

Figure 6 plots how the average IMDB review score varies with the logged box office sales for all movies in the action/adventure genres (Panel A, N=972) and for all movies in the drama genre (Panel B, N=2,448) in our database.[6] In the upper panels, the solid line is the average review score, and the dashed lines show the 95% confidence intervals based on 10,000 bootstrap simulation runs. The lower panels show the proportion of observations for different box office sales intervals.

In the action/adventure genres (Panel A), there is a middle dip in the association between movies' box office sales and average IMDB review scores. The average review score initially increases with box office sales for movies earning less than 100,000 USD. It declines with box office sales for movies with sales between 100,000 and 10 million USD (except for a kink around 5 million) and

[5] The number of theaters in which a movie opened is a conservative measure of maximum audience size because a movie can be played on multiple screens in one theatre, but the screen count is rarely published domestically.

[6] These plots exclude box office intervals with fewer than five movies, which excluded 12 movies.

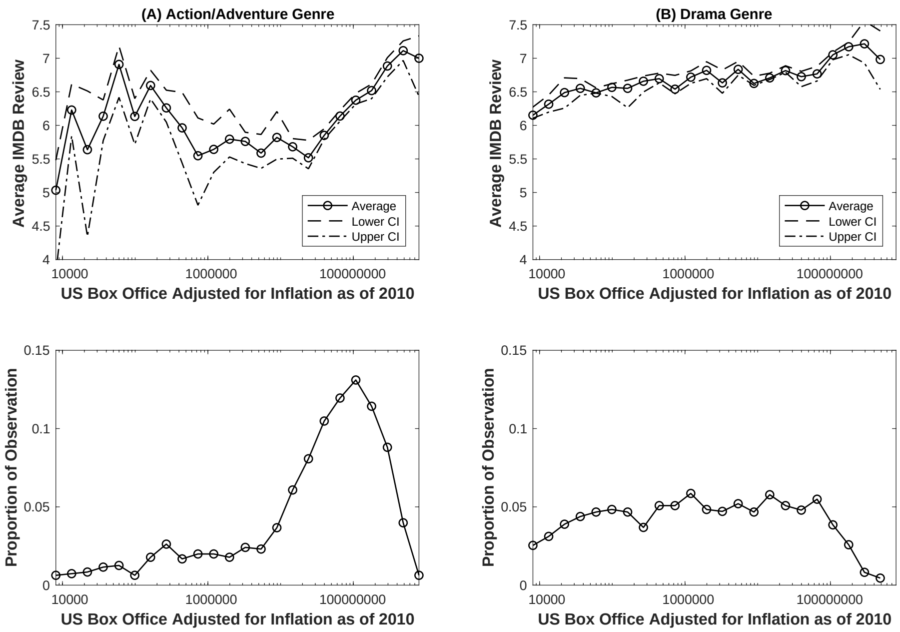

Figure 6 **Upper panels: association between box office sales and quality (average review scores in the IMDB database). The solid line is the average, and the dashed lines are 95% bootstrap confidence intervals based on 10,000 simulation runs. Lower panels: proportion of observations. Panel A is based on movies in the action/adventure genres, and Panel B on movies in the drama genre.**

then plateaus between 10 million and 25 million USD. Finally, it increases with box office sales for movies making more than 25 million USD, the only exception being a dip at the top for movies with sales above 800 million USD. In contrast, Panel B shows that in the drama genre, the association between box office sales and review scores increases monotonically, except for a dip at the top for movies with sales above 485 million USD.

Is the middle dip in the action/adventure genres' sales-quality association reliable? To explore this more systematically, we first fitted a cubic spline regression to the data. The results reveal a non-monotonic association in the action/adventure genres, and a monotonic association in the drama genre (see upper panels in Figure 7). The changes in sign roughly correspond with the change in slope observed in the non-parametric plots in the upper panel of Figure 6.

Is the sign change detected by the cubic spline reliable? To answer this question, we rely on a modified version of Simonsohn's recently published "two-lines" test, which was developed to test for possible u-shaped relationships that are less sensitive to functional form (i.e., quadratic or cubic) assumptions (Simonsohn 2018). In brief, our modified tests work by fitting two regression lines around the local minima and local maxima identified by the cubic spline, and testing whether the signs in the estimated lines differ (see Appendix C for details). To maximize the test's power to

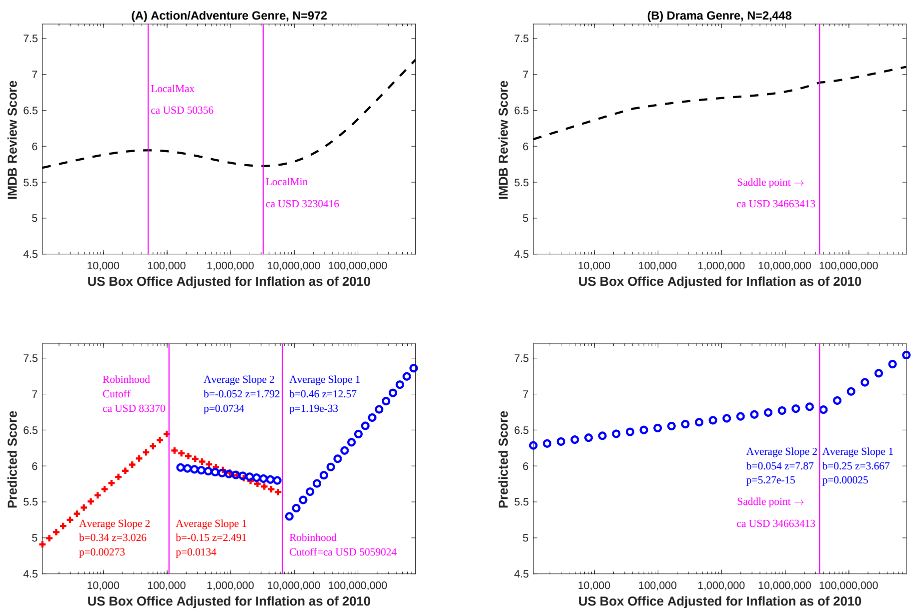

Figure 7 Upper panels: cubic splines fitted to data on box office sales and quality (average review scores in the IMDB database). Lower panels: results of applying a modified "two-line test" to our data. Panels A and B are based on movies in the action/adventure and drama genres, respectively.

detect sign changes, the regression lines are fitted around a new value—the so-called "Robinhood cut-off"—rather than around the identified local minima or maxima. The purpose of this step in the algorithm is to "reinforce the poorer" by allocating more observations to the side (above or below the original cut-off) where the regression slope has weaker statistical power.

The lower panels of Figure 7 show the results of applying the algorithm to the action/adventure and drama genres. The action/adventure genres are considered first. The cubic spline identifies a local minimum and a local maximum. The estimated slope for movies above the local minimum is positive (Z=12.57), while it is negative (Z=1.79) below the local minimum (the blue line). The estimated slope for movies below the local maximum is positive (Z=3.03), while it is negative (Z=2.49) for movies above the local maximum and below the local minimum (the red line). This pattern is consistent with a middle dip in the association between box office sales and average ratings of movies in the action/adventure genres. In contrast, in the drama genre, the estimated slopes for movies below and above the saddle-point at around 35 million are both positive, implying that the average rating is a monotonically increasing function of box office sales.

5.3. The Distribution of Outcomes and Investments

Why does a dip occur in the action/adventure genres but not in the drama genre? To explore the immediate causes of the dip, the upper panels in Figure 8 plot the distribution of box office

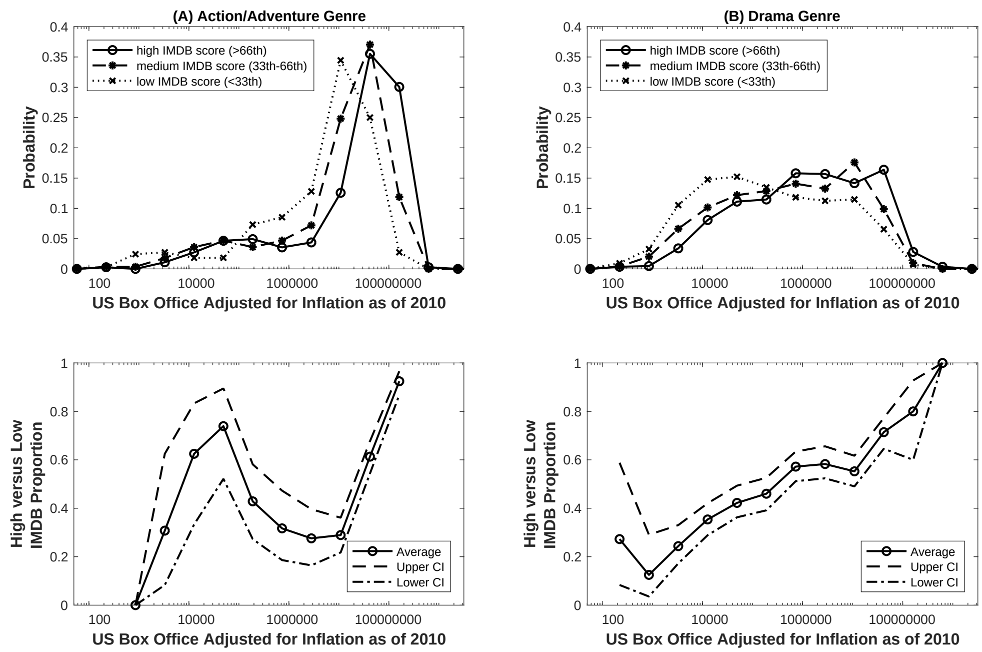

Figure 8 **Upper panel: box office sales distributions for movies at different levels of IMDB ratings. Lower panel: proportion of high- versus low-quality movies at different sales intervals. Panel A is based on movies in the action/adventure genres, and Panel B on movies in the drama genre.**

sales at three quality intervals. First, we consider the action/adventure genres. For movies in these genres (Panel A), the distribution of box office sales exhibits a clear bimodal pattern at all quality levels. Movies at all quality levels tend to do either relatively poorly or relatively well, whereas intermediate levels of sales are less frequent. To test whether the outcome distribution is bimodal, we fit (by maximum likelihood) a mixture of two normal distributions to logged box office sales. At all three levels of quality, this mixture distribution fits the data better than a single normal distribution, providing evidence of bimodality.[7] For high-quality movies the highest mode occurs at 122 million USD and the lowest at 1.4 million, for medium-quality movies the modes occur at 58 and 0.35, and for low-quality movies at 32 and 1.9 million USD. The bimodal distribution even for low-quality movies shows that a low-quality movie can, occasionally, garner high box office sales.

To illustrate in greater detail how the proportions of high- and low-quality movies change with sales, the lower panel in Figure 8 plots the proportion of high- versus low-quality movies (i.e., the number of high-quality movies divided by the total number of high- and low-quality movies) at different levels of sales. Consistent with our theoretical predictions (see Figure 3D), high-quality movies are relatively more common than low-quality movies at both low and high sales levels.

[7] The null hypothesis that the means and variances of the two distributions are identical is rejected at a p-value of 0.01.

This, in turn, is why the middle dip occurs: low-quality movies are more common than high-quality movies at extremely low sales levels, but there is also a wide region of medium levels of sales where low-quality movies are most common.

Next, we consider the drama genre, for which the bimodality of box office sales is less obvious. A mixture distribution still fits the data better than a normal distribution at all three levels of quality, providing evidence of bimodality also in the drama genre. However, the lower panel in Figure 8 reveals a crucial difference between the drama and action/adventure genres. In the latter, low-quality movies are more frequent than high-quality movies at low and medium levels of box office sales, whereas in the former, low-quality movies are the most frequent type of movie only at low levels of box office sales. Thus, low-quality movies in the drama genre seldom reach medium levels of box office sales.

Why do low-quality movies reach medium levels of box office sales in the action/adventure genres but not in the drama genre? The reason seems not to be that the level of financial investment or number of theaters in which a movie is launched matter less in the latter than in the former. A low-quality movie launched in many theaters is substantially more likely to garner high box office sales, regardless of genre. However, investment patterns differ significantly across the different genres (see Figure 9). The distribution of the number of theaters in which movies are initially launched is bimodal in the action/adventure genres, with the largest mode at 3,000 theaters and another at 1,000.[8] This finding reflects the fact that most action/adventure movies have major releases to avoid box office failure and to compensate for their high production costs. The distribution of the number of initial theaters in the drama genre has a tendency toward low values.

A final clue is that the correlation between average review scores/ratings and the number of theaters in which a movie is initially launched is 0.23 for the action/adventure genres but minus 0.01 for the drama genre. Hence, the general public's preferences are more predictable for most action/adventure movies, and moviemakers may be confident in placing big bets on movies in these two genres. However, even in these genres, these predictions are highly imperfect, implying that moviemakers may invest substantially in what turn out to be low-quality movies with limited sales potential. Overall, a dip emerges in the action/adventure genres because low-quality movies that have received investment may achieve higher box office sales than high-quality movies that have received less investment. In the drama genre, moviemakers are less likely to invest substantially in perceived winners and there is no dip.

[8] This is a truncated distribution, as we only have data on the initial number of theaters for a subset of movies that sold well, whereas the overwhelming majority of movies for which we have no data were launched in only a few, and certainly fewer than 1,000, theaters.

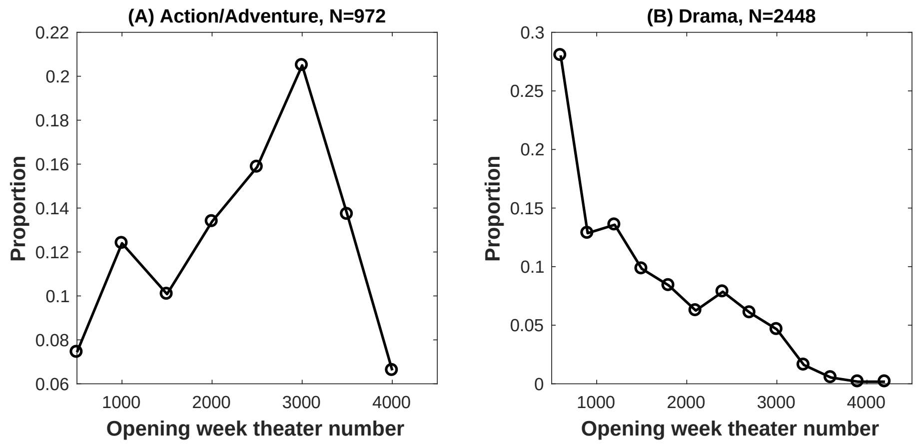

Figure 9 **Distribution of the number of theaters in which movies were initially launched**

6. Discussion

6.1. When Does a Dip Occur?

The models and the empirical illustration show that a dip in the quality-outcome association may occur when: i) both quality and other factors, which we label "resources", impact the outcome; ii) there is a bimodal distribution of resources (products/individuals receive either few or many resources), possibly generated by a reinforcing process. Moreover, the impact of resources is large enough for better resourced, low-quality products/individuals to achieve higher outcomes than more poorly resourced, high-quality products/individuals; and iii) allocation mistakes are possible, as low-quality products/individuals may receive a lot of resources and high-quality ones may not.

When do these conditions occur in reality? First, outcomes in many economic, cultural and social settings may be impacted by factors other than "quality". People or products that happen to be in the right place at the right time may perform better than superior alternatives that come later (Lieberman and Montgomery 1988, Salganik et al. 2006). However, this effect is not always strong enough to generate a dip. The advantages of early success may be fleeting, and high-quality late entrants may eventually outstrip early entrants (Liebowitz and Margolis 1995, Tellis et al. 2009, Zhu and Iansiti 2012, van de Rijt 2019). Factors beyond quality that impact on early success will have a larger impact on final outcomes when products and individuals that do well initially are given additional resources to increase their chances of future success (Merton, 1968), when early failures are eliminated and not given a second chance (Rosenbaum 1979), when products that become popular early on become well-known and acquire loyal followers, and when consumers/evaluators have strong incentives to select an offering that seems (even only marginally) better than others (Vincenz and van de Rijt 2016). For example, early outcomes have a significant effect on final outcomes in the movie industry, as a result of a) concentrating investment in movies that appear promising early on, and b) reducing the number of theaters in which a movie with poor opening-weekend box office sales is screened. In contrast, in settings where popularity only impacts evaluations, not awareness or resources allocated, being selected early on may only weakly increase the probability of being selected in the long term (Azoulay et al. 2014, van de Rijt 2019), and a dip is unlikely to be observed.

Second, a dip requires that the resources distributed differ substantially between actors, so that obtaining resources through means other than quality may propel a low-quality product/individual toward a better outcome. This bimodal resource allocation pattern may be generated by a strong reinforcing process, where initial success increases the chances of future success (Arthur et al. 1987). A bimodal resource allocation may also be the result of a deliberate choice to focus investments on promising candidates. In contrast, a dip is unlikely in settings where it makes sense to spread investments to a wide range of candidates, thus creating a level playing field. Third, a dip requires allocation mistakes, where low-quality products (individuals) are chosen over high-quality products (individuals) or are provided with more resources. Even minimal randomness or uncertainty may generate a dip: a dip will occur if there is some chance that two quality levels are mistaken (as in the resource allocation model), but the dip will then be small, in that it will only hold for a very narrow quality range. A dip occurring over a broad range of outcomes requires substantial randomness, in that the lowest-quality offering in this range has some chance of being confused with the upper-quality offering. Such randomness is realistic when there is high variance in outcomes owing to variability in the context and the presence of confounding factors (for example, whether other movies released on the same weekend are attractive) and when evaluations are based on small samples. In contrast, randomness in evaluations is less likely when outcomes have low variance and when quality can be accurately measured.

6.2. Implications for the Literature on Market Selection and Strategic Opportunities

If there is a dip, inferences about quality based on a naïve, more-successful-is-better heuristic may be misleading. Products or individuals with greater success but lower expected quality may be evaluated as having higher quality than those with less success but higher expected quality. Of course, not all evaluators care about expected quality. Investors (and funding agencies) may only care about continued success. If past success strongly increases the chances of future success, it may be rational to support a product that has succeeded in the past, regardless of its quality. However, expected quality matters if observers want to learn from successful products or identify talented people, and care about expected success in a *new* round of competition (for example, a new movie's expected box office sales).

A dip will only lead to mistaken inferences if not all evaluators are sophisticated enough to perfectly understand that a dip may occur. Sophisticated evaluators understand how $E[q_i|m_{i,t}]$ varies with $m_{i,t}$. It does not matter to them how quality varies with market share, whether increasing, decreasing or following a complicated pattern. They may also rely on other information, such as on firms' available resources, in making inferences about quality. Our claim (which may be incorrect in some contexts) is that not all evaluators understand or have access to the information necessary to correct for the dip. Naïve evaluators, who assume that greater success implies higher quality, will be misled by a dip. Sorenson and Waguespack (2006) offers evidence consistent with such naïve evaluations in the movie industry, showing that movie distributors prefer to work with partners whom they know well. There is also evidence that movies resulting from such collaborations have higher box office sales. However, the reason for this correlation is that distributors allocate more resources to these movies. When these resources are controlled for, the movies actually perform worse in box office sales. The implication is that some movie distributors are naïve and over-rely on misleading box office numbers without correcting for them.

Whether a dip leads to mistaken resource allocations also depends on whether a few sophisticated investors, who seek to identify undervalued actors, will make the dip disappear as a result of their arbitrage activities (thus helping to eliminate possible mismatches between resources and talent/quality). Matching of resources and product quality is likely to improve over time if investors are able to observe the outcomes of past competitions and attempt to identify "hidden gems" that did reasonably well with few resources. For example, moviemakers may be able to detect skilled directors who have performed well with few resources. More generally, if investors have access to information about the levels of resource received by different products and whether they did well initially, they should be able to determine more accurately which products and actors are of high quality. The dip is therefore less relevant to markets in which a significant amount of information is known about relevant actors (e.g., directors who have produced many movies), and more relevant to markets in which there is a constant inflow of new products/actors of uncertain quality.

If a dip persists and not all evaluators are sophisticated, this has interesting implications for evolutionary theories of the firm (Nelson and Winter 1982). Evolutionary theories of growth assume that firms that do well in competition are more efficient or produce higher-quality products than firms that do not. Profitable firms survive and expand, while unprofitable firms fail and contract. This differential growth is the mechanism by which selection operates and through which industries change over time (Nelson and Winter 1982). If a dip occurs and the resources to which firms have access depend on the outcomes of previous competition (i.e., firms are self-financed), the argument based on differential growth may break down. For example, a low-quality moviemaker may end up with a minor blockbuster, providing the funds necessary to bankroll the next movie. A high-quality moviemaker who has been unlucky, received few resources and produced a poorly selling movie will have few resources to invest in making the next movie. However, if the resources are provided by sophisticated investors, the dip may not matter, as such investors may see through the dip and allocate all their resources to firms with superior merit.

A dip also creates opportunities for more informed evaluators who understand the mechanism that generates it (Denrell et al. 2019). Such foresighted investors may benefit from understanding that quality is sometimes divorced from popularity (Csaszar and Laureiro-Martínez 2018). If market forces are too weak to eliminate dips, a dip may thus be a reliable source of arbitrage, and perhaps competitive advantage, through cognitive superiority (Cattani et al. 2018, Menon 2018)

6.3. Implications for the Literature on Status and Quality

The possibility of a dip has important implications for how status and quality may become decoupled (Gould 2002, Lynn et al. 2009). Suppose the status of actor j relates to the number of other actors who endorse or defer to that actor. Our model shows that if an endorsement is subject to a strong reinforcing process, there may be a dip in the status-quality association. Specifically, the model in Section 3 shows that a dip may occur if the probability that an evaluator will endorse actor j depends on both quality and the proportion of others who have already endorsed that actor. Existing models of status attainment can also be modified to generate a dip. Consider Lynn et al. (2009)'s model of the effects of social influence on status attainment. Following Gould (2002), they assume that the extent to which actor i endorses actor j depends on previous endorsements by other actors (endorsements also reflect a concern for reciprocity, p. 770, equation 4b). Lynn et al. (2009) assume that this social influence function is linear. Simulations of this model show that it does *not* generate a dip. However, a dip does occur if social influence is instead assumed to be a *non-linear* function of past attachments, matching our logit function (equation 1). Consistent with this assumption of a non-linear function, empirical studies show that social influence is indeed often a non-linear function of the number of past adopters (MacCoun 2012).

A dip represents a more fundamental decoupling of status and quality than shown in previous studies. While previous research has focused on situations where the correlation is close to zero (Lynn et al. 2009), our model shows that the correlation between status and quality may be negative for some status intervals. The existence of such a dip has interesting implications for interpretations of status and the persistence of a stratification system. Owing to the bimodal nature of outcomes in a system with a dip, many individuals will end up with a status level that does not reflect their true merit. Consider lucky individuals who achieve an unmerited high status. How will they react? One possibility is "imposter syndrome", if they sense that they have been awarded a higher status than they deserve. Another reaction is a desire to protect their privilege by erecting entry barriers and using their resources to increase their offsprings' chances of success. In systems with a dip, there are also many unlucky individuals whose status is much lower than they merit. These individuals are most likely to push to overturn the system, or rebel against it and create an alternative system of stratification. As an example of this type of rebellion, consider the Sundance Film Festival (Ortner 2012), which was founded in 1978 to attract film makers to Utah but quickly became a focal point for independent movies. Unless they are advertised and launched in many theaters, the box office sales of high-quality, independent movies are likely to be several orders of magnitude lower than the typical Hollywood blockbuster action movie. Frustrated by this inequality, independent moviemakers coordinated and sponsored the Sundance Film Festival as a venue where they could escape from the conventional use of box office sales as a status metric.

6.4. Implications for Performance Evaluation and Incentives

In a system with a dip, how can evaluators make more accurate inferences about quality from observed outcomes and avoid being fooled by the dip? One possibility is to rely on additional information about the resources to which actors have access. For example, moviemakers who have done well with a limited budget may be of higher quality than those who have achieved similar outcomes with large budgets. Even if evaluators do not have information about resources, they can use information about the sequence of outcomes as a proxy. For example, in the model presented in Section 3, we can condition on whether a product was chosen in the first period. Plotting the quality-market share association separately for products that were or were not chosen in the first period eliminates the dip in some cases. One example is when the market share is based on a limited number of periods (e.g., $t = 25$) and the first period has a considerable impact on all remaining periods.

Furthermore, if evaluators feel that a dip is undesirable, they may choose to limit the resources allocated to early winners, when uncertainty about quality is highest, to avoid a dip and allow for more accurate inferences about quality (see Bothner et al. (2011)). This may be especially important if evaluators are unsure about how much additional resource increases the chances of future success. If this information is known, evaluators can factor it in when evaluating outcomes; but if the effect is uncertain, allocating a large amount of resources to some individuals complicates evaluations of the outcomes. A dip also complicates performance-based incentives, because better performance does not necessarily merit more positive evaluations or rewards. Thus, individuals in the region of the dip may strive to worsen their performance (e.g., reduce it below the dip "danger zone") to be more favorably evaluated. To avoid such complications, management may decide to avoid performance-based incentives altogether (Milgrom and Roberts 1988).

6.5. Implications for Inequality

Our results may have important implications for interpreting social inequality. According to the meritocratic ideal, income reflects individual traits such as cognitive skill (Herrnstein and Murray 1994). Sociologists and some economists have argued that the association between individual traits and income is much weaker (Fischer et al. 1996, Bowles et al. 2001) because income is influenced by other factors including luck (Jencks 1979, Frank 2016), socio-economic background (Blau and Duncan 1967, Fischer et al. 1996) and adverse childhood experiences (Sharkey and Elwert 2011). Our model of the outcome-quality dip suggests a third possibility—the relationship between income and individual traits may be non-monotonic. A modified version of our model, in which income depends on individual traits and stochastic shocks and is subject to reinforcement, may generate an income-skill dip. Incomes may be subject to a process of cumulative disadvantage, as bad luck and initial setbacks are amplified in a vicious circle (Elman and Angela 2004, Schafer et al. 2011, Maroto 2012). For example, failure to graduate, combined with high debt, may leave college students in a poor financial situation, forcing them to accept any available job. If merit, such as high cognitive skill, cannot compensate for such bad luck, individuals with superior merit may be stuck at low income levels. Lucky or privileged individuals with medium cognitive skills may reach medium incomes, although high incomes may require high skills in many occupations. Overall, this suggests that the income-merit association may not only be low but negative for some medium income levels, an implication which can be tested with data on cognitive skills and income.

6.6. Limitations

One of the main limitations is empirical. Additional and better data to the movie box office are needed to assess how often a dip occurs, and when reinforcing processes are sufficiently strong. An experiment involving many participants adopting an object/product might also be useful to test more rigorously whether the mechanism in the model generates a dip. Furthermore, we assume that many people will not understand and will be misled by a dip, but this assumption should be examined experimentally. Theoretically, our mechanism needs further exploration. When does a dip occur? Is a bimodal allocation necessary, or can a dip occur with smoother distributions? We examined a non-linear Pólya urns model. Does a dip also occur for other reinforcing processes? In particular, can a preferential attachment process (Barabasi and Albert 1999) also generate a dip? The model in Section 3 assumed that consumers tend to choose products with the highest market share. This type of social influence arises through a variety of mechanisms, such as network externalities or because consumers rely on choices made by others to infer quality. If consumers rely on others' choices, the outcome depends on how sophisticated the consumers are. If they are sophisticated and understand the possibility of a dip, they may infer that a product with a low market share is of superior quality and choose this product. This would contradict the assumption that the value of b in the model is positive, which might undo the dip. Exit and market segmentation might also influence the dip. Actors at some quality levels may exit owing to anticipated (or past) poor evaluations by sophisticated evaluators. Exploring these dynamics formally would be an interesting avenue for future research.

7. Concluding Remarks

We have formally and empirically demonstrated how a bimodal resource allocation, typically enabled by reinforcing processes, can generate a middle dip in the outcome-quality association. The intuition is that all actors or products who have been provided with many resources early on will reach at least a medium-level outcome. Only low-quality actors or products will remain at a medium-level outcome despite being well-resourced. An outcome just below medium level is more impressive because it is likely to have been achieved despite having few resources. A dip complicates evaluations, because evaluators must understand that a low outcome may indicate higher expected quality than a medium-level outcome. To eliminate the dip, one should provide early winners with extra resources only when they are clearly better than the rest. Ironically, people rely on initial popularity precisely when quality is difficult to evaluate. When early winners are provided with substantial resources, a dip may occur, distorting the meritocratic ideal.

References

Arthur B (1989) Competing technologies, increasing returns, and lock-in by historical events. *The Economic Journal* 99(394):116–131.

Arthur B, Ermoliev YM, Kaniovski YM (1987) Path-dependent processes and the emergence of macro-structure. *European Journal of Operational Research* 30(3):294–303.

Asch SE (1955) Opinions and social pressure. *Scientific American* 193(5):31–35.

Azoulay P, Stuart T, Wang Y (2014) Matthew: Effect or fable? *Management Science* 60(1):92–109.

Baker WE, Faulkner RR (1991) Role as resource in the hollywood film industry. *American Journal of Sociology* 97(2):279–309.

Barabasi A, Albert R (1999) Emergence of scaling in random networks. *Science* 286(5439):509–512.

Baron J, Hershey J (1988) Outcome bias in decision evaluation. *Journal of Personality and Social Psychology* 54(4):569–79.

Bikhchandani S, Hirshleifer D, Welch I (1992) A theory of fads, fashion, custom, and cultural change in informational cascades. *Journal of Political Economy* 100(5):992–1026.

Blau PM, Duncan OD (1967) *The American Occupational Structure* (New York, NY: Wiley).

Bothner M, Podolny J, Smith E (2011) Organizing contests for status: The matthew effect versus the mark effect. *Management Science* 57(3):439–457.

Bowles S, Gintis H, Osborne M (2001) The determinants of individual earnings: Skills, preferences, and schooling. *Journal of Economic Literature* 34(4):1137–1176.

Brown AL, Camerer CF, Lovallo D (2012) To review or not to review? limited strategic thinking at the movie box office. *American Economic Journal: Microeconomics* 4(2):1–26.

Bushman RM, Indjejikian RJ, Smith A (1996) Ceo compensation: The role of individual performance evaluation. *Journal of Accounting and Economics* 21(2):161–193.

Callander S, Hörner J (2009) The wisdom of the minority. *Journal of Economic theory* 144(4):1421–1439.

Carroll GR, Harrison JR (1994) On the historical efficiency of competition between organizational populations. *American Journal of Sociology* 100(3):720–749.

Cattani G, Sands D, Porac J, Greenberg J (2018) Competitive sensemaking in value creation and capture. *Strategy Science* 3(1):632–657.

Chase ID (1980) Social process and hierarchy formation in small groups: A comparative perspective. *American Sociological Review* 45(6):905–924.

Christensen M, Knudsen T (2010) Design of decision-making organizations. *Management Science* 56(1):71–89.

Colyvas JA (2012) Performance metrics as formal structures and through the lens of social mechanisms: When do they work and how do they influence? *American Journal of Education* 118(2):167–197.

Correll SJ, Ridgeway CL, Zuckerman EW, Jank S, Jordan-Bloch S, Nakagawa S (2017) It's the conventional thought that counts: How third-order inference produces status advantage. *American Sociological Review* 82(2):297–327.

Csaszar F, Laureiro-Martínez D (2018) Individual and organizational antecedents of strategic foresight: A representational approach. *Strategy Science* 3(3):513–532.

Csaszar FA (2013) An efficient frontier in organization design: Organizational structure as a determinant of exploration and exploitation. *Organization Science* 24(4):1083–1101.

Dana J, Dawes R, Peterson N (2013) Belief in the unstructured interview: The persistence of an illusion. *Judgment and Decision Making* 8(5):512–520.

Dawes RM, Faust D, Meehl PE (1989) Clinical versus actuarial judgment. *Science* 243(4899):1668–1674.

De Vany A (2004) *Hollywood Economics: How Extreme Uncertainty Shapes the Film Industry* (New York, NY: Routledge).

Denrell J, Fang C, Liu C (2019) In search of behavioral opportunities from misattributions of luck. *Academy of Management Review* 44(4):896–915.

Denrell J, Liu C (2012) Top performers are not the most impressive when extreme performance indicates unreliability. *Proceedings of the National Academy of Sciences of the United States of America* 109(24):9331–9336.

Dubé JPH, Hitsch GJ, Chintagunta P (2010) Tipping and concentration in markets with indirect network effects. *Marketing Science* 29(2):216–249.

Elberse A (2013) *Blockbusters: Why Big Hits - and Big Risks - Are the Future of the Entertainment Business* (New York, NY: Henry Holt and Company).

Elberse A, Eliashberg J (2003) Demand and supply dynamics for sequentially released products in international markets: The case of motion pictures. *Marketing Science* 22(3):329–354.

Eliashberg J, Jonker JJ, Sawhney MS, Wierenga B (2000) Moviemod: An implementable decision-support system for prerelease market evaluation of motion pictures. *Marketing Science* 19(3):226–243.

Elman C, Angela MO (2004) The Race is to the swift: Socioeconomic origins, adult education, and wage attainment1. *American Journal of Sociology* 110(1):123–160.

Elster J, Moene KO (1989) Introduction. Elster J, Moene KO, eds., *Alternatives to Capitalism*, chapter 1, 1–35 (Cambridge, U. K.: Cambridge University Press).

Espeland WN, Sauder M (2007) Rankings and reactivity: How public measures recreate social worlds. *American Journal of Sociology* 113(1):1–40.

Eyster E, Rabin M (2014) Extensive imitation is irrational and harmful. *The Quarterly Journal of Economics* 129(4):1861–1898.

Fischer CS, Jankowski MS, Lucas SR, Swidler A, Voss K (1996) *Inequality by Design: Cracking the Bell Curve Myth* (Princeton, NJ: Princeton University Press).

Frank RH (2016) *Success and Luck: Good Fortune and the Myth of Meritocracy* (Princeton, NJ: Princeton University Press).

Galton F (1907) Vox populi. *Nature* 74(2):450–451.

Gan H, Park MS, Suh S (2020) Non-financial performance measures, ceo compensation, and firms' future value. *Journal of Business Research* 110(C):213–227.

Gould R (2002) The origins of status hierarchies: A formal theory and empirical test. *American Journal of Sociology* 107(5):1143–1178.

Gretz RT, Basuroy S (2013) Why quality may not always win: The impact of product generation life cycles on quality and network effects in high-tech markets. *Journal of Retailing* 89(3):281–300.

Groysberg B, Sant L, Abrahams R (2008) When 'stars' migrate, do they still perform like stars? *MIT Sloan Management Review* 50(1):41–46.

Harris M, Raviv A (1978) Some results on incentive contracts with applications to education and employment, health insurance, and law enforcement. *The American Economic Review* 68(1):20–30.

Herrnstein R, Loewenstein G, Prelec D, Vaughan W (1993) Utility maximization and melioration: Internalities in individual choice. *Journal of Behavioral Decision Making* 6(1):149 – 185.

Herrnstein RJ, Murray C (1994) *Bell Curve: Intelligence and Class Structure in American Life* (New York, NY: Simon and Schuster).

Hill BM, Lane D, Sudderth W, et al. (1980) A strong law for some generalized urn processes. *The Annals of Probability* 8(2):214–226.

Holbrook MB, Addis M (2007) Taste versus the market: An extension of research on the consumption of popular culture. *Journal of Consumer Research* 34(3):415–424.

Holmstrom B (1979) Moral hazard and observability. *Bell Journal of Economics* 10(1):74–91.

Hopwood A (1972) Empirical study of role of accounting data in performance evaluation. *Journal of Accounting Research* 10(1):156–182.

Jencks C (1979) *Who Gets Ahead? The Determinants of Economic Success in America* (New York, NY: Basic Books).

Karlin S, Rubin H (1956) The theory of decision procedures for distributions with monotone likelihood ratio. *The Annals of Mathematical Statistics* 27(2):272–299.

Katz M, Shapiro C (1986) Technology adoption in the presence of network externalities. *Journal of Political Economy* 94(4):822–841.

Keum DD, See KE (2017) The influence of hierarchy on idea generation and selection in the innovation process. *Organization Science* 28(4):653–669.

Kovacs B, Sharkey AJ (2014) The paradox of publicity: How awards can negatively affect the evaluation of quality. *Administrative Science Quarterly* 59(1):1–33.

Larrick RP, Soll JB (2006) Intuitions about combining opinions: Misappreciation of the averaging principle. *Management Science* 52(1):111–127.

Lazear E, Rosen S (1981) Rank-order tournaments as optimum labor contracts. *Journal of Political Economy* 89(5):841–64.

Lieberman M, Montgomery D (1988) First-mover advantages. *Strategic Management Jounral* 9(S1):41–58.

Liebowitz SJ, Margolis SE (1995) Path dependence, lock-in, and history. *Journal of Law, Economics and Organization* 11(1):205–226.

Liu C (2020) *Luck: A Key Idea for Business and Society* (Routledge).

Lynn FB, Podolny JM, Tao L (2009) A sociological (de) construction of the relationship between status and quality. *American Journal of Sociology* 115(3):755–804.

MacCoun RJ (2012) The burden of social proof: Shared thresholds and social influence. *Psychological Review* 119(2):345–372.

Manso G (2011) Motivating innovation. *Journal of Finance* 66(5):1823–1860.

March JG (1994) *Primer on Decision Making*. How Decisions Happen (Free Press).

March JG, Sutton RI (1997) Crossroads—organizational performance as a dependent variable. *Organization Science* 8(3):698–706.

Maroto M (2012) The scarring effects of bankruptcy: Cumulative disadvantage across credit and labor markets. *Social Forces* 91(1):99–130.

Marshall G, Mowen J (1993) An experimental investigation of the outcome bias in salesperson performance evaluations. *Journal of Personal Selling & Sales Management* 13(1):31–47.

Massey C, Thaler R (2013) The loser's curse: Decision making and market efficiency in the national football league draft. *Management Science* 59(7):1479–1495.

Mazmanian M, Beckman C (2018) Making your numbers: Engendering organizational control through a ritual of quantification. *Organization Science* 29(2):357–379.

Menon A (2018) Bringing cognition into strategic interactions: Strategic mental models and open questions. *Strategic Management Journal* 39(2):168–192.

Merton RK (1968) The Matthew effect in science. *Science* 159(3810):56–63.

Meyer JW, Rowan B (1977) Institutionalized organizations: Formal structure as myth and ceremony. *American Journal of Sociology* 83(2):340–363.

Meyer MW (2002) *Rethinking Performance Measurement: Beyond the Balanced Scorecard* (New York, NY: Cambridge University Press).

Milgrom PR (1981) Good news and bad news: Representation theorems and applications. *The Bell Journal of Economics* 12(2):380–391.

Milgrom PR, Roberts J (1988) An economic approach to influence activities in organizations. *American Journal of Sociology* 94(4):S154–S179.

Nair H, Chintagunta P, Dubé JP (2004) Empirical analysis of indirect network effects in the market for personal digital assistants. *Quantitative Marketing and Economics (QME)* 2(1):23–58.

Nelson RR, Winter SG (1982) *An evolutionary theory of economic change* (Harvard University Press).

Ortner SB (2012) Against hollywood: American independent film as a critical cultural movement. *HAU: Journal of Ethnographic Theory* 2(2):1–21.

Ouchi WG (1979) A conceptual framework for the design of organizational control mechanism. *Management Science* 25(9):833–848.

Park S (2004) Quantitative analysis of network externalities in competing technologies: the vcr case. *Review of Economics and Statistics* 86(4):937–945.

Purkayastha S (1998) Simple proofs of two results on convolutions of unimodal distributions. *Statistics & ProbabilityLetter* 39(1):97–100.

Rosenbaum JE (1979) Tournament mobility: Career patterns in a corporation. *Administrative Science Quarterly* 24(2):220–241.

Ross L, Nisbett RE (1991) *The person and the situation* (New York, NY: McGraw Hill).

Sah R, Stiglitz J (1986) The architecture of economic systems: Hierarchies and polyarchies. *American Economic Review* 76(4):716–27.

Salganik M, Dodds P, Watts DJ (2006) Experimental study of inequality and unpredictability in an artificial cultural market. *Science* 311(5762):854–856.

Schafer MH, Ferraro KF, Mustillo SA (2011) Children of misfortune: Early adversity and cumulative inequality in perceived life trajectories. *American Journal of Sociology* 116(4):1053–1091.

Shapira Z, Shaver J (2014) Confounding changes in averages with marginal effects: How anchoring can destroy economic value in strategic investment assessments. *Strategic Management Journal* 35(10):1414–1426.

Sharkey P, Elwert F (2011) The legacy of disadvantage: Multigenerational neighborhood effects on cognitive ability. *American Journal of Sociology* 116(6), ISSN 0003-1224.

Simon HA (1947) *Administrative Behavior* (New York, NY: Free Press).

Simonsohn U (2018) Two lines: A valid alternative to the invalid testing of u-shaped relationships with quadratic regressions. *Advances in Methods and Practices in Psychological Science* 1(4):538–555.

Skvoretz J, Faust K, Fararo TJ (1996) Social structure, networks, and E-state structuralism models. *Journal of Mathematical Sociology* 21(1):57–76.

Sorenson O, Waguespack DM (2006) Social structure and exchange: Self-confirming dynamics in hollywood. *Administrative Science Quarterly* 51(4):560–589.

Starbuck W (2005) Performance measures: Prevalent and important but methodologically challenging. *Journal of Management Inquiry* 14(1):280–286.

Sterman JD (1989) Misperceptions of feedback in dynamic decision making. *Organizational Behavior and Human Decision Processes* 43(3):301–335.

Surowiecki J (2004) *The Wisdom of Crowds* (New York, NY: Doubleday).

Tellis G, Yin E, Niraj R (2009) Does quality win? Network effects versus quality in high-tech markets. *Journal of Marketing Research* XLVI(2):135–149.

Townley B, Cooper D, Oakes L (2003) Performance measures and the rationalization of organizations. *Organization Studies* 24(4):1045–1071.

van de Rijt A (2019) Self-correcting dynamics in social influence processes. *American Journal of Sociology* 124(5):1468–1495.

van de Rijt A, Akin IA, Willer R, Feinberg M (2016) Success-breeds-success in collective political behavior: Evidence from a field experiment. *Sociological Science* 3(1):940–950.

van de Rijt A, Kang SM, Restivo M, Patil A (2014) Field experiments of success-breeds-success dynamics. *Proceedings of the National Academy of Sciences* 111(19):6934–6939.

Vincenz F, van de Rijt A (2016) Arbitrary inequality in reputation systems. *Scientific Reports* 6(38304):1–5.

Weibull JW, Mattsson LG, Voorneveld M (2007) Better May be Worse: Some Monotonicity Results and Paradoxes in Discrete Choice Under Uncertainty. *Theor Decis* 63(2):121–151.

Wernerfelt B, Karnani A (1987) Competitive strategy under uncertainty. *Strategic Management Journal* 8(2):187–194.

Zhu F, Iansiti M (2012) Entry into platform-based markets. *Strategic Management Journal* 33(1):88–106.

Zuckerman EW, Kim TY (2003) The critical trade-off: identity assignment and box-office success in the feature film industry. *Industrial and Corporate Change* 12(1):27–67.

Appendices

Appendix A: Fixpoint Analysis

LEMMA EC.1. *Suppose the utility of selecting product 1 in period t is $u_{1,t} = aq_1 + bm_{1,t-1} + \varepsilon_{1,t}$ and the utility of selecting product 2 is $u_{2,t} = a0.5 + b(1 - m_{1,t-1}) + \varepsilon_{2,t}$, where $\varepsilon_{j,t}$ are independent draws from an error distribution. Whenever the error distribution $f(\varepsilon)$ is an unimodal distribution (i.e., $f(\varepsilon)$ is increasing for $\varepsilon < \beta$ and decreasing for $\varepsilon > \beta$ for some value β) and has infinite support in at least one tail, then the probability of selecting product 1, $P_{1,t}$, is i) an increasing function of $m_{1,t-1}$ and q_1 ii) $0 < P_{1,t} < 1$ for all $m_{1,t-1} \in [0,1]$, iii) a convex function for $m_{1,t-1} < 0.5 - a(q_1 - 0.5)/2b$ and a concave function for $m_{1,t-1} > 0.5 - a(q_1 - 0.5)/2b$, iv) $P_{1,t} = 0.5$ whenever $m_{1,t-1} = 0.5$ and $q_1 = 0.5$.*

Proof The probability of selecting product 1 is $P_{1,t} = Pr(u_{1,t} > u_{2,t})$ which equals

$$Pr(u_{1,t} > u_{2,t}) = Pr(\varepsilon_{1,t} - \varepsilon_{2,t} > -a(q_1 - 0.5) - 2b(m_{1,t-1} - 0.5)) \tag{EC.1}$$

i) It follows that $P_{1,t}$ is an increasing function of $m_{1,t-1}$ and q_1. ii) Moreover, $0 < P_{1,t} < 1$ for all $m_{1,t-1} \in [0,1]$ since $Pr(\varepsilon_{1,t} - \varepsilon_{2,t} > -a(q_1 - 0.5) + 2b(m_{1,t-1} - 0.5)) \in (0,1)$ for all $m_{1,t-1} \in [0,1]$ when $f(\varepsilon)$ has infinite support in at least one tail.

iii) Purkayastha (1998) Theorem 2.2 shows that if X_1 and X_2 are independent random variables with the same unimodal distribution then $X_1 - X_2$ is also unimodal with mode equal to zero. Hence, $\varepsilon_{1,t} - \varepsilon_{2,t}$ is unimodal with mode equal to zero, i.e., the density of $\varepsilon_{1,t} - \varepsilon_{2,t}$, $g(\varepsilon_{1,t} - \varepsilon_{2,t})$, is an increasing function for $\varepsilon_{1,t} - \varepsilon_{2,t} < 0$ and a decreasing function for $\varepsilon_{1,t} - \varepsilon_{2,t} > 0$ Moreover, the random variable $\varepsilon_{1,t} - \varepsilon_{2,t}$ has mean zero and has a symmetric distribution, which follows from the fact that $\varepsilon_{1,t}$ and $\varepsilon_{2,t}$ are identically distributed.

The derivative of $P_{1,t}$ with respect to $m_{1,t-1}$ equals

$$\frac{\delta}{\delta m_{1,t-1}}[1 - G(-a(q_1 - 0.5) - 2b(m_{1,t} - 0.5))] = 2bg(-a(q_1 - 0.5) - 2b(m_{1,t-1} - 0.5)), \tag{EC.2}$$

where $g(d)$ is the density of $d = \varepsilon_{1,t} - \varepsilon_{2,t}$. The second derivative is $-4b^2 g'(-a(q_1 - 0.5) + 2b(m_{1,t-1} - 0.5))$. Because $g()$ is unimodal, with $g'(d) > 0 (< 0)$ for $d < 0 (> 0)$, it follows that $P_{1,t}$ is a convex (concave) function of $m_{1,t}$ whenever $-a(q_1 - 0.5) - 2b(m_{1,t-1} - 0.5) < 0 (> 0)$. That is, $P_{1,t}$ is a convex function whenever $m_{1,t-1} < 0.5 - a(q_1 - 0.5)/2b$ and a concave function whenever $m_{1,t-1} > 0.5 - a(q_1 - 0.5)/2b$.

iv) If $m_{1,t-1} = 0.5$ and $q_1 = 0.5$, then $P_{1,t} = Pr(\varepsilon_{1,t} - \varepsilon_{2,t} > 0)$. Because $\varepsilon_{1,t} - \varepsilon_{2,t}$ is symmetric around zero it follows that $P_{1,t} = 0.5$.

Theorem EC.1. *Suppose the utility of selecting product* 1 *in period* t *is* $u_{1,t} = aq_1 + bm_{1,t-1} + \varepsilon_{1,t}$ *and the utility of selecting product 2 is* $u_{2,t} = a0.5 + b(1 - m_{1,t-1}) + \varepsilon_{2,t}$*, where* $\varepsilon_{j,t}$ *are independent draws from an unimodal error distribution with infinite support in at least one of the tails. Let* $g(d)$ *be the density of* $d = \varepsilon_{1,t} - \varepsilon_{2,t}$*. Let* m_1^* *be the market share firm 1 converges to as* $t \to \infty$*. i) Whenever* $b < 1/2g(0)$*,* $E[q_1|m_1^*]$ *is an increasing function of* m_1^* *for all* $m_1^* \in [0,1]$*. ii) Whenever* $b > 1/2g(0)$*,* $E[q_1|m_1^*]$ *is an increasing function of* m_1^* *for all* $m_1^* \in (0, m_l)$*, where* $0 < m_l < 0.5$*, and* $E[q_1|m_i^*]$ *is an increasing function of* m_1 *for all* $m_1^* \in (m_h, 1)$*, where* $0.5 < m_h < 1$*, but* $E[q_1|m_i^*]$ *"dips" in the sense that* $E[q_1|m_1^* = m_l] > E[q_1|m_1^* = m_h]$*.*

Proof Let $C(z_1, q_1)$ be the probability that the customer in period $t+1$ chooses the product of firm 1 when $m_{1,t} = z_1$ and the quality of firm 1 is q_1. That is $C(z_1, q_1) = Pr(u_{1,t+1} > u_{2,t+1})$. To simplify the notation we often write the choice probability as $C(z_i)$ and only include the quality of the agent when necessary for the argument. When $\varepsilon_{j,t}$ are independent draws from an unimodal error distribution with infinite support in at least one of the tails, then it follows by Lemma EC.1 that $0 < C(z_1) < 1$ for all $z_1 \in [0,1]$. It follows that there is at least one fixpoint of the mapping $z_i = C(z_i)$ in the interval $z_1 \in [0,1]$ (z_1 cannot lie below $C(z_1)$ for all values $z_1 \in [0,1]$ or lie above $C(z_1)$ for all values $z_1 \in [0,1]$).

Let X be the set of stable fixpoints of the mapping $C(z_1) = z_1$, i.e., points at which $C(z_i)$ crosses z_i from above. Let Y be the set of unstable-fixpoints of the mapping $C(z_1) = z_1$, i.e., points at which $C(z_i)$ crosses z_i from above. Theorem 6.1 in Hill et al. (1980) imply that whenever $0 < C(z_1) < 1$ for all $z_1 \in [0,1]$, then, for each $r \in X$, there is a positive probability that z_1 converges to r, while, for each $r \in Y$, the probability that z_1 converges to r is zero.

Let z_1^1 be the "first" fixpoint of the mapping $C(z_1) = z_1$; "first" in the sense that z_1^1 is closest to zero: $z_1^1 < z_1^k$ for any other fixpoint k. Because $C(z_1 = 0) > 0$, $C(z_i)$ crosses z_i from above at z_1^1 ($C(z_1) > z_1$ for all $z_1 < z_1^1$). Another fixpoint $z_1^2 > z_1^1$ can only exists if $C(z_1)$ increases faster than z_1 does, otherwise $C(z_1)$ remains below z_1 for all $z_1 > z_1^1$. Moreover, $C(z_1)$ can only increase faster than z_1 does if the derivative of $C(z_1)$ with respect to z_1 is larger than 1 for some $z_1 \in (0,1)$. The derivative of $C(z_1)$ with respect to z_1 equals (see Lemma EC.1)

$$\frac{\delta C(z_1)}{\delta z_1} = 2bg(-a(q_1 - 0.5) - 2b(z_1 - 0.5)). \qquad \text{(EC.3)}$$

Because $g()$ is unimodal with mode equal to zero (Lemma EC.1), the derivative reaches a maximum at the value of z_1 for which $z_1 = 0.5 - a(q_1 - 0.5)/2b$. The value of the derivative at this value of z_1 is $2bg(0)$. A necessary condition for having more than one fixpoint is thus that $2bg(0) > 1$, i.e., $b > 1/2g(0)$.

If $b < 1/2g(0)$, the derivative of $C(z_1)$ with respect to z_1 is always below 1, implying there is only one fixpoint of the mapping $C(z_1) = z_1$. If there is only one fixpoint, the value of z_1 at which the fixpoint occurs is increasing in q_1 because $C(z_1, q_1)$ is increasing in q_1 (Lemma EC.1i). It follows that whenever $b < 1/2g(0)$, there is a unique fixpoint that the market share of firm 1 converges to as $t \to \infty$ and the value of this fixpoint is increasing in q_1. Thus, whenever $b < 1/2g(0)$, $E[q_1|m_1^*]$ is an increasing function of m_1^* for all $m_1^* \in [0,1]$.

Suppose now that $b > 1/2g(0)$. Consider the special case of $q_i = 0.5$. We first show that for this special case there are three fixpoints (two of which are stable). First, note that there is a fixpoint at $z_1 = 0.5$ since $C(z_1 = 0.5, q_1 = 0.5) = 0.5$ (Lemma EC.1iv). Denote this fixpoint by z_1^m. Second, there exists another fixpoint at $z_1 > 0.5$. To show this, note that because $\delta C(z_i)/\delta z_i > 1$ at $z_i = 0.5$, there exists a region $z_1 \in (0.5, 0.5+\varepsilon)$ in which $C(z_1) > z_1$. Because $C(z_1)$ is an increasing function and $C(z_1 = 1) < 1$, there must exist another fixpoint $z_1^h > 0.5$. There cannot be more than one fixpoint for $z_1 > 0.5$, however. The reason is that $C(z_1)$ is a concave function for $z_1 > 0.5$ (Lemma EC.1iii), implying that the value of the derivative is a decreasing function. Hence, once $\delta C(z_1)/\delta z_1$ drops below one it stays below one, implying that $C(z_1)$ remains below z_1 for all $z_1 > z_1^h$. A similar argument shows that when $b > 1/2g(0)$ there exists a third fixpoint z_1^l in the interval $z_1 \in (0, 0.5)$. A fixpoint of the mapping $z_1 = C(z_1)$ is stable at $z_1 = z_1^*$ if $\delta C(z_1)/\delta z_1$ evaluated at $z_1 = z_1^*$ is smaller than one and unstable otherwise. It follows that when $b > 1/2g(0)$ then the two fixpoints z_i^l and z_i^h are stable while z_1^m is unstable. If $b < 1/2g(0)$, however, there is only one fixpoint and because $\delta C(z_1)/\delta z_1 < 1$ for all z_1 this fixpoint is stable.

We next show that whenever $b > 1/2g(0)$, there are also two stable fixpoints for values of q_1 close to 0.5. That is, there exists an interval around $q_1 = 0.5$, $I = (0.5 - e, 0.5 + e)$ where $e \in (0, 0.5)$, such that for all $q_1 \in I$ the mapping $z_1 = C(z_1)$ has two stable fixpoints in $z_1 \in [0,1]$. To show this, consider first qualities just above 0.5. We know that $C(z_1 = 0, q_1 = 0.5 + e) > 0$ (Lemma EC.1ii). We also know that $C(z_1 = 0.5, q_1 = 0.5 + e) > C(z_1 = 0.5, q_1 = 0.5) = 0.5$ (Lemma EC.1). Let z_1^* be a value of $z_1 < 0.5$ at which $C(z_i, q_1 = 0.5) - z_i$ is the most negative, i.e., where $C(z_i, q_1 = 0.5)$ is the farthest below z_i (see Figure EC.1A). Note that at z_1^* the choice probability for the higher quality level will be higher: $C(z_1^*, q_1 = 0.5 + e) > C(z_1^*, q_1 = 0.5)$. By continuity, however, we can find some increase in quality from 0.5, i.e., find some value of $e \in (0, 0.5)$ such that $C(z_1 = z_1^*, q_1 = 0.5 + e) < z_1^*$, i.e., so that the choice probability at a market share of z_1^* is still below z_1^* (see Figure EC.1A). For this value of e, the function $C(z_1, q_1 = 0.5 + e) - z_1$ crosses the zero line at least once before z_1^*, because it started (at $z_1 = 0$) above zero. The function $C(z_1, q_1 = 0.5 + e) - z_1$ has to cross the zero line again before $z_1 = 0.5$ because $C(z_1 = 0.5, q_1 = 0.5 + e) - 0.5 > C(z_1 = 0.5, q_1 = 0.5) - 0.5 = 0$. It cannot cross the zero line more than twice for $z_1 \le 0.5$, however, because $C(z_1, q_1 = 0.5 + e)$ is increasing and has a derivative with respect to z_1 which is initially increasing and then decreasing in z_1 (Lemma EC.1iii). Finally, there has to be another fixpoint at market share above 0.5 ($z_1 > 0.5$) because $C(z_1 = 1, q_1 = 0.5 + e) < 1$. This fixpoint is stable because the function $C(z_1, q_1 =$

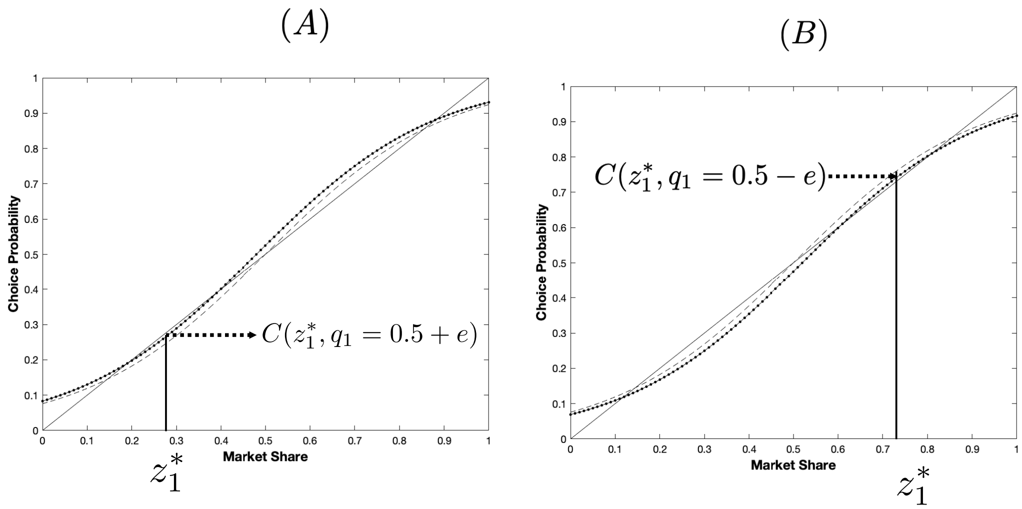

Figure EC.1 Illustration of Proof

$0.5+e)-z_1$ crosses the zero line from above (the derivative is decreasing after the second unstable fixpoint). Thus, there exists some value of e such that the function $C(z_1, q_1 = 0.5+e) - z_1$ has two stable fixpoints, one occurring before $z_1 = 0.5$ and one occurring after.

Consider next qualities just below 0.5. We know that there is some value $z_1^b < 0.5$ such that $C(z_i = z_1^b, q_1 = 0.5) = z_1^b$, i.e., there is a fixpoint before $z_1 = 0.5$ when $q_1 = 0.5$. Because $C(z_1^b, q_1 = 0.5 - e) < C(z_1^b, q_1 = 0.5)$ and $C(z_1 = 0, q_1 = 0.5 - e) > 0$, there has to be a fixpoint below $z_1 = 0.5$ for $q_1 = 0.5 - e$ (see Figure EC.1B). There can only be one fixpoint before $z_1 = 0.5$, however, since $C(z_1, q_1 = 0.5 - e) < C(z_1, q_1 = 0.5)$, i.e., the choice probability is always lower for lower quality and for quality equal to 0.5 the second fixpoint is at $q_1 = 0.5$. Let z_1^* be a value of $z_1 > 0.5$ at which $C(z_i, q_1 = 0.5) - z_i$ is the most positive, i.e., where $C(z_i, q_1 = 0.5)$ is the farthest above z_i (see Figure EC.1B). Note that at z_1^* the choice probability for the lower quality level will be lower: $C(z_1^*, q_1 = 0.5 - e) < C(z_1^*, q_1 = 0.5)$. By continuity, however, we can find some decrease in quality from 0.5, i.e., find some value of $e \in (0, 0.5)$ such that $C(z_1 = z_1^*, q_1 = 0.5 - e) > z_1^*$, i.e., so that the choice probability at a market share of z_1^* is still above z_1^*. Because $C(z_1 = 0, q_1 = 0.5 - e) > 0$ and $C(z_1 = 1, q_1 = 0.5 - e) < 1$, there has to be three fixpoints for this value of e. Due to the concavity of the choice function after the third fixpoint there can only be three fixpoints and the first and the third are stable. Thus, there exists some value of e such that the function $C(z_1, q_1 = 0.5 - e) - z_1$ has two stable fixpoints, one occurring before $z_1 = 0.5$ and one occurring after.

This argument shows that whenever $b > 1/2g(0)$ there exists some interval around $q_1 = 0.5$ such that for qualities in this interval there are two stable fixpoints, one below 0.5 and one above, both increasing in q_1.

Next we show that for firms with qualities outside this interval there exists only one fixpoint. Consider first firms with qualities $q_1 > 0.5$ which have no fixpoint at $z_1 < 0.5$, i.e., for which there is no value of $e \in (0, 0.5)$ and $z_1 < 0.5$ such that $C(z_1, q_1 = 0.5 + e) < z_1$. For these firms, there is at least once fixpoint at some z_1 since $C(z_1 = 0, q_1 = 0.5 + e) > 0$ and $C(z_1 = 1, q_1 = 0.5 + e) < 1$. This fixpoint has to occur at a value $z_1 > 0.5$. Only one fixpoint can occur because the inflection point of $C(z_1, q_1 = 0.5 + e)$, at which the second derivative changes from positive to negative, occurs at a value $z_1 < 0.5$ for $q_1 > 0.5$ (Lemma EC.1iii). For all $z_1 > 0.5$, the derivative is thus decreasing. Hence, once the function $C(z_1, q_1 = 0.5 + e)$ crosses the line z_1 it remains below. Because it crosses the line from above, the fixpoint is stable.

Consider next firms with qualities $q_1 < 0.5$ for which there is no $e \in (0, 0.5)$ such that $C(z_1, q_1 = 0.5 - e) \geq z_1$ for some value $z_1 > 0.5$. For these firms, there is at least one fixpoint since $C(z_1 = 0, q_1 = 0.5 + e) > 0$ and $C(z_1 = 1, q_1 = 0.5 + e) < 1$. This fixpoint has to occur at a value $z_1 < 0.5$ by assumption. Only one stable fixpoint can occur because the derivative of $C(z_1, q_1 = 0.5 + e)$ is increasing for all $z_1 < 0.5$ when $q_1 < 0.5$ (Lemma EC.1iii).

Overall, when $b > 2$ there are three regions of quality. i) For firms with qualities $q_1 \in [0, 0.5 - e^l)$ there is one stable fixpoint, below 0.5, the value of which is increasing in q_1. ii) For firms with qualities $q_1 \in (0.5 - e^l, 0.5 + e^h)$, there are two stable fixpoints, one below and one above 0.5, the values of which are increasing in q_1. iii) For firms with qualities $q_1 \in (0.5 + e^h, 1]$ there is one stable fixpoint, above 0.5, the value of which is increasing in q_1.

These three regions of quality map into three regions of market share. Consider firms with quality $q_1 \in [0, 0.5 - e^l)$, where there is one fixpoint. Each quality level maps onto a unique market share, $m_1^* \in (0, m_l)$, that a firm with this quality converges to. Moreover, no other firm with quality larger than $0.5 - e^l$ can converge to such a market share because its choice probability is larger than m_l at $z_1 = m_l$. There is thus a one-to-one mapping between qualities $q_1 \in [0, 0.5 - e^l)$ and the market share, $m_1^* \in (0, m_l)$, where $m_{l,1} < 0.5$, a firm converges to. It follows that $E[q_1|m_1^*]$ is an increasing function of m_1^* for all $m_1^* \in (0, m_l)$. For similar reasons, there is also a one-to-one mapping between quality $q_1 \in [0.5 + e^h, 1)$ and the market share a firm converges to, $m_1^* \in (m_h, 1)$, where $0.5 < m_h < 1$. It follows that $E[q_1|m_1^*]$ is an increasing function of m_1^* for all $m_1^* \in (m_h, 1)$.

Consider, next, qualities in the interval $q_1 \in (0.5 - e^l, 0.5 + e^h)$, for which there are two stable fixpoints. Focus first on the lower fixpoint for each q_1. It is also increasing in q_1. Let m_{low} be the largest value of this lower fixpoint, associated with quality equal to $0.5 + e^h$. We know that $m_{low} < 0.5$. We also know that the lowest fixpoint a firm in the interval $q_1 \in (0.5 - e^l, 0.5 + e^h)$ can converge to is higher than the highest fixpoint a firm in the interval $q_1 \in (0, 0.5 - e^l)$ can converge to. There is thus a one-to-one mapping between the value of the lowest fixpoint for firms in the interval $q_1 \in (0.5 - e^l, 0.5 + e^h)$ and their level of quality. Focus next on the larger fixpoint for each

q_1. It is also increasing in q_1. Let m_{larg} be the *lowest* value of this lower fixpoint, associated with quality equal to $0.5 - e^l$. We know that $m_{larg} > 0.5$. We also know that the *largest* fixpoint a firm in the interval $q_1 \in (0.5 - e^l, 0.5 + e^h)$ can converge to is lower than the lowest fixpoint a firm in the interval $q_1 \in (0.5 + e^l, 1)$ will converge to. There is thus a one-to-one mapping between the value of the largest fixpoints for firms in the interval $q_1 \in (0.5 - e^l, 0.5 + e^h)$ and their level of quality.

Overall then, each level of market share is associated with a unique level of quality that can converge to this level of market share (but a given level of quality may be associated with two different levels of market share, see Figure 4). For $m_1^* \in (0, m_{low})$, where $m_{low} < 0.5$, the level of quality associated with m_1^* is an increasing function of m_1^*. For $m_1^* \in (m_{larg}, 1)$, the level of quality associated with m_1^* is also an increasing function of m_1^*. However, the level of quality associated with $m_1^* = m_{low}$ is lower than the level of quality associated with $m_1^* = m_{larg}$ because $m_{larg} > 0.5$ is associated with a firm with quality below 0.5 while $m_{low} < 0.5$ is associated with a firm with quality above 0.5.

Theorem EC.2 **(Logit Choice Model)**. *Suppose the utility of selecting product* 1 *in period* $t+1$ *is* $u_{1,t+1} = aq_1 + bm_{1,t} + \varepsilon_{1,t+1}$ *and the utility of selecting product 2 is* $u_{2,t+1} = a0.5 + b(1 - m_{1,t}) + \varepsilon_{2,t+1}$*, where* $\varepsilon_{j,t+1}$ *are independent draws from an extreme value distribution with density* $f(\varepsilon) = Exp(-\varepsilon - e^{-\varepsilon})$*. The probability of selecting product 1 in period* $t+1$ *is then*

$$P_{1,t+1} = \frac{1}{1 + e^{a(q_1 - 0.5) + 2b(m_{1,t} - 0.5)}}. \tag{EC.4}$$

Let m_1^* *be the market share firm 1 converges to as* $t \to \infty$*. i) Whenever* $b < 2$*,* $E[q_1|m_1^*]$ *is an increasing function of* m_1^* *for all* $m_1^* \in [0,1]$*. ii) Whenever* $b > 2$*,* $E[q_1|m_1^*]$ *is an increasing function of* m_1^* *for all* $m_1^* \in (0, m_l)$*, where* $0 < m_l < 0.5$*, and* $E[q_1|m_i^*]$ *is an increasing function of* m_1^* *for all* $m_1^* \in (m_h, 1)$*, where* $0.5 < m_h < 1$*, but* $E[q_1|m_i^*]$ *"dips" in the sense that* $E[q_1|m_1^* = m_l] > E[q_1|m_1^* = m_h]$.

Proof For the logit model, the density of $d = \varepsilon_{1,t} - \varepsilon_{2,t}$, equals $g(d) = e^{-d}/(1 + e^{-d})^2$ (e.g., Maddala, 1983, p. 60). Thus, $g(0) = 0.25$ and $1/2g(0) = 2$. The conclusion follows by Theorem EC.1.

Theorem EC.3 **(Probit Choice Model)**. *Suppose the utility of selecting product* 1 *in period* $t+1$ *is* $u_{1,t+1} = aq_1 + bm_{1,t} + \varepsilon_{1,t+1}$ *and the utility of selecting product 2 is* $u_{2,t+1} = a0.5 + b(1 - m_{1,t}) + \varepsilon_{2,t+1}$*, where* $\varepsilon_{j,t+1}$ *are independent draws from a normal distribution with mean zero and variance one. Let* m_1^* *be the market share firm 1 converges to as* $t \to \infty$*. i) Whenever* $b < \sqrt{\pi}$*,* $E[q_1|m_1^*]$ *is an increasing function of* m_1^* *for all* $m_1^* \in [0,1]$*. ii) Whenever* $b > \sqrt{\pi}$*,* $E[q_1|m_1^*]$ *is an increasing function of* m_1^* *for all* $m_1^* \in (0, m_l)$*, where* $0 < m_l < 0.5$*, and* $E[q_1|m_i^*]$ *is an increasing function of* m_1^* *for all* $m_1^* \in (m_h, 1)$*, where* $0.5 < m_h < 1$*, but* $E[q_1|m_i^*]$ *"dips" in the sense that* $E[q_1|m_1^* = m_l] > E[q_1|m_1^* = m_h]$.

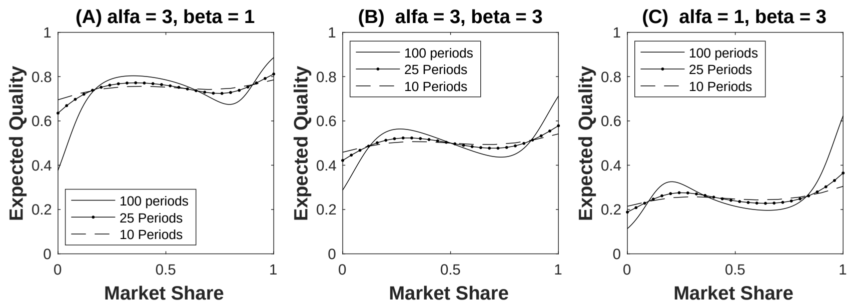

Figure EC.2 **Average quality associated with different market share levels when** $b = 2.5$ **and when quality was drawn from a beta distribution with parameters a)** $(\alpha = 3, \beta = 1)$ **(skewed towards high values) b)** $(\alpha = 3, \beta = 3)$ **(mean** 0.5 **and concentrated around 0.5) or c)** $(\alpha = 1, \beta = 3)$ **(skewed towards low values).**

Proof For the probit choice model, the density of $d = \varepsilon_{1,t} - \varepsilon_{2,t}$, is a normal density with mean zero and variance equal to two, i.e., $g(d) = \frac{1}{\sqrt{2\pi 2}} Exp[-\frac{d^2}{4}]$. Thus, $g(0) = \frac{1}{2\sqrt{\pi}}$ and $1/2g(0) = \sqrt{\pi}$. The conclusion follows by Theorem EC.1.

Appendix B: Robustness Analyses

Quality Distribution

The model in section 3 assumed that quality was uniformly distributed but the basic result—a dip occurs when b is large enough—holds for many alternative assumptions about the quality distribution. Suppose, for example, that quality was drawn from a beta distribution with parameters a) $(\alpha = 3, \beta = 1)$ (skewed towards high values) b) $(\alpha = 3, \beta = 3)$ (mean 0.5 and concentrated around 0.5) or c) $(\alpha = 1, \beta = 3)$ (skewed towards low values). Figure EC.2A-C shows that the dip occurs, when $b = 2.5$, at about the same level of market share as in Figure 1B. Moreover, Theorem EC.2 in Appendix A shows that a dip occurs, in the long-run (as $t \to \infty$), whenever $b > 2$ for any continuous quality distribution with support $(0, 1)$.

Many Products

The model in section 3 analyzed competition between two products, with one having a fixed level of quality. A dip also occurs in a model with many competing products, each with its quality drawn from a uniform distruibution. To illustrate this, suppose there are n products. The quality of product i is q_i, where q_i is drawn from a uniform distribution between zero and one. The utility of product i in period t is $u_{i,t} = aq_i + bm_{i,t-1}$, where $m_{i,t}$ is the market share of product i in period $t-1$ (we set $m_{i,0} = 1/n$). The probability that product i is selected in period t is the multinomial logistic: $P_{i,t} = e^{u_{i,t}} / (\sum_{j=1}^{n} e^{u_{j,t}})$. Such a model also generates a dip if b is large enough, as illustrated

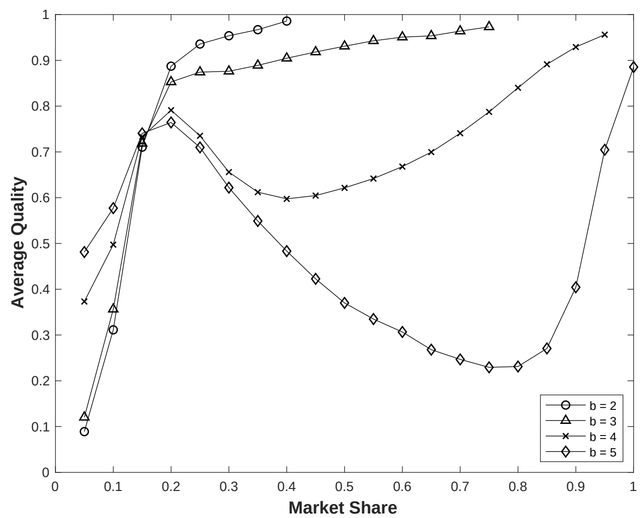

Figure EC.3 **How average quality varies with market share for different values of b (each line is based on 1 million simulations of 1000 periods in which $a = 1$ and $n = 10$).**

in Figure EC.3. Figure EC.3 shows how average quality varies with the market share obtained after 1000 periods for different values of b when $n = 10$ (ten actors) and $a = 1$. Higher market share indicates higher average quality only when $b = 2$ or $b = 3$. For $b = 4$ or $b = 5$, higher market share does not indicate higher average quality but the function relating market share and quality "dips".

Alternative Choice Models

The model in section 3 assumed that consumer choice followed the logit choice model but a dip also occurs for other assumptions about the choice process. The crucial aspect of the logit choice model, that gives rise to the dip, is the initial convexity and the eventual concavity of the logit choice function, i.e., $1/(1 + exp(-b(m - 0.5)))$ is convex for $m < 0.5$ and concave for $m > 0.5$. If b is large enough ($b > 2$), then the derivative at $m = 0.5$ is larger than 1, and these properties will guarantee that there will be three fixpoints of the function $1/(1 + exp(-b(m - 0.5))) - m$: one at a value $m < 0.5$, one at $m = 0.5$, and one at $m > 0.5$ and the first and the third of these fixpoints will be stable.

What other choice models share these essential features? As Lemma 1 in Appendix A shows, a large class of so called 'random utility models' do share these features and also generate a a dip if b is large enough (Theorem EC.1). The logit model can be derived as a random utility model in which individual i selects product 1 if $u_{i,1} + \varepsilon_{1,i} > u_{i,2} + \varepsilon_{2,i}$, where $\varepsilon_{j,i}$ are independently drawn from an extreme value distribution with density $f(\varepsilon) = Exp(-\varepsilon - e^{-\varepsilon})$. Lemma EC.1 in Appendix A shows that any random utility model in which individual i selects product 1 if $u_{i,1} + \varepsilon_{1,i} > u_{i,2} + \varepsilon_{2,i}$, where $\varepsilon_{j,i}$ are independently drawn from a *unimodal* distribution has the same essential features (convex for $m < 0.5$ and concave for $m > 0.5$) and also generates a dip (Theorem EC.1). The intuition is that if $\varepsilon_{j,i}$ are drawn from an unimodal distribution, values of $\varepsilon_{2,i} - \varepsilon_{1,i}$ close to zero are more likely than large or negative values. As a result, if $u_{i,1}$ is much larger than $u_{i,2}$, the probability that product 1 will be chosen is close to one and will not increase much if $u_{i,1}$ increases by a given amount.

This class of random utility models with unimodal error terms is quite broad and includes the logit but also the probit choice model, the two most commonly used choice models. Indeed, Theorem EC.3 shows that for the probit choice model, a dip occurs whenever $b > \sqrt{\pi}$. What does this class of random utility models exclude? It excludes choice models in which the error terms are drawn from a bimodal distribution. If the error terms were drawn from a bimodal distribution, with a density which reaches a minimum at zero, then the choice function would be concave for $m < 0.5$ and convex for $m > 0.5$, implying there is only one fixpoint and there is no dip.

A dip occurs in the short-run (for low values of t) even for choice models that do not imply a dip asymptotically (as $t \to \infty$). For example, a dip occurs in the short run for a linear probability model: $P_{1,t+1} = wq_1 + (1-w)m_{1,t}$. For example, a dip occurs after 100 periods whenever $0 < w < 0.243$. If $w = 0.1$, a dip occurs even after 20,000 periods. Why does a dip occur even when the choice model is linear? The reason is that if w is low the outcome in the first period has a long-term effect. For example, suppose $w = 0.2$ and consider a product with quality $q_1 = 0.5$. If this product is not chosen in the first period the choice probability in the second period decreases to 0.1. Unless the product is selected in the second period, the choice probability remains at this low level. On average it takes ten periods before this product is selected again. Even if the product is selected after ten periods, the choice proportion will remain low, resulting in a low choice probability. If the product had been selected in the first period, however, the choice probability would have increased to 0.9 and the product would have been very likely to be selected in the second period and the choice proportion would likely have remained high. As this illustration shows, the outcome in the first period can have a long-term impact even when the choice model is linear. The outcome in the first period does not have a permanent effect, however, if the choice model is linear. Rather, the choice proportion will eventually converge to q_1 as $t \to \infty$ regardless of the outcome in the first period and regardless of the value of w. The reason is that $z^* = q_1$ is the only solution for $z \in (0,1)$ to the equation $z = wq_i + (1-w)z$.

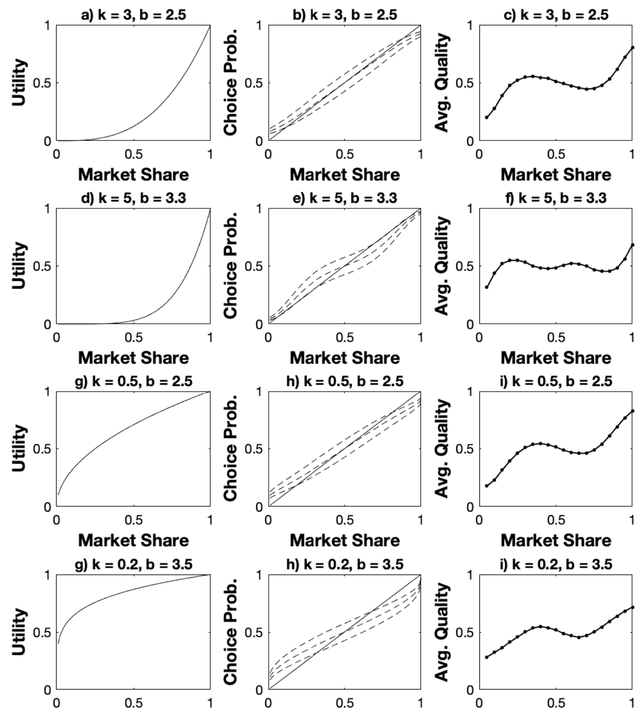

Figure EC.4 **How how i) $m_{j,t-1}^k$ varies with $m_{j,t-1}$ ii) how the choice probability varies with the market share for three different levels of quality (equal to 0.8, 0.5, and 0.2) and iii) how average quality, in period 100, varies with the obtained market share (based on 500,000 simulations) for different combinations of k and b (when $a = 1$).**

Concave or convex market share effects

The model in section 3 assumed that utility was a linear function of past market share. Consider a slightly more general model where $u_{j,t} = aq_j + bm_{j,t-1}^k$. If $k < 1$, utility is a concave function of market share and if $k > 1$ utility is a convex function of market share. Computations show that a dip occurs, when b is high enough in both cases, but if k differs significantly from one a dip only occurs if b is high. The intuition is that when k differs significantly from one, the utility of a product is not impacted by changes in market share at high (concave functions) or low (convex function) values of market share. Figure EC.4 illustrates the impact of convexity and concavity. Figure EC.4 shows how a) $m_{j,t-1}^k$ varies with $m_{j,t-1}$ b) how the choice probability varies with the market share for three different levels of quality (equal to 0.8, 0.5, and 0.2) and c) how average quality, in period 100, varies with the obtained market share (based on 500,000 simulations). When $k = 3$, we still get a dip when $b = 2.5$ even if $m_{j,t-1}^k$ is a convex function, because the logit function, $1/(1 + exp(-b(x)))$, is concave for $x > 0$. Even when $k = 5$ we get a dip when $b = 3.3$; in fact we get two dips. Consider, next, the case with a concave function. When $k = 0.5$ a dip occurs when $b = 2.5$, again because the logit function, $1/(1 + exp(-b(x)))$, is convex for $x < 0$. When $k = 0.2$ a dip occurs with $b = 3.5$.

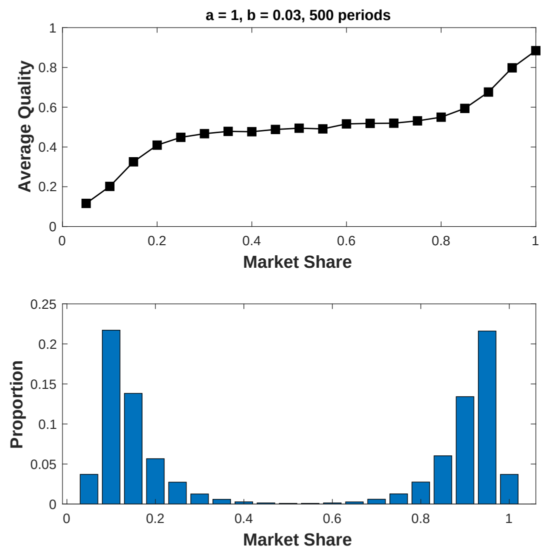

Figure EC.5 How a) average quality varies with market share in period 500 and b) a histogram of the obtained market share when $a = 1$ and $b = 0.03$.

Numbers versus proportion

Suppose $u_{j,t} = aq_j + bn_{j,t-1}$ where $n_{j,t-1}$ is the number of consumers that have selected product j during periods 1 to $t-1$. If b is large enough, this model also implies that product 1 either obtains a very high or very low market share but this model does not generate a dip. Rather, average quality is an increasing function of the obtained market share. To illustrate this, Figure EC.5 shows a) how average quality varies with the market share of product 1 in period 500 and b) a histogram of the resulting market share, when $a = 1$ and $b = 0.03$ (based on 1 million simulations). The market share of product 1 tends to be close to zero or one but there is no dip. The crucial difference between this model and the model in section 3 is that in this model quality does not matter in the long-run. When t is large, $n_{1,t-1} - n_{2,t-1}$ will become a very large positive or negative number and the term $b(n_{1,t-1} - n_{2,t-1})$ will dominate over $a(q_1 - 0.5)$ in the choice probability. Thus, in the long-run, the choice probability will become one or zero, regardless of the quality of product 1.

To understand the dynamics of this model, it is useful to consider the extreme case when only quality matters when $|n_{1,t-1} - n_{2,t-1}| < c$ (i.e., suppose that $b = 0$ and $a = 1$ whenever $|n_{1,t-1} - n_{2,t-1}| < c$) and quality does not matter whenever $|n_{1,t-1} - n_{2,t-1}| > c$ (consumers choose the leading product with probability one whenever $|n_{1,t-1} - n_{2,t-1}| > c$). For this extreme version of the model, quality only matters initially and only determines whether product 1 or product 2 wins in the long-run and the time until one of the barriers is reached ($n_{1,t-1} - n_{2,t-1} > c$ or $n_{1,t-1} - n_{2,t-1} < c$).

The higher the value of q_1, the higher the probability that product 1 wins in the long-run (i.e., the barrier $n_{1,t-1} - n_{2,t-1} > c$ is crossed) and the more quickly this will happen. In the long-run, market share converges to one or zero, and the probability that a product 1 with quality q_1 reaches a market share equal to one is the probability that the barrier $n_{1,t-1} - n_{2,t-1} > c$ is crossed before $n_{1,t-1} - n_{2,t-1} < c$. This probability is an increasing function of q_1. Hence, expected quality will be higher for products that converge to a market share equal to one than for products that converge to a market equal to zero.

Appendix C: Statistical Tests for a Dip

Is the change in sign detected by the cubic spline statistically significant? This question can be answered by fitting a cubic spline, calculate confidence intervals for the predicted review score, and test whether there are box office values for which the effect of box office is significantly negative. As Uri Simonsohn has recently demonstrated, however, such a test has low statistical power and can both fail to detect a change in sign and falsely detect it when it is not present (Simonsohn 2018). The problem is that such a test is sensitive to departures from the assumed functional form (i.e., quadratic or cubic). Simonsohn shows that an algorithm with superior statistical power, for detecting whether a function is u-shaped, is an algorithm that fits separate lines: one for high values of x and one for low values. We adopted this method to test whether a dip occurs by fitting two lines around each "kink" in the data and testing whether the signs around the kinks differ.

In brief, the modified algorithm works as follows (see Simonsohn 2018, for technical details). First, we fit a cubic spline regression to the data. Based on the cubic spline regression, we identify any local maxima and minima, i.e., x-values where the predicted y-value is a local maximum or a local minimum. For example, in the action/adventure genres the cubic spline regression identifies a local maximum at a box office of about 50,000 dollars and a local minimum at about 3 million dollars. Second, we apply Simonsohn's algorithm to test whether the change in sign around a local maximum or minimum is significant. When we use this algorithm, we only rely on data around the local maximum or minimum. That is, to test whether the change in sign around the local minimum at 3 million is significant, we use data on a) all movies in the action/adventure genres with box office above 3 million and b) all movies with box office below 3 million and above 50,000 dollars, i.e., above the local maximum. Similarly, to test whether the change in sign around the local maximum at around 50,000 is significant, we use data on a) all movies in the action/adventures genres with box office below 50,000 and b) all movies with box office above 50,000 and below 3 million, i.e., below the local maximum. For the movies from the drama genre, where the cubic spline shows a monotonic pattern with only one change in slope at around 35 million, we use data on all movies above or below this change in slope.

The purpose of these first two steps, which is our modification of Simonsohn's method, is to create two sets of data, each of which is then analyzed by Simonsohn's original method. Third, we fit a linear regression with separate slopes above and below the cutoff used (i.e., the local minimum or maximum). Fourth, to maximize the power of the test to detect a change in sign, we change the cutoff, from the local maximum or minimum, to a new value, called a "Robinhood Cutoff". The purpose of changing the cutoff is to allocate more observations to the side (above or below the original cutoff) where the regression slope was less significant. For example, in this fourth step, the cutoff changes from the local minimum at around 3 million to a value of about 5 million (see Figure 7A, Lower Panel), to allow for more observations in the middle range of box office.

Appendix D: Simulations and Numerical Integration

Figures 1 to 3 are not based on simulations but on numerical integration of the equations specifying the expected level of quality of firm 1 given the market share obtained after t periods. We first explain the basis for the numerical integrations and then show how to replicate the results using a simple simulation program.

Suppose firm 1 has market share $m_{1,t}$ after t periods. We wish to calculate $E[q_1|m_{1,t}]$. A given level of market share can always be translated in the number of consumers out of t that choose the product of firm 1, n_t. We thus seek the expected level of q_1 given that firm 1 is chosen k times in t periods. This can be computed as

$$E[q_1|n_t = k] = \int_{q_1=0}^{q_1=1} q_1 f(q_1|n_t = k)dq_1 \tag{EC.5}$$

where $f(q_1|n_t = k)$ is the conditional density of the quality distribution given that firm 1 chosen k times in t periods. Bayes rule implies that

$$f(q_1|n_t = k) = \frac{Pr(n_t = k|q_1)f(q_1)}{Pr(n_t = k)} = \frac{Pr(n_t = k|q_1)f(q_1)}{\int_{q_1=0}^{q_1=1} Pr(n_t = k|q_1)f(q_1)dq_1}, \tag{EC.6}$$

where $Pr(n_t = k|q_1)$ is the probability that a firm 1 with quality q_1 is chosen k times in t periods and $f(q_1)$ is the density of the quality distribution. Using Bayes rule $E[q_1|n_t = k]$ can thus be written as

$$E[q_1|n_t = k] = \frac{\int_{q_1=0}^{q_1=1} q_1 Pr(n_t = k|q_1)f(q_1)dq_1}{\int_{q_1=0}^{q_1=1} Pr(n_t = k|q_1)f(q_1)dq_1}. \tag{EC.7}$$

To compute $E[q_1|n_t = k]$, we need to compute $Pr(n_t = k|q_1)$, for any value of q_1. Computing $Pr(n_t = k|q_1)$ can be done recursively, starting from period one. In the first period, $Pr(n_1 = 1|q_1) = 1/(1+e^{-a(q_1-0.5)})$. The probability that firm 1 is chosen in the second period depends on the market share after period one: $P_{1,2} = 1/(1 + e^{-a(q_1-0.5)-2b(m_{1,1}-0.5)})$. There are only two possible values of $m_{1,1}$, however: firm 1 was chosen or not in the first period. To compute $Pr(n_2 = k|q_1)$ we a) compute the choice probability for each of these two possible outcomes in the first period and b) weigh these two possibilities by their probabilities, i.e., weigh them by $Pr(n_1 = 1|q_1)$. Proceeding in the same manner, $Pr(n_t = k|q_1)$ can be computed for any period t. When $Pr(n_t = k|q_1)$ have been computed, we then compute $E[q_1|n_t = k]$ by numerical integration of the numerator and denominator of equation EC.7.

Alternatively, a simple simulation program can compute $E[q_1|m_{1,t}]$ (and modification of this program can be used to compute Figures 2 and 3). The Matlab program below simulates the competition between firm 1, with quality q_1, and firm 2, with quality 0.5, each for t periods. This simulation is repeated ns times. The result is an array with ns market shares ("ms") and another array with ns quality levels ("qual"). Using these two arrays, one can plot the average quality level within different market share intervals, as an approximation to $E[q_1|m_{1,t}]$.

```
ns = 50000; % number of simulations
t = 100; % number of periods
a = 1; % parameter in choice rule
b = 2.5; % parameter in choice rule

ms = zeros(ns,1); % market share
rr = rand(t,ns); % uniform random numbers

%-----draw qualities from uniform distribution
qual = rand(ns,1);

%-------the main simulation loop for each pair of firms
for i = 1:ns
    n = 0; % number of times firm 1 get chosen
    q = qual(i,1); %  the quality of firm 1
    m = 0.5; % initial market share set to 0.5

for j = 1:t  % simulation loop for each period
    pr = 1/(1+exp(-a*(q-0.5)-2*b*(m-0.5))); % equation 1
    if(rr(j,i) < pr)
        n = n+1;  % if gets chosen, then add 1
    end
    m = n/j; % market share = proportion of periods firm 1 was chosen
```

```
end % end of the periods

ms(i,1) = n/t; % save the market share this firm 1 obtained
end % end of the simulations
```

References

Hill BM, Lane D, Sudderth W, et al. (1980) A strong law for some generalized urn processes. *The Annals of Probability* 8(2): 214-226.

Maddala, GS (1986). *Limited-dependent and Qualitative Variables in Econometrics* (Cambridge, England: Cambridge University Press).

Purkayastha, S. (1998). Simple proofs of two results on convolutions of unimodal distributions. *Statistics & Probability Letter* 39: 97-100.

Simonsohn U (2018) Two lines: A valid alternative to the invalid testing of u-shaped relationships with quadratic regressions. *Advances in Methods and Practices in Psychological Science* 1(4): 538-555.