Received: 13 November 2014 | Revised: 2 May 2016 | Accepted: 29 July 2016 | Published on: 25 March 2018

DOI: 10.1002/smj.2769

WILEY STRATEGIC MANAGEMENT JOURNAL

SPECIAL ISSUE ARTICLE

Transaction surplus superiority in canonical market segments: Using the profit map to guide positioning and investment choices across price-rivalry regimes

Steven Postrel

Paul Merage School of Business, University of California at Irvine, Irvine, California

Correspondence
Steven Postrel, Paul Merage School of Business, University of California at Irvine, Irvine, CA 92697.
Email: spostrel@gmail.com

Research Summary: Strategy needs an "intermediate measure" of competitive strength easily linked to management choices. "Competitive advantage" fails to accomplish this because most formulations of it violate important criteria and because these formulations are so diverse. I show that one intermediate measure, *transaction surplus superiority* (TSS), potentially solves both problems. Critically, TSS is defined independently of prices and its usefulness is independent of the degree of price competition: In "canonical" market segments, firms with TSS always make superior segment profits (gross of sunk investments) compared to rivals. When combined with a restriction to "incentive-compatible pricing behavior," TSS limits or solves the theoretical problem of indeterminate rivalry among oligopolists. Finally, TSS allows precise explanation of when sunk investments to improve efficiency are stimulated or suppressed by the "spur of competition."

Managerial Summary: Executives need a conceptual "instrument panel"—"intermediate measures" of competitive strength—to help choose their products and processes. Profit itself is too causally remote from these choices to serve. "Competitive advantage" has been defined in ways that either just mean superior profits or that depend on strong assumptions about price competition that frequently don't apply. Transaction Surplus Superiority (TSS) solves these problems. In an important special case, a "profit map" completely describes each firm's profitability for any combination of price competition and TSS. For choices where sunk cost differences typically aren't important (e.g., product color or inventory policy), TSS is a universal guide to action. TSS helps decide whether to increase price competition to gain market share, and it

largely determines whether a given investment in efficiency makes sense.

KEYWORDS

competitive advantage, modeling, oligopoly theory, rivalry, value-added

1 | INTRODUCTION

A primary purpose of theory in business strategy is to help us reason forward from managerial choices to the impact of these choices on payoffs. Our field is a classic example of a "science of the artificial" (Simon, 1996) in which analytical theory serves as an "engineering science" to support the activity of design—in this case, the design of firm strategies. But strategic contests take place in a wide range of product markets, from aromatherapy to zippers. Available managerial actions, as well as the linkage between these actions and payoffs, are context-specific. Even seemingly similar types of moves depend on situational specifics. Ford's decision to switch its flagship product, the F-150 pickup truck, from steel to aluminum (Taylor, 2014) occurred in a different context from Kraft's decision to switch its flagship macaroni-and-cheese product from "artificial" to "natural" ingredients (Best, 2016). The only way to provide useful insight across such a diverse set of environments is to abstract away from the details, identifying common patterns that might help explain and predict performance.

Our task in strategy is thus to find mappings from (a) the diverse contextual details of a given business setting to (b) a context-independent set of "intermediate" positional measures of competitive strength *that in turn map to (c) performance.* The idea is that if a firm can score well on this intermediate measure in its product market, by whatever context-specific means, then it is more likely to achieve superior performance for reasons that can be generalized beyond that particular situation. It has been common in strategy to call this sort of measure "competitive advantage," but that use of the term is by no means universal nor has it been consistent from one author to another, nor have those usages generally captured all the important aspects an intermediate measure of competitive strength ought to possess. An intermediate measure serves as a "position evaluation function" somewhat like those used in traditional computer chess programs—it takes more-or-less observable features of the current situation and computes from them an index of relative competitive strength that is in turn causally correlated to relative payoffs. It does *not* serve as an ultimate causal explanation for success, as, for example, a chess player's mental resources and training regimen might do. Nor does it merely restate the fact of success, as would a player's Elo rating or tournament record—the intermediate measure is part of the instrument panel, not the scorecard (Jensen, 2001, p. 19).

Based on this understanding of the task, I introduce a set of criteria for a useful, coherent intermediate measure to guide positioning and investment decisions. I then propose a concept, *transaction surplus superiority* (TSS) that meets these criteria.

One criterion, *independence of prices*, deserves special attention. It has long been known that the degree and type of price competition varies across industry contexts; the last 40 years of economic theory has also provided a number of plausible models in which, even in a single context, a

wide range of pricing behaviors are possible in Nash equilibrium. Furthermore, competitive success over rivals can be observed both for firms that charge relatively high prices (e.g., Apple and Tiffany & Co.) and others that charge relatively low ones (e.g., Huawei and Wal-Mart). It follows that a good intermediate measure must not only avoid including prices directly but also have *its relevance to profitability be robust to a wide range of assumptions about the nature of price competition.* Current formal approaches to assessing competitive strength (usually called "competitive advantage")—for example, the "value-based approach" or the resource-based view—fail to do this, generally assuming that all-out price competition is inevitable in the absence of capacity constraints.

The first contribution of this study is to show that TSS provides a sufficient intermediate target for managers' decisions under a wide range of assumptions about pricing rivalry. In "canonical" market segments that permit product and cost differentiation and have constant returns to scale (gross of sunk investments), surplus analysis allows the entire set of profit possibilities for every firm to be summarized using a new tool, the "profit map." This map displays each firm's payoff and market share for any level of prevailing price competition, and shows that firms with TSS over rivals always earn more. Therefore, theories that posit performance differences as being caused by differences in economic efficiency, whether based on superiority of resources, capabilities, knowledge, or anything else, *must* have their effects mediated through TSS.

The second contribution of this study is to attack a fundamental and long-standing problem in industrial organization economics and business strategy: the indeterminacy of the level of price competition in oligopolies (Fisher, 1989, pp. 114–117). Applying TSS to the *incentives* of leading firms and adopting the assumption that any firm could unilaterally start a price war and thereby drag the market into Bertrand competition, it turns out that the most efficient firm (with TSS over all rivals) has veto power over the constellation of offers made to buyers. As firm heterogeneity (in terms of the leader's TSS) increases, there eventually comes a point where segment competition must turn Bertrand with the leader choosing to take over the entire segment. Moreover, the following counterintuitive proposition holds: Actions by a segment leader (or any other force) that *raise* the costs of trailing firms can benefit customers and society as a whole; "raising rivals' costs" can no longer be treated as an evil per se and must be evaluated on a case-by-case basis.

The third contribution made here is to characterize the incentive for a firm to make sunk investments in surplus-creating assets. This characterization allows, among other things, an assessment of the extent to which investment is stimulated or suppressed by "the spur of competition." A somewhat unexpected finding under all pricing regimes is that competition (either more firms or lower starting price-cost margins) uniformly *reduces* the incentive to invest *when the investing firm cannot anticipate attaining TSS by its action.* On the other hand, if (a) the focal firm can attain a large enough TSS through its investment and (b) incentive-compatible pricing applies, then more initial competition actually encourages the firm to invest—the "spur of competition" operates.[1] Once again, counterintuitively, weakness of non-leading firms may be conducive to both overall and

[1]This finding is loosely analogous to results in the theory of patent races, where a "replacement" effect gives potential entrants or members of a competitive industry a bigger incentive to invest in new technology than incumbents while an "efficiency" effect goes in the opposite direction. In those models, the former effect dominates for drastic innovations that would wipe out the old technology but the latter dominates for incremental innovations that would allow the old and new technologies to both survive (Arrow, 1962; Reinganum, 1984; Tirole, 1988, pp. 391–396). Unlike the models considered in this paper, the patent-race literature's replacement effect depressed the incentives of the initial leader to invest and increased the incentive for initially non-participating entrants. In this paper, with incentive-compatible pricing incumbents tend to be more stimulated by competition to invest than entrants, especially with strong initial surplus positions. There is also an echo of the finding in Lippman and Rumelt (1982) that higher efficiency of incumbent firms raises the surplus-gain threshold needed to make an entrant's sunk investment profitable.

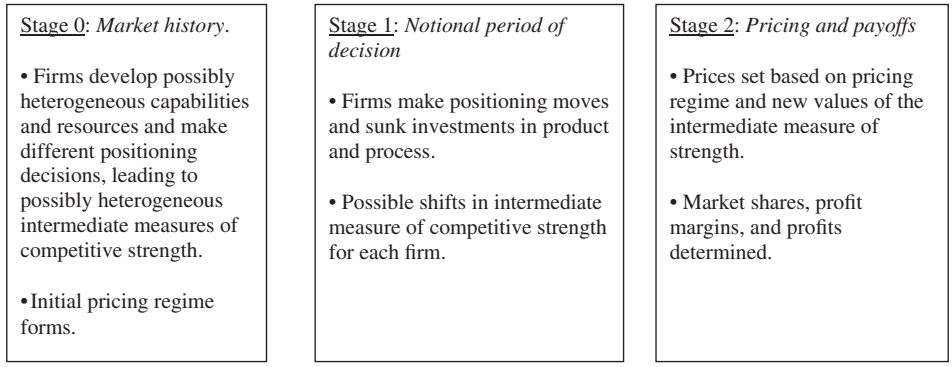

FIGURE 1 Timing and ordering of events

customer economic welfare if it stimulates investment and leads to more aggressive pricing by the focal post-investment leader.

Formally, one can think of this positioning framework in terms of a modified class of two-stage games in which firms make product, marketing, process, and capacity decisions in the first stage, determining TSS, and engage in price competition in the second stage.[2] Figure 1 lays out the sequence of events. The modification is that prior to the first stage the market is already considered to be in some specific pricing equilibrium that the first-stage choice might or might not disturb, so the analysis takes places in the midst of "market history" rather than starting from the beginning of time. In addition, pre-existing asymmetries among firms in resources, capabilities, or reputations are taken as given at the positioning (first) stage, as they are inherited from the past. The TSS concept requires no assumptions about the *reasons* why firms might be heterogeneous (e.g., resource superiority, better positioning choices, etc.) entering Stage 1, focusing entirely on the *consequences.* Nor does it assume any particular durability of this heterogeneity, as it separates the definition of superiority from the sustainability of that superiority so that the two can vary independently.

2 | "COMPETITIVE ADVANTAGE" AND ITS DISCONTENTS

At this point, a reader may be wondering why "competitive advantage" has barely been mentioned. The reason is the current state of the literature. In some strategic discourse, "competitive advantage" plays the role of an intermediate measure as described above (part of the firm's instrument panel). In other writings, the term has been used to describe various sorts of superior performance (part of the firm's scorecard). These are distinct conceptual categories. As the next section will explain, for purposes of informing positioning strategy via an intermediate measure of competitive strength, the scorecard definition must be set aside. Scholars attached to this scorecard definition of "competitive advantage" will therefore likely prefer to adopt "transaction surplus superiority" for the concept advanced in this article. (It remains unclear why "superior performance" is not an adequate term for describing superior performance, but perhaps its unadorned directness lacks the requisite touch of mystery.)[3]

[2]The two-stage game approach has a substantial heritage in industrial organization economics and strategy, as described by Brandenburger and Stuart (2007, p. 540). Recent examples include Asmussen (2015) and Costa, Cool, and Dierickx (2013).

[3]Siggelkow and Wibbens (2015) try to unpack performance so as to isolate the impact of firm resources and introduce a "ladder of competitive advantage" containing three rungs. *Transaction advantage* is closely related to transaction surplus superiority as defined in this paper; *gross value appropriation,* intended to be a measure of returns to a firm's "idiosyncratic resources;" and *lifetime*

Even within the instrument panel category of usage, different authors specify different subtle shades of meaning. Thus, one can never be sure what a finding about "competitive advantage" means unless one reads the "fine print" of a particular work's definition. This print may be fine indeed, sometimes requiring the reader to parse a particular formal model. While this situation is a "mere" problem of terminological conventions, it is not a benign state of affairs—when people use the same term to mean many different things and/or call the same thing by many different terms, it is hard either to have a conversation or to efficiently present one's findings. As Oxley, Rivkin, and Ryall (2010, p. 379) have suggested, one criterion for a piece of theory in strategy to be high quality should be that "The theoretical claims [of the work] are unambiguous: interpretation of its terms, premises and conclusions does not vary from scholar to scholar." The strategy field is clearly failing to meet that test with respect to "competitive advantage." At least one scholar (Lieberman, 2010) has called for abandonment of the term in research settings for precisely this reason.

Note that no individual scholar can unilaterally satisfy this criterion of common interpretation. Social consensus on a terminological convention is necessary for its adoption, supported by writers, readers, editors, and reviewers. This article advances a specific concept for an intermediate measure of competitive strength, justifies it based on an explicit set of desiderata, and shows how the concept can be applied to reproduce known results as well as generate new insights about profitability, price competition, and sunk investments. For scholars who think about "competitive advantage" in instrument-panel terms, this "transaction surplus superiority" definition therefore ought to fill the bill and identifying the two should be unproblematic. Were a consensus to form that "competitive advantage is TSS" there could be little objection, but the substantive issues addressed in this article should not be left obscure pending such an outcome. Nor should a reader's understanding of these issues be blocked by hang-ups about surrendering a cherished usage of "competitive advantage." While connotations and historical usage ought to have some bearing on terminology, so long as a term is consistently used throughout a field and the concept to which it refers is sufficiently clear and precise, wrangling over which exact word to use is not very productive (Popper, 1966, pp. 13–20).

Therefore, the rest of this article will stick to calling the concept "transaction surplus superiority" (or TSS) and describe it as a candidate "intermediate measure of competitive strength." After the TSS concept has been explained and its usefulness shown in the analysis of canonical market segments, I will discuss its relationship to extant definitions of "competitive advantage."

3 | CRITERIA FOR A USEFUL INTERMEDIATE MEASURE

A useful intermediate measure of competitive strength should help us understand competitive strategic problems: It must be a (a) *medium-term,* (b) *product-market*, and (c) *forward-looking* objective for managers to pursue. "Medium-term" positioning decisions normally comprise choices about target markets, product attributes, process attributes, and capacity levels. Meaningful superiority cannot be created by easily reversed tactical moves (Ghemawat, 1991), and these short-term tactics (such as

investor value appropriation, which captures the discounted value of all cash flows accruing to investors from firm founding until bankruptcy or acquisition. The latter two are "scorecard" items not suitable for an intermediate measure of competitive strength in positioning analysis. Encouragingly, each concept is precisely defined and distinguished. Note that the authors chose to call none of these measures "competitive advantage," deferring to the wide range of usage found in the literature for that term. There have been previous attempts to isolate parts of firm profitability and attribute them to management acumen, such as Economic Value Added (EVA) or "residual income"; these have been used mostly as scorecards and as part of compensation plans, with some theoretical justification (Rogerson, 1997; Shaked, Michel, & Leroy, 1997).

prices) will have to be chosen taking the level of the intermediate measure as given. On the other hand, there are long-term, hard-to-adjust aspects of the firm (such as culture, reputation, or deep-seated capabilities) that often must be taken as fixed at the time positioning decisions are made; these long-term factors will affect the relationship between medium-term choices and the realized value of the intermediate measure.

"Product market" means that the intermediate measure captures a firm's relative strength against rivals in a given customer domain. More subtly, it means that the unit of analysis (the "firm") includes whatever combination of stakeholders and assets is, by convention or for the purpose of the analysis, to be treated as a purely cooperative team. An intermediate measure of competitive strength should not be confused with other indices that might determine the distribution of rents among the members of a single team.

"Forward-looking" implies that the purpose of the intermediate measure is to help managers think about the prospective payoffs of different choices, rather than impute past results to specific resources or decisions. While an understanding of history may matter for future-oriented assessments, the *definition* of the intermediate measure ought not to depend on these imputations. There must be some practical way (at least within a simple model) to link firms' positioning choices today to their measured competitive strength tomorrow.

Bearing these needs in mind, I propose the following (non-exhaustive) list of criteria for a useful intermediate measure. While this is not a formal axiomatic system, it follows the spirit of that approach by suggesting criteria with high face validity and conformance with the discussion above.

3.1 | Not a trivial tautology (NT)

The intermediate measure is supposed to help us understand *why* some strategies are more successful than others. It therefore should be an antecedent to business success rather than a trivial tautological re-description of it—a dial on the instrument panel rather than a scorecard. A sophisticated decomposition of accounting and financial data might identify differences in firms' cost or revenue structure and suggest causal hypotheses that could inform strategy making (Siggelkow & Wibbens, 2015). But such a decomposition, without linking to context-specific operational and marketing facts, conveys no causal information about why one strategic choice might be preferred over another. Moreover, it should be logically possible to overpay for superiority according to the intermediate measure, a possibility central to analyses of strategic factor markets (e.g., Chatain, 2014 or Asmussen, 2015) where one wants to know if pursuit of it in a specific case is worthwhile. Tautological definitions of performance do not permit that distinction.

3.2 | Complete ordering (CO)

It should be possible to make pairwise comparisons between direct rivals and say which one scores better on the intermediate measure of competitive strength or if neither has superiority over the other. Moreover, Firm A can have superiority over firm B and simultaneously be inferior to firm C. Without this feature, we would not be able to compare strategies across all firms. While this criterion sounds benign, not all definitions of "competitive advantage" meet it. For example, Peteraf and Barney (2003) compare each firm only to a hypothetical breakeven competitor, while MacDonald and Ryall (2004) compare each player only to zero payoff, forming only a partial order in each case.

3.3 | Managerial control (MC)

An intermediate measure that compares the competitive strength of rivals ought to show how managerial choices, playing out through product markets, affect that comparison. Concepts that focus on immutable conditions (e.g., past sunk investments in resources) may sometimes have descriptive power but nevertheless miss the point of the exercise. These conditions ought to inform positioning strategy insofar as they affect a firm's relative likelihood of success with different product and process choices, but they are not useful intermediate targets because they are not amenable to being changed by those choices. This criterion debars the resource-based view's trinity of resource value, rarity, and inimitability from consideration as an intermediate measure. It also means that empirical studies of the intermediate measure will have to cope with the context-specific nature of causality, digging into operational and marketing detail along the lines of a case study rather than using cross-context, large-sample methods.

3.4 | Operational guidance and generality (OG)

The intermediate measure ought to offer a target for managerial action, something more operational than "superior profits" but not so specific that it can't be applied across the whole range of product and service businesses where strategic thinking is desirable; and, of course, the measure must be generally conducive to superior profits (gross of sunk investments) so as to be a goal worth pursuing. Much of the rest of this article is devoted to demonstrating that transaction surplus superiority passes this test.

3.5 | Independence of sustainability (IS)

One key to meeting these constraints is to avoid conflating the relative competitive strength of firms with the persistence of that relationship. Sometimes these different concepts get blurred together, confusing the issues.[4] The intermediate measure should embrace the possibility that the ordering among rivals may be ephemeral. A difference between firms on the intermediate measure might be part of a strong equilibrium after all rational attempts at imitation have been made or it might be a fleeting disequilibrium event, depending on the context. An intermediate measure of strength is especially vital in turbulent settings where firms must adapt their positions rapidly.

3.6 | Independence of prices (IP)

The intermediate measure must not be defined in terms of the prices firms charge nor should its impact on profit depend on specific assumptions about industry pricing behavior. There are four main reasons for this:

1. Prices are tactical variables under the direct control of management, and may be set wishfully—either very high, hoping for customers to purchase anyway, or very low, hoping to cover one's costs. The measure of competitive strength ought not to be defined in terms of these hopes or

[4]To the extent that "competitive advantage" has been used to describe the intermediate measure in the literature, Peteraf and Barney (2003) goes a long way in ameliorating this problem. But note McGrath (2013), whose title suggests an "end to competitive advantage" even though the substance of its argument is entirely about a purported decline in the sustainability of advantages. The impact of greater transience of competitive advantages, however, would actually reverse the apparent implication of the book's title, making analysis of competitive advantage a more frequently urgent task than previously.

wishes; if it were, it would be vulnerable to manipulation since any firm can choose high or low prices whenever it wants to.

2. If competitive strength is about prices, is it high prices or low prices that convey it? One can easily see that winning firms may be in either position, suggesting that prices are not fundamental to advantage. *What is fundamental is the ability to prosper at whatever price turns out to be optimal for the firm.*
3. Modern industrial organization has not developed a tight predictive theory about the level of price competition that can be expected in a given product market. To the contrary, advances in game-theoretic models have led to a plethora of distinct Nash equilibrium outcomes that apply in circumstances distinguishable only by unobservable boundary conditions. The repeated-game (or supergame) literature, for example, produces the Folk Theorem, whereby any oligopoly outcome from collusion to Bertrand can be sustained if the discount rate is low enough (Tirole, 1988). Quick-response models, where rivals rapidly match price undercutting, thus deterring the tactic, easily lead to near-collusive outcomes (Anderson, 1984; Radner, 2003; Bhaskar, 2007). This central consideration drives much of the analysis presented below, where the level of price competition is directly parameterized and results tested for robustness against variation of that parameter. As mentioned in the introduction, failure to achieve this independence greatly limits alternative approaches to "competitive advantage."
4. Strategy scholars are often interested in studying the degree of forbearance rivals show toward one another. A definition of "competitive advantage" or any other intermediate measure that presupposes one level of rivalry—typically, in the literature, all-out price competition—makes it useless in such circumstances.

This criterion is only relevant in situations where prices are posted in some committed way to buyers. If instead each transaction is subject to unrestricted individual bargaining, as is assumed in much of the cooperative game theory literature, then there is no set price to be independent of and useful intermediate measures would of necessity ignore IP.[5] Moreover, definitions of "competitive advantage" such as Peteraf and Barney (2003), which implicitly require perfect competition in the product market to identify a "breakeven" competitor, fail to meet this criterion.

3.7 | Transaction focus (TF)

The unit of analysis for the intermediate measure must be chosen to avoid the aggregation of dissimilar cases. A firm may be superior to a rival with respect to some potential transactions but be inferior with respect to others. Averaging over both sets and concluding that one firm has superiority in the aggregate is misleading analytically and managerially. We really do not want to be in the position of saying without qualification that Ford has superiority over Porsche or vice versa. Rather, we want to segment potential auto purchase transactions and say which firm is superior within which segment.

The implications of the Transaction Focus criterion may be more far-reaching than is immediately obvious. Any firm facing a downward sloping demand curve for its product is likely to find that its relative competitive strength, as tallied by the intermediate measure, varies all along the curve. Such a firm will therefore not have a single value of the intermediate measure but a whole set

[5]It should be borne in mind, however, that the assumption of unrestricted bargaining is *not* weaker and more general than a direct assumption about the degree of price rivalry in a non-cooperative model; in fact, in situations like the ones studied in this article, the unrestricted bargaining assumption is tantamount to assuming all-out Bertrand price competition.

TABLE 1 Notation

S_i	Total surplus creation of firm i with buyer $= V_i - C_i$
V_i	Buyer's willingness to pay for i's product if no rivals present
C_i	Firm i's incremental cost of executing transaction with buyer
P_i	Firm i's price
b_i	Firm i's offer to the buyer $= V_i - P_i$
$\underline{b}$	Prevailing offer in a market segment
N	Potential number of firms in a segment
M	Number of firms with positive market share in a segment
K	Sunk cost of a surplus-increasing investment
s	Surplus increase from an investment
y	S_2/S_1 where S_1 is the leader's surplus creation and S_2 is the second-best surplus
π_i	Gross profit of firm i

of values, each corresponding to a distinct market segment. An average value of the measure across segments can be constructed, but it will not necessarily be true that superiority in this average value translates into superior performance.[6]

Some may find this criterion a hard pill to swallow, both theoretically and empirically. It rules out, when aggregating over heterogeneous customers, employing the intermediate measure (often, "competitive advantage") to generate simple assessments of relative competitive strength across the whole market. But this limitation is just facing up to reality; simple assessments would be misleading. For example, one model in Makadok (2010) has two firms at the ends of a line segment of customers, with a "transportation cost" of selling that is linear in the distance from customer to firm. Each firm has its own constant unit cost function, and the difference in unit costs is labeled the "competitive advantage" of the lower-cost firm, *excluding* the transportation cost to any given customer from this putative intermediate measure. (The transportation cost is instead interpreted as a separate parameter measuring "rivalry restraint," although the analysis assumes one-shot Nash equilibrium behavior.[7])

That definition allows a blanket claim of superior competitive strength to be made, but by losing transaction focus it leaves a critical explanatory gap: When the higher-cost firm is not driven from the market completely, it profitably sells to a set of customers near its end of the line segment. For those customers, the sum of production and transportation costs is lower when purchasing from the higher-unit-cost firm than when purchasing from its rival. The transaction focus criterion states that for these customers, an intermediate measure ought to show that the higher-unit-cost, nearer firm has superior competitive strength. In that way, the intermediate measure would be able to explain why the higher-unit-cost firm was able to profitably sell to that set of nearby customers; the transaction-focus-violating definition cannot. A change in transportation cost is a change in

[6]A "strong-second" firm with a high average strength across segments might have high payoffs if price competition were soft, but with perfect price discrimination and a winner-take-all situation in each segment it would be knocked out everywhere by its higher-variance rivals.

[7]The situation reveals an ambiguity or duality in how "rivalry" is understood. On the one hand, we have the basic idea of the level of forbearance (or lack of it) given all cost and demand curves. That is the sense used in this article. On the other hand, there is also a tradition equating the level of "rivalry" with the degree of cross-price demand elasticity between rivals. The transportation cost in the linear market segment is inverse to this elasticity. I find this latter use of the term troubling, as the very same "rivalry" effects can be found for a monopolist facing no rivals at all when its demand elasticity changes (as they would for a monopolist placed at one end of the Makadok, 2010 linear segment). Since these two types of "rivalry" can be varied independently by a modeler, and likely by reality, strategy scholars could avoid future confusion by distinguishing them explicitly.

economic efficiencies, not a change in "rivalry" in the sense of forbearance; it would affect a monopolist with no rivals in the same way it would a duopolist.

This article focuses on understanding "canonical" market segments in which firms are heterogeneous but buyers are not. Aside from being a natural, simple case that must be thoroughly understood before more-complex settings can be mastered, this setup allows us to aggregate up to firm market share and profits without violating transaction focus. The analysis can be conducted entirely in terms of a "representative" customer. The natural extension to more-realistic market settings would be to consider a discrete number of distinct canonical segments or a continuum of customers laid out on a line segment, circle, or plane. In such multi-segment settings, the TF criterion highlights that *ease of price discrimination* is one of the most critical contingent factors for analysis, even though it is rarely given prominence in strategy research. With perfect price discrimination, each canonical market segment can be analyzed independently, thus allowing the methods of this article to be carried over without amendment.

4 | TRANSACTION SURPLUS SUPERIORITY AS AN INTERMEDIATE MEASURE OF COMPETITIVE STRENGTH

One constant feature of any transaction that takes place in a market economy is that both sides (buyer and seller, or firm and customer) must believe they are better off if they perform an exchange than if they do not. We can think of measuring how much each side gains in monetary terms (in effect, how much they would have to be paid to forego the transaction). The sum of these gains from trade for each party is the *surplus* (denoted by S) created by that transaction (see Table 1 for a glossary of the variables used in the theory).

Because we want to perform competitive analysis, we are interested in situations where more than one potential transaction partner is available: Seller 1 or seller 2 may both be able to exchange with a particular buyer. With this setup, consider the following definition:

Seller 1 has *transaction surplus superiority (TSS)* over seller 2 with respect to a specific transaction if and only if the surplus from a transaction between 1 and the buyer is greater than the surplus from a transaction between 2 and the buyer, i.e., if and only if $S_1 > S_2$.

It is immediately apparent that this definition satisfies the criteria Complete Ordering, Independence of Sustainability, Independence of Prices, and Transaction Focus. CO is satisfied because the relationship "has at least as much surplus" completely orders the sellers. There is no mention of the persistence of this relationship, so IS, avoiding the conflation of competitive superiority itself with sustainability of that superiority, is accomplished. As the concept of surplus makes no reference to prices, which act to divide up the surplus rather than determine its size, IP is satisfied. Finally, because the definition refers to advantage with respect to a specific transaction, TF is met as well; seller 1 could have transaction surplus superiority over 2 in dealing with buyer A but be inferior when dealing with some other buyer B.

In order to see how well this definition satisfies the other criteria, I will define a concept of willingness to pay, or *value* (V) and then characterize buyer behavior in terms of it. We can then understand surplus as being generated by the difference between this value and the *cost* (C) to the seller to provide it. Firms' medium-term policy choices about product attributes, process configuration, and capacity levels affect both willingness to pay and firms' costs. The causal links from these choices to surplus are commonly referred to as *cost and value drivers.* Firms' choices operate through these drivers to determine relative surplus and hence TSS as defined above. Thus, (4), Managerial Control, is assured.

The next step will be to show how superior surplus creation, under a wide range of plausible competitive behaviors, leads to higher expected profit from a transaction, so that the transaction surplus superiority concept satisfies Operational Guidance. Finally, I will establish that TSS is not synonymous with higher profits, and so is not a trivial tautology.

4.1 | Value and buyer's surplus

The buyer places some value, V_1, on the product or service (hereafter I will use "product" to mean "product or service") offered by seller 1. This value measures the buyer's *willingness to pay* for seller 1's product. V is measured in dollars or some appropriate monetary unit.[8,9] Since we are going to want to compare the Vs of rival firms, the level of V for firm 1 cannot be allowed to depend on the V (or price) of rival firm 2 or else we would end up in an infinite regress, where we could not define the value of any product without defining the values of all others.[10] Accordingly, the full definition is:

> V is the maximum amount the buyer is willing to pay for a firm's product or service, *assuming that none of the other sellers to be analyzed is available to the buyer.*[11,12]

Value is only part of the explanation of buyer behavior, however. The actual economic surplus captured by a transacting buyer is the *difference* between his value for the product and the price he had to pay to get it. Thus, I define $V_i - P_i$ as the *buyer's surplus* from transacting with seller i; when considering the prospects prior to transacting, I will call these amounts *offers.* The default assumption about buyer behavior is that buyers are "fully responsive" and will *always take the best positive offer they get, choosing randomly in the case of ties.*[13] Note that this assumption about buyer behavior implies nothing about the level of price competition, which depends entirely on the

[8]Other writers (e.g., Ghemawat, 1991) use the term "benefit" instead of value, reserving "value" for what I call surplus. Hoopes, Madsen, and Walker (2003) agrees with the usage here. No strong argument can be made for either nomenclature; I prefer sticking to the standard language of economics in discussing consumer and producer surplus, as well as the ordinary meaning of "valuable."

[9]A number of empirical methods for assessing V have been developed in both marketing and economics, including estimation of random logit demand models (Berry, 1994, Train, 2003), auction mechanisms (Wertenbroch & Skiera, 2002), and direct calculation of benefits (Best, 2000, Chapter 4).

[10]This nuance is not explicit in the otherwise similar benefit/value definitions of Adner and Zemsky (2006); Besanko, Dranove, and Shanley (2000); Ghemawat (2006); Hoopes et al. (2003); or Peteraf and Barney (2003).

[11]Imagine drawing a boundary around the focal firm and all the rival and substitute products and services that one wishes to include in the analysis. V is defined as the maximum willingness to pay for a given product if everyone else inside this *analytical boundary* were to disappear. Providers outside the analytical boundary (and their prices), however, are very much considered in the determination of V; they are substitutes that constrain willingness to pay for the focal products.

[12]An inapt drawing of the analytical boundary can cause distortions in the calculation of TSS by distorting the excluded alternatives. If products 1 and 2 are relatively similar to one another compared to product 3 and the boundary is drawn so as to include products 2 and 3 but not 1, then product 3's relative value will be exaggerated because 3's value is less depressed than 2's by 1 lurking outside the boundary. To avoid such distortions, the analytical boundary should be drawn so that excluded substitutes are roughly equal in their "distance" from included products. I am indebted to Ed Fox for this point.

[13]There are two caveats about this assumption. First, it rules out the possibility of the buyer bargaining with the seller(s). In some environments, bargaining is the more natural way to model the situation, in which case the weaker predictions of bargaining theory will have to suffice. Second, the assumption imposes a kind of regularity and consistency in buyer behavior that may seem unrealistic. But the default assumption is an approximation that can be replaced by a more realistic response model where the probability of the buyer picking offer 1 over offer 2 increases as the superiority of offer 1 over offer 2 increases. Such "attraction models" are common in the marketing literature, especially logit brand-choice models (Train, 2003, Chapter 6). Besanko, Gupta, and Jain (1998) perform an empirical analysis of equilibrium pricing and buyer behavior in the yogurt and ketchup markets combining logit brand-choice behavior with the surplus creation idea discussed here. Their results are consistent with the claim that "competitive advantage," defined as TSS, is an important precursor to superior profitability. Since this paper intends no contribution to the brand-choice literature, it will focus on the simpler, if less realistic, default assumption of fully responsive buyers.

level of forbearance firms display in making aggressive offers to steal one another's potential customers—we do not assume Bertrand pricing behavior even in the one-shot game.

4.2 | Cost and seller's surplus

Executing a transaction with a given buyer incurs a cost *C*. Since the firm earns revenue *P* from the exchange, its contribution received, or *seller's surplus*, is $P - C$. Adding together the buyer's surplus and the seller's surplus causes price to cancel out, so the sum is total surplus $V - C$. Thus, total surplus is independent of prices.

C may depend upon the number and type of other transactions that the firm engages in at the same time or in the future (i.e., there may not be constant returns to scale, volume, or scope). The main motivation for the Transaction Focus criterion is to eliminate aggregation bias, so in general one should focus on a given transaction and hold constant all the other transactions the firm intends to perform. *C* should thus be thought of as the *incremental* cost of dealing with this particular buyer's needs.[14]

It is also very important that the costs built into *C* include only those expenditures actually contingent on performing the hypothetical transaction at the time of decision. Thus, all previously incurred sunk expenditures should be ignored in the construction of *C*, since they will not be recovered by forbearing from the transaction. This exclusion of past sunk costs is critical if TSS is to reliably guide managerial decision-making rather than merely provide historical analyses of profitability, meeting the Not a Trivial Tautology, Managerial Control, and Operational Guidance criteria.

5 | TRANSACTION SURPLUS SUPERIORITY, REASONABLE PRICING BEHAVIOR, AND RELATIVE PERFORMANCE IN THE CANONICAL CASE

In order to see how transaction surplus superiority relates to performance while maintaining agnosticism about the toughness of price competition we need a setting where (a) the toughness of price competition is easily defined and (b) TSS is not affected by the pricing game played after sunk investments and strategic choices have been made. Ideally, this setting should be simple enough to render visible the mechanisms relating TSS, price competition, and performance. To that end, I will define a "canonical" case of post-investment price rivalry.

The canonical case for thinking about transaction surplus superiority and profitability comprises

1. A set of identical potential customers (with a collective mass of one) trying to fulfill a certain need, each of which has value V_j for firm *j*'s solution to his problem.
2. *Fully responsive* customer behavior, where buyers compare each firm's *offer* $b_j = V_j - P_j$, always choose the largest positive one, and randomly choose among the best offers if there is a tie. Note that firms with different prices may be offering the same amount of surplus, due to the differences in value across firms.
3. A set of *N* firms, each with its own value V_j and constant unit cost C_j, with no capacity constraints and with common knowledge of all parameters.

[14]In cases where the transaction at issue is just one of a large set of *homogeneous* transactions, where all the relative Vs are the same and there are no buyer-specific differences in *C*, it is possible to treat *C* as the *average* cost per transaction in the set. But this simplification will only work with such homogeneity, when there is no risk of aggregation bias. That is exactly what will be accomplished later by looking at "canonical market segments."

This setup is the simplest and most widely relevant case for analysis. If one cannot understand the canonical case, more complex situations will be hopeless. In a two-step process, with sunk investment decisions that affect V and C being made in the first stage, one can still model increasing returns without creating a causal link from market share to our intermediate measure of strength, TSS. Constant returns in the second stage ensure that market share may be the result, but never the cause, of TSS.[15] A thorough understanding of canonical segments under varying price regimes and with simple investment possibilities appears to be an essential building block to any theory of positioning strategy.

Homogeneity of buyers means that the demand side of the canonical case models a *market segment,* a group of customers with a distinct set of needs and preferences.[16] Much as white light is made up of a spectrum of pure colors, a firm's potential demand is composed of a set of canonical segments. In any specific empirical context, overall demand can be approximated either by aggregating segments into a finite number of discrete groups or by postulating a continuous distribution of canonical types. The total number of purchases in a single canonical segment does not change with the size of the offers made (unless offers become negative, in which case there are no purchases); all customers buy exactly one unit each from whatever firms make the best offer.[17]

Note that *there is no price-competition assumption built into the canonical case.* This feature contrasts with the literature on "competitive advantage" in both the resource-based and value-based streams, as these are committed to Bertrand or perfect-competition behavior in a setting such as the canonical market segment, where each firm's capacity exceeds quantity demanded. Here we do not impose such aggressive pricing behavior—Nash equilibrium behavior—even in the modeled one-shot pricing stage at the end of the game.[18]

In general, firms are differentiated on V within the homogeneous customer segment, so firms making the same offer will by necessity charge different prices P; instead of a "law of one price" we have a "law of one offer." Within this "price agnostic" or "rivalry agnostic" framework it makes sense for firms to avoid needlessly pricing themselves out of profitable sales or needlessly losing money on unprofitable sales. We can formalize this idea by defining *reasonable pricing behavior for firm i* as obeying

$$\text{Min}\left\{\text{Max}_j b_j, S_i\right\} \leq b_i \leq S_i \tag{1}$$

The left-hand expression in Equation (1) is the lesser of the highest offer out in the market and i's total surplus. This first inequality means that firm i will not be undersold and driven from the market unless rival offers exceed its total surplus creation, in which case the second inequality tells us that it offers exactly that total surplus. Reasonable pricing behavior doesn't pin down the exact level of price competition. It merely ensures that firms are neither (a) so aggressive as to incur losses in a quest for positive market share nor (b) so passive that they voluntarily accept zero sales when they could profitably match or exceed the best rival offer. We will maintain the assumption of reasonable pricing behavior throughout this article, including during later discussions of incentive-compatible pricing.

[15]The canonical case trades off modeling of industry-level demand elasticity, customer heterogeneity, and economies or diseconomies of scale (gross of sunk investments) in order to enable clear modeling of (a) product and process heterogeneity (the results of differing capabilities and positioning choices) across an arbitrary number of rivals and (b) different levels of tacit collusion among rivals. Sunk investments in Stage 1 allow for increasing returns to be captured despite the constant-returns assumed post-investment.

[16]Stuart (2015) also uses a segment-specific approach to defining "value gaps" that are equivalent to the definition of TSS employed in this article.

[17]In some contexts, a canonical segment may be a reasonable approximation of the entire market, depending on market definition. When it is not, it is important to realize that a firm's profit in a single canonical segment is only part of its total profit in that market.

[18]For those insistent on Nash equilibrium, the last stage can be interpreted as an approximation to a multistage pricing game in which quick-response or Folk Theorem equilibria with prices above Bertrand levels are equilibria.

Greater *behavioral rivalry* (holding fixed the number of firms) is defined as the prevailing offer being higher. Greater *structural rivalry* (holding fixed the prevailing offer) is defined as a larger number of rival firms. The degree of behavioral rivalry can thus vary from perfect collusion, where each firm j sets $P_j = V_j$ (equivalently sets $b_j = 0$), to differentiated Bertrand competition, where the most economically efficient firm knocks all others out of the market with an offer just bigger than the total surplus of the second-best firm, $b_{leader} = S_{secondbest} + \varepsilon$ or $P_{leader} = V_{leader} - S_{secondbest} - \varepsilon$. (If there is more than one firm tied for greatest surplus, then the Bertrand offer is the entire surplus.)

We want to know if TSS reliably leads to higher profits in the canonical case. Because customers are identical, firm i creates the same surplus S_i for every potential buyer, although the firms will generally differ from one another in the surplus they create, i.e., S_i will differ from S_j. (This is an advantage of the V-P-C setup; it's easy to think about varying-efficiency, differentiated rivals competing head to head.) Define the index over firms so that $S_1 \geq S_2 \geq \ldots \geq S_N$. Because of the constant unit cost assumption in the canonical case, no firm's surplus creation depends on its output level or market share, an important simplification. We can now state

Proposition 1 (Offer/Share Cutoff) *In the canonical case with reasonable pricing behavior for all firms, there exists* $M \in \{1,\ldots,N\}$ *and* $\underline{b}$ *such that.*
(a) $b_j = \underline{b}$ for $j = 1, \ldots, M$ *and* $\underline{b} > b_k$ *for* $k = M + 1, \ldots, N$, *with* $0 \leq \underline{b} \leq S_j$ *for all* $j \leq M$ *and* $\underline{b} > S_k$ *for all* $k > M$.
(b) The distribution of market shares is fully described by M *with all* $i \leq M$ *attracting* 1/M *of the market and* $i > M$ *having zero market share.*

Proof Please see Appendix 1.

Part A says that the set of "top M surplus creators" all make the same offer, which the bottom N-M firms cannot match because their surpluses are too small. Note that M could range from 1, meaning that firm 1 knocks everyone else out of the market, to N, meaning that every firm is able to participate, capturing the full range of possible levels of price competition. Part B says that the M firms with surplus big enough to operate in the market segment split the customers evenly, while the N-M firms at the bottom have zero market share.[19]

Proposition 1 tells us how offers and market shares will be configured under a wide range of pricing behaviors. In a Bertrand regime with a single leader, the leader captures the whole market with an offer $\underline{b}$ that is just greater than the next-highest surplus in the industry. (With multiple firms tied for highest total surplus and Bertrand pricing, only the leaders have positive market share and the market offer is equal to that highest surplus.) At the other extreme of a perfectly collusive regime, $\underline{b} = 0 \leq S_N$, $M = N$, and every firm has market share $1/N$. Market offers in between these two extremes leave $1 < M < N$ for $N > 2$. Now we turn to the implications for profit:

Proposition 2 (TSS and Gross Profit Superiority) *Let* π *represent profit (gross of past investment). In the canonical case with reasonable pricing behavior,* $\pi_i - \pi_j = (S_i - S_j)/M$ *for all firms* i *and* j *with positive market share, so the profit difference between these firms is exactly proportional to TSS. A firm* i *with positive market share has profit difference with a firm with zero market share of* $\pi_i - 0 = (S_i - \underline{b}/)M$, *which is a positive affine function of TSS.*

Proof In general, $P - C = S - \underline{b}$. So for firms with positive market share, $\pi_i - \pi_j = [(S_i - \underline{b}) - (S_j - \underline{b})]/M = (S_i - S_j)/M$. The comparison with a firm with zero market share is trivial.

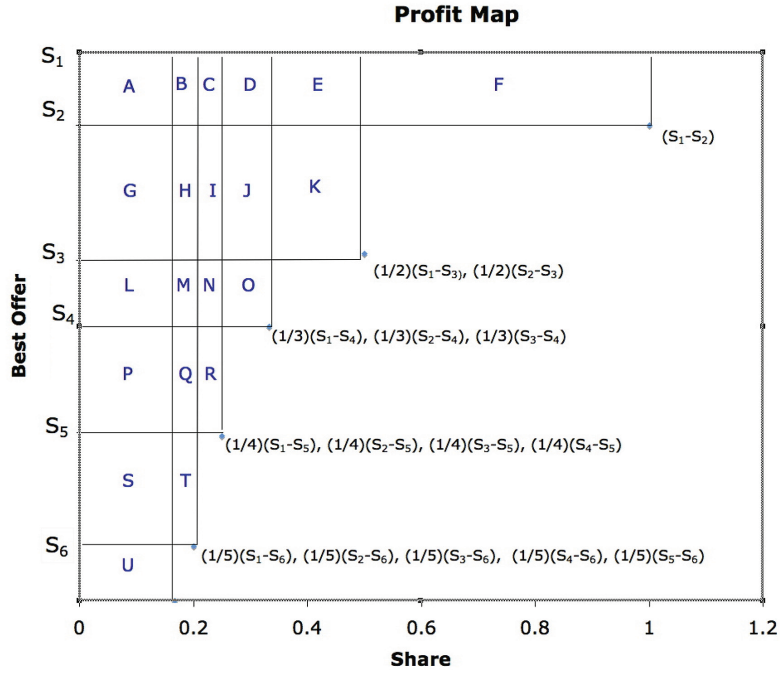

FIGURE 2 Profit map

Proposition 2 shows that *regardless of the degree of price competition*, in the canonical case with reasonable pricing behavior profit differences between positive-market-share firms are a linear function of TSS. Increasing the level of price competition (weakly) reduces M, the number of positive-market-share firms, so that each one has a higher share. It also reduces the profit per transaction $S_i - \underline{b}$ of each such firm, reflecting the usual tradeoff between volume and profit margins. In a Bertrand case, if the strongest firm has strictly higher surplus than any rival, only that strongest Firm 1 garners positive sales and it earns profit $S_1 - \underline{b} = S_1 - S_2$ which is just 1's transaction surplus superiority over 2.

Figure 2 introduces a new tool for understanding competition in canonical segments—the *profit map*. In this case, it shows a hypothetical example with six firms. The horizontal axis measures market share, which varies from zero to one. The vertical axis measures the size of the highest offer $\underline{b}$ in the market, which ranges from zero to S_1. The "profit staircase" in the diagram shows what each firm earns as the toughness of price competition varies. Under perfect collusion, with $\underline{b} = 0$ each firm has a market share of one-sixth and earns profit per unit equal to its surplus creation, so Firm 1 earns A + G + L + P + S + U, Firm 2 earns G + L + P + S + U…, and Firm 6 earns U. With very soft but not perfectly collusive price competition, we might have $0 < \underline{b} = a < S_6$ with each firm still having one-sixth the market but now earning a less per unit than under perfect collusion, i.e., Firm 1 would earn A + G + L + P + S+ (U − a) and so on.

If $\underline{b}$ were to reach the critical threshold of S_6 (actually an infinitesimal amount above it) then Firm 6 would not sell anything under reasonable pricing behavior, because it would have to incur losses to attain positive market share. The remaining firms would each have one-fifth of the market and Figure 2 shows that Firm 1 would earn A + G + L + P + S + B + H + M + Q + T. Compared to perfect collusion, $\underline{b} = S_6$ costs Firm 1 the region U (lower margin) and adds

[19]The *N-M* "priced out" firms could be thought of as "latent" rivals operating in other market segments or as potential entrants, depending on the specific context being modeled.

regions B, H, M, Q, and T (increased market share). Firm 2 earns G + L + P + S + H + M + Q + T, which represents a loss of U and a gain of regions H, M, Q, and T relative to perfect collusion. The pattern repeats for Firms 3, 4, and 5 which each gives up U and gains the regions in the second "column" that lie below its surplus level. The series of expressions in Figure 2 starting at the lower right corner of T gives the areas of these regions in terms of shares and margins for each firm with positive market share, so Firm 1 earns $(1/5)(S_1 - S_6)$, Firm 2 earns $(1/5)(S_2 - S_6)$, and so on.[20] Raising $\underline{b}$ continues the pattern of chopping off rows and adding columns to surviving firms' profits. At the last "stair," with $\underline{b}$ just above S_2, we have Bertrand competition: Firm 1 has a market share of one, all other firms have zero shares, and Firm 1 earns $A + B + C + D + E + F = S_1 - S_2$.

The only "exogenous" part of Figure 2 is the number of firms and the distribution of surplus levels among them. Hence, the number of firms and this distribution uniquely identifies a canonical-case market-segment profit map. The profit map thus depicts a context-free relationship between the equally context-free concepts of surplus and profit. With an additional assumption about the level of price competition, the map gives the exact distribution of profits. Context-specific information is required only at an earlier stage when studying how choices of product and process generate different surplus levels—that is, when one is engaged in the analysis of cost and value drivers for the segment.

It follows from this analysis that superiority in total surplus creation for a transaction—TSS—is correlated with superior performance and so meets the Operational Guidance criterion. But is the fit *too* good so that TSS violates the Not Trivially Tautological criterion? Four arguments suggest that, unlike definitions that directly build in the idea of supernormal profits, this definition is not vacuous.

First, there is a causal story that links surplus differentials to profits, invoking the concept of value, which is not part of the formula for profits. Profit differences are the result of surplus differences but are not the same thing. Transaction surplus superiority is thus the first answer to the question "why do some firms in an industry segment perform better than others?"[21]

Second, the assumption of equilibrium pricing behavior need not hold. If a firm prices incorrectly, then it can underperform or even be knocked out of the market. Third, the C that is subtracted from V to compute surplus does not include any sunk costs that may have been incurred in the past in order to produce the strategic position a firm now enjoys. There is nothing in the analysis to guarantee that these past sunk expenditures were justified even if a firm attains TSS as a result of them. It is not only logically possible but arguably observable that large past sunk investments are sometimes regretted even though they create superiority in the sense advocated here.

Finally, the need for *commitment* (Ghemawat, 1991, 2006) can sometimes render TSS irrelevant in determining relative profits. For example, consider a non-canonical market where recurring fixed costs are large relative to the amount of demand, so that only one firm can operate at a time without incurring losses. Even for a firm with substantial transaction surplus superiority, if a rival can credibly commit to staying in this market in the face of losses then it may be rational for the firm with advantage to exit (or not enter the market in the first place).

So TSS should be correlated with profits under a wide variety of competitive conditions. At the same time, the link is causally non-trivial, subject to proper managerial perception and execution of pricing tactics, compatible with mistakes in sunk investment that lead to overpaying for superiority, and sometimes overridden by differences in commitment.

[20]Between the levels of the staircase represented by two consecutive surplus levels, increases in $\underline{b}$ leave market shares unchanged and shrink each firm's profit per unit. Such "intermediate" prevailing offers are thus collectively inefficient for the firms in the industry.

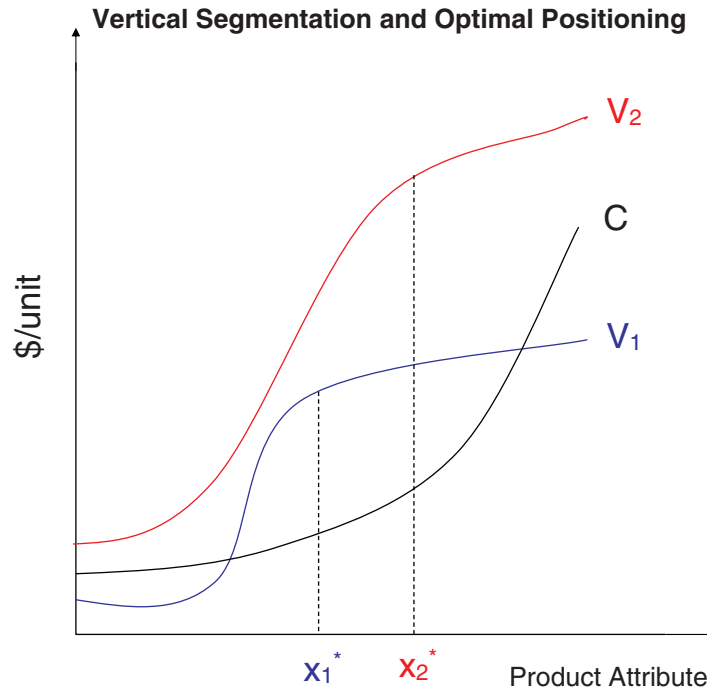

FIGURE 3 Vertical segmentation and optimal positioning

5.1 | Strategic implications of TSS in the canonical case: positioning with forced choices and value/cost tradeoffs

Propositions 1 and 2 establish the importance and relevance of the TSS intermediate measure for firm performance when the level of price competition is allowed to take on any value between collusion and Bertrand. In order to show how TSS meets the Managerial Control and Operational Guidance criteria, this section shows how TSS informs "forced choice" first-stage decisions about competitive positioning, those where all options incur equal levels of sunk cost so that tradeoffs of *V* and *C* are the only consideration.[22]

The abstraction provided by the TSS concept clarifies and unifies positioning analysis. One can decide whether to improve (or degrade) a product attribute in order to raise (or lower) *V* while raising (or lowering) *C* simply by investigating whether the increase (or decrease) in *V* would exceed the increase (or decrease) in *C*. If the answer were favorable, then the shift would be worthwhile regardless of any assumptions about price competition or even about the number of firms—proposition 2 ensures that result.[23] In principle, *every* product and process choice the firm makes

[21]There are three basic approaches to explaining observed firm heterogeneity with firms experiencing different payoffs. First, one could have non-equilibrium (i.e., rationally random or remediably mistaken) behavior with weak selection pressure. Second, one could have path dependence in which today's heterogeneity is the byproduct of yesterday's (e.g., differences in resource inheritances). Third, one could have an equilibrium that is necessarily symmetry breaking, as in duopoly models with customers who vary in their willingness to pay for quality. TSS is consistent with any of these mechanisms.

[22]Of course, many strategy choices require the firm to decide whether or not to incur an exogenous sunk cost, or to choose the level of an endogenous sunk cost, in order to increase transaction surplus. We will take a first cut at analyzing such choices in the seventh main section of this article.

[23]In some circumstances it may be optimal for a firm to reduce TSS in one segment in order to avoid cannibalizing sales in another segment, e.g., when Intel launched its low-powered Celeron chip as a "fighting brand" against AMD's lower-powered and cheaper offerings.

(where sunk costs are the same across options) can be assessed with this "surplus engineering" test, an enormous unification and simplification of a host of diverse topics in operations, marketing, and strategy. To the extent that heterogeneous pre-existing resources or capabilities matter, they do so only by affecting transaction surplus superiority.

Figure 3 illustrates a canonical-segment positioning choice about the level of a product attribute x (or composite of attributes that represents "quality"). The horizontal axis shows the level of this attribute while the vertical axis registers dollars per transaction, which is the metric for both V and C. The two V curves representing different canonical segments 1 and 2 as well as the C curve all rise as the level of the attribute increases; the former because the attribute is defined to be something customers want and the latter because increasing desirable attributes usually costs more.[24] The concavity of the V curves reflects diminishing marginal value of the attribute while the convexity of the C curve shows increasing marginal cost.

Optimal positioning in canonical segment 1, shown as point x_1^*, drives V_1 and C as far apart as possible. The optimal point occurs where marginal V and marginal C (the slopes of the two curves in Figure 3) are equal. For two firms positioned both below or both above x_1^*, the one that is closer to x_1^* will always have TSS over the one that is farther away. The identical arguments hold in segment 2. The relevant positioning difference between the two segments is that the *marginal* value of x is higher for segment 2 customers than it is for segment 1 customers, causing x_2^* to exceed x_1^*. (A vertical shift in either direction for a V curve leaves the optimal position unchanged.)

Simply knowing that the qualitative shapes and solution are likely to follow the pattern of Figure 3 is very helpful in finding successful positions. For one thing, it shows that *at the segment level* the Porter (1985) advice to avoid "getting stuck in the middle" is exactly wrong—the optimal position in the segment is very likely to be at an interior point given the respective concavity and convexity of the V and C curves. As a consequence, with more than two distinctly positioned firms, an optimally positioned player would likely have neither the highest value nor the lowest cost but would be intermediate on both.

6 | INCENTIVE-COMPATIBLE PRICING AND TSS IN THE CANONICAL CASE: THE WEAK LEADERSHIP PRINCIPLE

The second contribution of this study attacks a long-standing problem in industrial-organization economics. As discussed above, there exists a plethora of economic models that can rationalize any degree of price competition ranging from Bertrand to collusion. Up to this point, therefore, the analysis has agnostically treated the level of price competition as an exogenous parameter defined by the size of the "prevailing offer" $\underline{b}$. It turns out, however, that the economics literature's default assumption—that firms are completely symmetric—obscures plausible lower bounds on the degree of price competition. In strategy, where firm heterogeneity is expected, we can use the TSS concept to find these bounds.

6.1 | The weak leadership principle

In a canonical market segment with firm subscripts ordered by surplus in the usual way, compare Firm 1's profits under any prevailing offer $\underline{b}$ (where M firms split the market) with its Bertrand profits if it prices to just knock Firm 2 (and hence all other firms) out of the market:

[24]In the long term, advances in technology and organization my make it possible to increase quality and reduce operating costs at the same time, but in Figure 3's medium-run presentation that would look like a downward shift in the C curve.

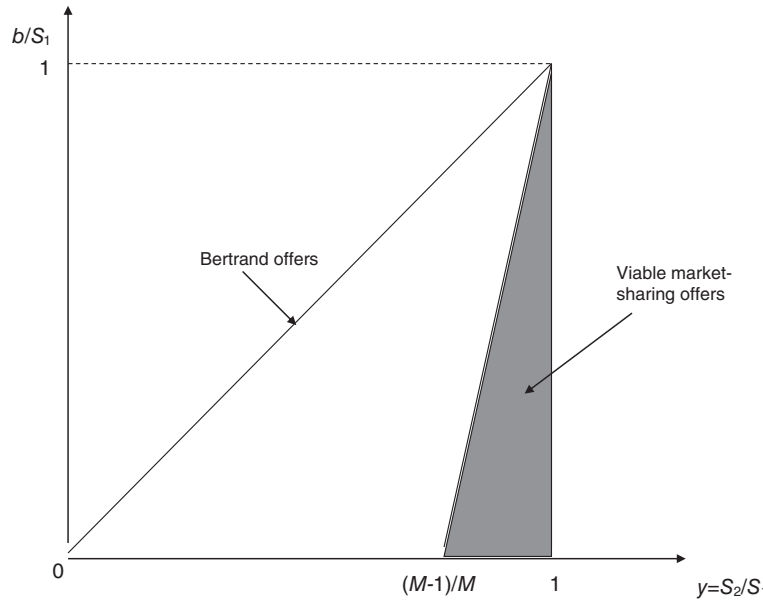

FIGURE 4 TSS and viable offers under the weak leadership principle

$$\pi_1\left(\underline{b}\right) - \pi_1(\text{Bertrand}) = \left[\left(S_1 - \underline{b}\right)/M\right] - (S_1 - S_2) \tag{2}$$

Figure 2 shows that for a best prevailing offer $\underline{b}$ just above, say, S_4, Firm 1 would earn A + B + C + D + G + H + I + J + L + M + N + O = (1/3)(S_1 − S_4). By contrast, Firm 1's Bertrand profit is given by A + B + C + D + E + F = S_1 − S_2. In going from just knocking out firms 4, 5, and 6 to Bertrand, Firm 1 gives up G + H + I + J + L + M + N + O and gains E + F. So evaluated at $\underline{b}$ just above S_4 the expression in Equation (2) would be S_2 − (1/3)S_4 − (2/3)S_1. Expression (2) tends to be negative (i.e., Firm 1 tends to favor the Bertrand outcome) when Firm 1's transaction surplus superiority is higher, when the number of firms splitting the market under soft competition is higher, and when the prevailing offer under soft competition is higher.[25]

If Equation (2) is negative, then Firm 1 can unilaterally raise its offer to $b_1 = S_2 + \varepsilon$ (where ε is an infinitesimal positive number) and improve its profitability. It need not worry about rivals' responses, as no rival could match its offer without incurring losses. So, for a big enough transaction surplus superiority and/or a fragmented enough market, the leading firm should price to capture the whole market segment. Now define:

> **Weak Leadership Principle (WLP)** *With reasonable pricing behavior as in Equation* (1)*, no prevailing offer* $\underline{b}$ *can survive in a canonical segment as an outcome if it makes Equation* (2) *negative,* i.e., *if it is more profitable for the leading firm to drive out all of its rivals with a Bertrand offer. A Bertrand offer, however, can survive even if it is less profitable than some smaller alternative offer. Formally,* $\underline{b} \in \{[0, \text{Max}\ \{S_1 - M[S_1 - S_2],0\}], S_2 + \varepsilon\}$.

The WLP says that the prevailing offer is either Bertrand, $S_2 + \varepsilon$, or else it falls in the interval from 0 to $S_1 - M(S_1 - S_2)$ (when this expression derived from Equation (2) is non-negative). Any

[25]The finding that increasing competitive advantage eventually makes the leader of an N-firm segment favor Bertrand pricing over any softer alternative is consistent with Makadok's (2010, p. 362) duopoly analysis comparing Cournot and Bertrand equilibria.

prevailing offer bigger than this expression but smaller than the Bertrand offer is ruled out—the leader could unilaterally impose Bertrand and do better.

This principle is "weak" because it does not empower the leader to arrange a "soft landing" at any prevailing offer greater than $\underline{b}$ but less than $S_2 + \varepsilon$. It is easy to start a price war, but it may not be easy to coordinate the behavior of the rest of the firms in the segment, some of which might interpret (or misinterpret) a modest price cut as precipitating a general grab for market share and drive offers up once discipline has been broken. The principle also does not rule out a Bertrand offer prevailing even when other offers might be better for the leader. The leader can only take the prevailing offer or else "go nuclear." The WLP is perhaps the smallest concession from pure "offer agnosticism" possible that can still account for the incentive effect of firm heterogeneity. (Appendix 2 provides conditions under which a Strong Leadership Principle, where the leading firm directly imposes its most-favored offer, might apply and shows that transaction surplus superiority is even more valuable when that principle operates.)

Despite these limitations, the WLP by itself meaningfully constrains possible outcomes in fragmented segments. With ten potential rivals and Firm 2 creating anything less than 90% of the surplus of Firm 1, for example, there does not exist *any* offer $\underline{b}$, even zero (perfect collusion), that would survive the Bertrand outcome. With only two firms in the segment, if Firm 2 generates less than half the surplus of Firm 1, then no non-Bertrand outcome satisfies the WLP again.

> **Proposition 3 (WLP and Rivalry)** *Let* $S_2 = y^*S_1$, *with y in the interval* [0,1], *Under the WLP, Bertrand is certain when* $y < [M - 1]/M$. *For* $y > (M - 1)/M$ *there will be a range of prevailing offers* $\underline{b} \in [0, [1 - [1 - y]M)S_1]$ *for which soft price competition prevails over Bertrand (provided the outcome starts there).*

Proof Set $\underline{b} = 0$ in Equation (2) and substitute $S_2 = y^*S_1$, then solve for the critical *y*.

When the WLP is an appropriate assumption, proposition 3 describes how firm heterogeneity interacts with industry fragmentation to limit the range of rivalry outcomes. We find, for example, with $y = ¾$ and $M = 3$, that soft price competition with $\underline{b} \in [0, (1/4)S_1]$ survives the Bertrand alternative under the WLP. But if we added a firm so that $(M - 1)/M = 4/5$, then we would be forced back into Bertrand rivalry. *In other words, the greatest integer less than or equal to 1/(1 − y) is the maximum number of firms that can be supported in a canonical segment under the WLP.* Here we have a "structural" restriction on market structure dictated not by increasing returns to scale but by firm heterogeneity and transaction surplus superiority, a business-strategy contribution to a long-standing problem in the theory of industrial organization.

Figure 4 gives a geometric display of proposition 3. The horizontal axis plots *y*, the ratio of S_2 to S_1, while the vertical axis plots the ratio of the prevailing offer $\underline{b}$ to S_1. The set of viable offers given *y* is the union of the 45° line representing Bertrand competition $\underline{b}/S_1 = y$ with the shaded triangle representing smaller offers that the leader likes at least as much as Bertrand competition. For *y* below the cutoff $(M - 1)/M$ the only WLP-conforming offers are Bertrand; for any larger *y*, a continuum of offers rising vertically from the horizontal axis to the hypotenuse of the shaded region is also viable.

The WLP also predicts that the entry of less-efficient-than-incumbent firms into a segment could precipitate a sudden decrease in prices if it induces the leader to switch regimes and knock out all the other firms. (In such a case, the entering firms would have made a mistake by coming into the segment since any sunk investment they made to do so would be wasted.)

The WLP implies that a leader starting with a lower-than-Bertrand prevailing offer whose TSS increases sufficiently will choose to cut prices. That means buyers should be eager to see such surplus improvements on the part of the segment leader, because they would benefit by a reduction in the price they are required to pay. Perhaps more surprisingly, buyers would also benefit from a successful effort by the segment leader to raise the costs (or, more generally, lower the surplus) of the second-best firm sufficiently to tip the segment into Bertrand competition. Paradoxically, under the WLP, making the second-best firm *less* efficient could increase the total surplus created and captured by all parties in the segment.[26]

Under the WLP, gaining transaction surplus superiority in a segment has even more value for the leader than it does under the agnostic pricing initially considered. TSS under the WLP allows the leader to exploit opportunities that trailing firms cannot; a second-best firm might prefer, say, that the prevailing offer rise from the fifth-best surplus to the third-best, but the WLP gives it no power to make that happen.

The WLP (and the strong leadership principle described in Appendix 2) suggest that greater leader superiority over actual *and latent* segment rivals leads to more aggressive price competition and a more concentrated market structure in that segment. The WLP also suggests that the entry of more rivals in a given segment ought to be correlated with more-aggressive pricing behavior by the leader and lower profits for the followers. In fact, a second-best duopolist operating in a segment might find that an increase in latent segment rivals (say, because more firms start to operate in related segments and their sales spill into this one) causes the leader to jack up its offer enough to knock that second-best firm (and all others) out of the segment.

7 | SUNK INVESTMENT CHOICES AND TRANSACTION SURPLUS SUPERIORITY

The third contribution of the transaction surplus superiority framework is to show how irreversible investments to increase surplus should be assessed. Unlike the "forced choices" discussed above, where a firm would always want to maximize TSS because all options incurred the same sunk cost, here we consider circumstances where one choice (investing to raise S) involves greater sunk cost than another. Since TSS explicitly excludes sunk costs from its definition of surplus, whether a given sunk investment pays depends on what impact it can be expected to have on the investing firm's market share and contribution per transaction under the assumed pricing regime.

7.1 | Sunk investments to increase surplus with agnostic pricing

We will start with the agnostic pricing regime, where before and after investment there is a prevailing offer $\underline{b}$ in the canonical segment. Consider a firm i with an exclusive, last-mover, firm-specific first-stage sunk investment opportunity that costs K and will lead to an increase of surplus to $S_{i2} = S_i + s$, where the $i2$ subscript refers to i's surplus in the second stage. (The exclusivity assumption might hold because only one firm has access to the know-how or resources needed, because all other firms have already made the improvement, or because the investment is only relevant to the idiosyncratic resource and capability position of this firm. The last-mover assumption

[26]The WLP also has counterintuitive implications for merger analysis. Two trailing firms that argue their merger should be permitted on grounds of increased (though not leading) efficiency may be arguing for an inferior social outcome. If they would be forced to exit the segment absent the merger, then making them strong enough to force the leader to accommodate their presence would lower the average efficiency of a transaction in the segment.

is in lieu of building an infinite-horizon model with some postulated arrival process for investment opportunities for each firm.) Because of the sunk cost, in order to know whether the improvement is worthwhile one must know how many firms have positive market share, how big $\underline{b}$ is (i.e., how tough price competition is), whether $S_i + s$ would be below $\underline{b}$, and, if not, by how much $S_i + s$ would exceed $\underline{b}$. The following proposition describes how these parameters interact to determine the answer.

> **Proposition 4 (Exclusive Sunk Investment Rule With Agnostic Pricing)** *In the canonical case with a prevailing offer* $\underline{b}$*, a firm i with an exclusive, last-mover opportunity to raise its surplus by* s *at the expense of incurring a sunk cost* K *earns strictly positive profits from this investment if and only if* $MK < Min\ \{s, S_i - \underline{b} + s\}$ *where* M *is the number of firms with positive market share if i were included in that group.*

Proof Firm i either has zero or positive market share prior to the investment. If it starts with zero share and zero profits, because $\underline{b} > S_i$, then the investment would be optimal if and only if $(S_i - \underline{b} + s) / M > K$, or $(S_i - \underline{b} + s) > MK$. If i starts out with positive market share and hence positive profit $(S_i - \underline{b}) / M$ with $\underline{b} < S_i$, then investment would be optimal if and only if $[(S_i - \underline{b} + s) / M] - K - (S_i - \underline{b}) / M = (s / M) - K$ is greater than 0 or $s > MK$. Since Min $\{s, S_i - \underline{b} + s\}$ is s when i starts with positive market share and $S_i - \underline{b} + s$ when i starts with zero market share, the inequality expresses these decision rules.

Proposition 4 has four significant implications. (a) It is possible for an investment to convey transaction surplus superiority and still be a bad idea if K is high enough relative to s because the former is not included in the TSS definition, and a unique investment to raise firm surplus is more likely to be profitable if its sunk cost is smaller and its surplus gain is higher. (b) If the investment can't raise a firm's surplus beyond the level of the prevailing offer then the investment is valueless. (c) Since Min $\{s, S_i - \underline{b} + s\} = S_i - \underline{b} + s$ only for firms that have $S_i < \underline{b}$ and so start with zero segment share, proposition 4 shows that the incentive to invest is never greater for these out-of-the-segment firms than for firms with initially positive share. (d) Finally, it shows that investment is more likely to be profitable when there are fewer rivals and when price competition among those rivals is less intense. But how can improving one's economic efficiency be more profitable when the market is less competitive? What about the spur of competition?

The reason lies in a subtle feature of the setup for propositions 1, 2, and 4, where the level of price competition is set exogenously as *a fixed prevailing offer size*, rather than a function relating the prevailing offer to the industry surplus distribution. Thus, a firm can't gain market share post-investment by raising $\underline{b}$ to knock out other firms. This vital aspect of agnostic pricing implies that a greater or smaller post-investment gap between S_{12} and S_{22} (compared to the pre-investment gap between S_1 and S_2) does not move the prevailing offer even if the initial offer was set "at the Bertrand level."

Since with agnostic pricing each firm's margin also depends only on the prevailing offer and its own surplus creation, each firm's absolute profit is completely unaffected by the surplus creation of its rivals. An improvement in a rival's surplus therefore neither harms nor helps a given firm; each firm "runs its own race." Their relative profitability depends on their relative surplus, but each firm can tend to its own business without considering rival efficiency. The proposition depicts an extreme form of Williamson's (1991) dictum that "economizing" is more important than "strategizing." Such simplicity departs once we consider the WLP.

7.2 | Sunk investments under the weak leadership principle

The WLP potentially changes the incentives to make sunk investments. Such investments at a given prevailing offer may affect the sign of Expression (2) and thereby change the toughness of price competition, in turn reinforcing or diminishing the payoff to the investment. Consider the exclusive investment opportunity discussed in proposition 4, where only firm i has the option to spend K in order to raise its surplus by s.

Suppose we start with two active firms with $S_1 = 8$ and $S_2 = 6$, while $\underline{b}_1 = 1$. (Note that this initial offer satisfies the WLP, because 1's Bertrand profit $(8 - 6) = 2$ is less than its market-sharing profit $(8 - 1)/2 = 7/2$.) Firm 2 starts out earning $(6 - 1)/2 = 5/2$. If Firm 2 has an investment option to sink $K = 8$ to raise its surplus by $s = 12$, and it takes this option giving it a new surplus of 18, then the WLP says that it should grab the whole segment because $(18 - 8) = 10$ is larger than the profit from keeping the initial offer and splitting the market, $(18 - 1)/2 = 17/2$. But the investment is not worthwhile for Firm 2 because its gross profit increase of $10 - (5/2) = 15/2$ is less than the cost of the investment, 8. Hence the investment should not be made even though it generates TSS.

But now let the starting position include a third Firm 3 with $1 < S_3 < 6$. The initial offer $\underline{b}_1 = 1$ is still supportable under the WLP because $(8 - 6) = 2 < (8 - 1)/3 = 7/3$, but now the investment would look good for Firm 2: Pre-investment profit would be $(6 - 1)/3 = 5/3$, post-investment Firm 2 would "go Bertrand" under the WLP because $(18 - 8) = 10 > (18 - 1)/3 = 17/3$, and the gain in gross profit $10 - (5/3) = 25/3$ exceeds 8, the sunk investment cost. So here the spur of competition in moving from two to three active firms actually does stimulate investment. Moving from market sharing to a Bertrand outcome under the WLP is more desirable when the market-sharing outcome is more competitive and hence less profitable.

In general, competition can only spur investment under the WLP when that investment creates a leader with greater transaction surplus superiority than before and as a result impels that stronger leader to move the market from soft competition to Bertrand competition. The surplus increase s has to be big enough to take the investing firm's TSS beyond the point at which market sharing is optimal, as in the numerical example above. Imposing the WLP both before and after such a potential investment leads to the following characterization:

> **Proposition 5 (WLP Exclusive Sunk Investment Rules)** *Under the WLP, in the canonical case with a prevailing offer* $\underline{b}_1 \in [0, S_2 + \varepsilon]$ *in the first stage with* M *active firms and second-stage* $\underline{b}_2$ *equal either to* $\underline{b}_1$ *or to* $S_{22} + \varepsilon$*, a profit-maximizing firm i with an exclusive, last-mover opportunity to raise its surplus by* s *at the expense of incurring a sunk cost* K *should adopt the following rules:*
> *(a) When* $\underline{b}_1 = S_2 + \varepsilon$ *(first-stage Bertrand with* $M = 1$*):*

- *If* $i > 1$*, so that the investor is initially not active in the segment, then* i *should invest iff*:
  $S_i + s > 2K + S_2$ *when* $(S_1 + S_2)/2 < S_i + s \leq 2S_1 - S_2$ *so i maintains its offer at* S_2.
  $S_i + s > K + S_1$ *when* $S_i + s > 2S_1 - S_2$ *so i switches its offer to* $S_1 + \varepsilon$ *under WLP.*
- *If* $i = 1$*, so that the investor is the active leader under Bertrand, then* i *should invest iff* $s > K$.

> *(b) When* $\underline{b}_1 \in [0, S_1 - M(S_1 - S_2)]$ *(first-stage non-Bertrand offer consistent with WLP).*

- *If* $S_i < \underline{b}_1$ *so* i *is initially inactive in the segment, then* i *should invest iff.*
  $S_i + s > (M + 1)K + \underline{b}_1$ *when* $S_i + s \leq [(M + 1)S_1 - \underline{b}_1]/M$ *so* i *maintains* $\underline{b}_1$.
  $S_i + s > K + S_1$ *when* $S_i + s > [(M + 1)S_1 - \underline{b}_1]/M$ *so* i *switches offer to* $S_1 + \varepsilon$ *under WLP.*
- *If* $\underline{b}_1 \leq S_i \leq S_2$, *so* i *is initially active but not the leader, then* i *should invest iff.*
  $s > MK$ *when* $S_i + s \leq [MS_1 - \underline{b}_1]/(M - 1)$ *so* i *maintains* $\underline{b}_1$.
  $s > K + S_1 - [(M - 1)/M]*(S_i - \underline{b}_1)$ *when* $S_i + s > [MS_1 - \underline{b}_1]/(M - 1)$ *so* i *switches to* $S_1 + \varepsilon$ *under WLP.*
- *If* $S_i = S_1$, *so* i *is the leader, then* i *should invest iff.*
  $s > MK$ *when* $S_1 + s \leq [MS_2 - \underline{b}_1]/(M - 1)$ *so* i *maintains* $\underline{b}_1$
  $s > K + S_2 - [(M - 1)/M]*(S_1 - \underline{b}_1)$ *when* $S_1 + s > [MS_2 - \underline{b}_1]/(M - 1)$ *so* i *switches offer to* $S_2 + \varepsilon$ *under WLP*

Proof Please see Appendix 1.

Part (b) of proposition 5's characterization of optimal investment rules under the WLP reveals that *when investments yield large enough transaction surplus superiority in a segment*, greater competition (more firms or a higher initial common offer) make that investment more likely to be profitable. In each of the cases—initially inactive after the "when", initially active but not leading, and initially the leader—the second pair of inequalities after the "when", showing when competitive leadership is sufficient to trigger Bertrand behavior, reveal that investment is more likely to be profitable when M is higher and/or $\underline{b}_1$ is higher. The reason is that *the status quo market-sharing payoff declines with more firms and/or a higher prevailing offer* but the post-investment single-active-firm Bertrand outcome is unaffected by these factors. Hence, the gain to converting from the status quo to the leadership-Bertrand situation goes up with status-quo competition.

On the other hand, investments in a canonical segment that raise surplus but do not increase TSS up to the Bertrand threshold tend to be discouraged by the presence of greater competition. More competition means that the increase in surplus per transaction is applied to a smaller share of the segment, yielding a smaller total gain. It follows that *more concentrated canonical segments are likely to see higher rates of incremental surplus improvement than more fragmented ones, while more fragmented segments are likely to see higher rates of drastic surplus improvement compared to more concentrated ones.*

Finally, proposition 5 also shows that the incentive to invest (weakly) decreases in the competitive strength of the second-best firm post-investment. The investment rules in all the cases where the investor switches to Bertrand competition feature this next-strongest surplus (either the original segment leader or the original second-best if the original leader has the investment opportunity) on the right-hand side of a greater-than sign, meaning that the higher they are the less likely the inequality is to hold. This finding, together with the others, suggests that *surplus-increasing investments that create or increase transaction surplus superiority are most profitable in segments populated by a large number of weak rivals and least profitable in segments with a small number of strong rivals.*[27]

8 | RELATION TO PREVIOUS LITERATURE

The term most often invoked in the strategy literature to describe an intermediate measure of relative firm strength has been "competitive advantage." But as noted in an earlier section, competitive

[27]This finding reinforces in an investment context the concern raised in the previous section that antitrust policies favoring market structures with "strong second" firms may work against economic efficiency and customer welfare.

advantage has been an ambiguous and equivocal concept through most of the history of the strategy field. Textbook definitions have ranged from superior profitability (Besanko, Dranove, & Shanley, 2000) to inimitability of one's strategy (Hitt, Ireland, & Hoskinsson, 2007); the scholarly literature has been similarly varied.

On the other hand, many authors have adopted the basic idea of a larger wedge between value and cost as an intermediate measure of competitive strength. This general *V-C* orientation toward "competitive advantage" has been gaining ground of late in both research articles and teaching materials (e.g., Ghemawat & Rivkin, 2014). But whatever it is called, there has been little recognition that an intermediate measure of competitive strength should be transaction focused, independent of sustainability, and robust to the pricing regime. As a result, previous intermediate measures have differed from transaction surplus superiority in crucial respects and none has been suitable for analyzing canonical market segments under multiple pricing scenarios.

8.1 | The resource-based and capability views

Transaction surplus superiority is highly complementary to the RBV and capability-based views. TSS is the essential mediating construct between longer-term differences in what a firm is good at and how that translates into payoffs through deployment in product markets. Moreover, as shown in this article's findings about sunk investments, understanding TSS and its impact on competition is critical for assessing the suitability of those investments. The RBV literature has evolved its definition of "competitive advantage" from Barney (1991) evoking "a value-creating strategy not simultaneously being implemented by any current or potential competitors" to Peteraf and Barney (2003) assigning possession of competitive advantage to a firm that "is able to create more economic value than the marginal (breakeven) competitor in its product market" where economic value is defined as "the difference between the perceived benefits gained by the purchasers of the good and the economic cost to the enterprise." While this latter concept is not transaction focused, does not completely order the competitors, and builds in perfect competition to define the breakeven competitor, its spirit is broadly similar to TSS, which can easily be assimilated to RBV and capability theories without creating any specific conflicts.

8.2 | The value-based approach

A partial exception to the trend toward the relative *V-C* formulation of "competitive advantage" is the "value-based" stream of literature initiated in strategy by (Brandenburger & Stuart, 1996, 2007) and MacDonald and Ryall (2004). This literature implicitly divides into two branches, one that dissolves the structural distinction between sellers and buyers and one that maintains it.

The first branch, exemplified by MacDonald and Ryall (2004), says that when the minimum amount a player can receive in a bargaining process [no matter how bad a bargainer it might be] exceeds zero then the player has a "competitive advantage." Formally, if all feasible (i.e., allocating no more than the total surplus available) and stable (i.e., immune to defection by any coalition) distributions of surplus give a player positive payoffs, then it has competitive advantage.[28] This definition is not product-market oriented, allowing a firm to have "competitive advantage" over a supplier or buyer. It also does not provide a complete ordering, and is not transaction focused, two of our seven desirable criteria for an intermediate measure of competitive strength. (One might also question the highly idiosyncratic usage of "competitive advantage" in this work, as it seems to run

[28]In this stream of research, the concept of "added value" (or "marginal product") plays a role somewhat similar to transaction surplus superiority. Added value for a player *i* is the difference in total payoff for the coalition of all parties versus the total payoff for the

counter to the already-noted stricture of Oxley et al. (2010) that theoretical terms ought to be chosen for their community-wide understanding.)

The second branch of cooperative game theory, including Adner and Zemsky (2006) and Stuart (2015), retains the structural difference between buyers and sellers. Both of these papers end up using intermediate measures equivalent to transaction surplus superiority, even meeting the transaction focus criterion by allowing "added value" or "value gaps" to vary by market segments. Stuart (2015)'s findings about profit distribution among rivals coincide with the ones in this article for the case of Bertrand competition with constant returns to scale.[29]

But there is one substantive area of disagreement that causes both branches of the value-based framework to make different predictions from the TSS approach, regardless of their precise definition of an intermediate measure. In the canonical market segments studied in this paper, the value-based approach insists that *only Bertrand competition is possible.* This conclusion follows from the assumption of unrestricted bargaining and the stability requirement for equilibrium—because there are no capacity constraints in a canonical segment, any non-leading firm has zero added value because the leader can serve all customers at greater or equal efficiency. So the value-based approach rules out all but one point in the agnostic-pricing range and also rules out the Weak and Strong Leadership Principles.

It follows that the TSS framework will often give different answers from the value-based approaches about how the pricing stage of the game might play out. As a result, it will also give different answers about how investments should be made in the first stage. Given the wide range of pricing behavior empirically observed and theoretically postulated in different industry settings, as noted earlier in discussing supergames and quick-response-type equilibria, adopting the value-based approach may entail importing unrealistically strong or inappropriate pricing assumptions into the analysis. This importation eliminates study of how firm interdependence affects rivalrous behavior, including such staple topics as multimarket contact and repeated interaction.[30]

8.3 | Interaction of rivalry restraint and transaction surplus superiority

Makadok (2010, 2011) has propounded a framework for studying the interactions of what he terms "four theories of profit" in strategy. Two of those theories are "rivalry restraint" and "competitive advantage," and these are isolated for extended analysis in Makadok (2010). In a classic duopoly model,[31] competitive advantage is defined there as possession of lower constant unit cost than one's rival. With both firms facing a common demand curve for their homogeneous products, this definition would be equivalent to TSS for each customer on the demand curve. Makadok (2010) finds that the profit interaction of rivalry restraint and competitive advantage is *negative*—competitive advantage is worth less when rivalry is restrained and vice versa. This conclusion could alternatively be framed as supporting the idea of "the spur of competition," where firms facing more rivalry have a greater incentive to seek competitive advantage.

coalition of everyone except *i*. Firms with higher added values are likely to earn greater payoffs, and positive added value is a necessary condition for a firm to be guaranteed positive payoff, i.e., to have "competitive advantage" according to the definition in MacDonald and Ryall (2004).

[29]In addition, Stuart (2015)'s sophisticated analysis includes results beyond the canonical segment, considering capacity limits (so that all buyers may not be served) as well as decreasing or increasing returns to scale in a segment.

[30]Moreover, if one were to consider multiple market segments, the value-based approach mandates an assumption of perfect price discrimination, a case that is very important to study but far from covering the universe of possibilities.

[31]The discussion of the Transaction Focus criterion earlier in this article described just one model in Makadok (2010), a Hotelling-style spatial competition model. Most of that paper uses an undifferentiated linear-demand duopoly model.

By contrast, I find that under agnostic pricing, where rivalry is exogenous to TSS, an increase in surplus creation is *more* valuable when pricing is less aggressive and when there are fewer competitors, i.e., when rivalry is "more restrained." Under the weak leadership principle, where rivalry is partly endogenous to transaction surplus superiority, the profit interaction between TSS and rivalry restraint is also *positive* until the leader's advantage becomes so large that it pays for it to drive all others out of the canonical segment.

But the switch point between accommodation and Bertrand competition under the WLP does match the Makadok (2010, p. 362) result, as noted in footnote 25: The larger is the leader's transaction surplus superiority, the more likely it is to want to "put the hammer down" and capture the entire segment, under both the weak and strong leadership principles. Under the strong leadership principle described in Appendix 2, the leader's preferences determine the actual offers made, and proposition A2 shows that a firm's most-preferred prevailing offer is (weakly) increasing in its surplus-creation. Thus, stronger firms tend to prefer bigger prevailing offers (which entail lower prices) than do weaker firms, implying that TSS induces leaders to stop restraining rivalry; conversely, more-restrained rivalry (fewer firms and smaller offers) reduces the incentive to make sunk investments that create transaction surplus superiority, which is consistent with the Makadok (2010) conclusion.

9 | CONCLUSION

The heart of business strategy is the quest for methods to project the performance consequences of firms' product, process, capacity, and customer targeting choices. This quest seeks intermediate measures that are comprehensibly linked to management choices through situation-specific functions or drivers and that are causally linked in turn to firm performance. Such a measure can act as a guide or instrument-panel indicator of the firm's competitive strength; many authors have called this sort of measure "competitive advantage," but there has not been convergence to a single precise definition of that term. Here we have seen that transaction surplus superiority, especially when coupled to the analysis of canonical market segments, can usefully function as the desired intermediate measure of competitive strength (or competitive advantage for those willing to accept that identification).

The first finding of this paper is that transaction surplus superiority satisfies seven important criteria for an intermediate measure of competitive strength. These make it a uniquely suitable target for formulating business strategy, one that can be applied across a wide range of industry contexts to provide directional guidance across the vast majority of medium-term managerial decisions. It provides a universal guideline for all management choices across alternatives with equal sunk-cost implications: Pick the option that maximizes surplus creation. All tradeoffs between willingness to pay and incurred variable costs, such as picking better and more expensive materials versus worse but cheaper ones, are covered by this rule. TSS retains its structure and relevance under any degree of price competition ranging from perfect collusion to all-out rivalry and it mediates the impact of resources and capabilities on payoffs. In canonical market segments the profit map gives a complete description of how TSS and the level of price rivalry allocate market shares and profits across firms.

A second set of findings, and a new concept for both economics and strategy, is that the degree of transaction surplus superiority in a market can actually affect the pricing regime, linking strategy's long-standing interest in firm heterogeneity with oligopoly theory's even longer-standing problem of predicting the degree of rivalry found in a given market context. Through the Weak and Strong Leadership Principles it may be possible to apply incentive compatibility bounds on how soft

competition can be when firms have different abilities to create surplus. In the case of the former principle we can provide precise bounds on how efficient the second-best firm must be to deter the leader from driving all rivals out of the segment.

The third set of findings concerns sunk investment decisions, such as whether to engage in a training program, develop better production equipment, or invest R&D expenditures into new products or processes. Here, transaction surplus superiority still provides critical guidance but is not dispositive by itself. TSS is not always worth paying for, and the level of rivalry affects the payoff. If pricing behavior is exogenous to firms' surplus creation, then more rivalry depresses investment in surplus creation. If pricing behavior depends in reasonable ways on the incentive of the firm with overall transaction surplus superiority, then for investments that can turn a firm from a follower or a slight leader into a dominating leader, more rivalry tends to stimulate investment. Smaller-impact investments, however, are still depressed by rivalry.

There is much room for further research. Most obviously, the single-segment analyses presented here should be extended to cover multi-segment competition. A formal statement of when it is better to emphasize mean advantage across segments (being pretty strong everywhere) versus variance (being the leader in some segments and a laggard in others) would be a main goal of that extension. A key modeling choice in the multi-segment analysis is how much price discrimination is allowed across segments. Departing from the canonical segment case to cover second-stage games with capacity limits or with increasing returns to scale is another area for extension.

Empirical or experimental tests of the price leadership principles proposed here also present a potentially fruitful area for research. It might be possible to tell if bigger differences in rivals' surplus creation per unit are correlated with smaller average price-cost margins or if efficiency-enhancing mergers by trailing firms raise or lower price-cost margins. Experimentally, it would be interesting to see to what extent (and under what conditions) varying offer-matching speed, initial offers, number of rivals, and leader transaction surplus superiority induce leaders to follow the weak or strong leadership principles.

ACKNOWLEDGEMENTS

I would also like to thank the editor and reviewers for numerous helpful suggestions. This work has benefitted from conversations over the years with Christian Asmussen, David Besanko, Edward Fox, Javier Gimeno, David Hoopes, Marvin Lieberman, Richard Rumelt, Gordon Walker, and Peter Zemsky. In addition, I received useful insights from seminar participants at Copenhagen Business School, the Cox School of Business at Southern Methodist University, the Anderson School of Management at University of California at Los Angeles, and panelists at the Academy of Management and Strategic Management Society annual meetings. Finally, I owe a debt to the many MBA students over the years whose questions helped me identify weak spots in the framework and improve the exposition of (simpler versions of) these ideas.

REFERENCES

Adner, R., & Zemsky, P. (2006). A demand-based perspective on sustainable competitive advantage. *Strategic Management Journal*, *27*, 215–239.

Anderson, R. (1984). *Quick response equilibria*. Berkeley, CA: University of California C.R.M. I.P No. 323.

Arrow, K. (1962). Economic welfare and the allocation of resources to inventions. In R. Nelson (Ed.), *The rate and direction of inventive activity* (pp. 609-626). Princeton, NJ: Princeton University Press.

Asmussen, C. (2015). Strategic factor markets, scale-free resources, and economic performance: The impact of product market rivalry. *Strategic Management Journal*, *36*(12), 1826–1844.

Barney, J. (1991). Firm resources and sustained competitive advantage. *Journal of Management*, *17*(1), 99–120.

Berry, S. (1994). Estimating discrete-choice models of product differentiation. *RAND Journal of Economics*, *25*, 242–262.
Besanko, D., Dranove, D., & Shanley, M. (2000). *Economics of strategy* (2nd ed.). New York, NY: Wiley.
Besanko, D., Gupta, S., & Jain, D. (1998). Logit demand estimation under competitive pricing behavior: An equilibrium framework. *Management Science*, *44*, 1533–1547.
Best, J. 2016. Did you even notice your Kraft mac & cheese has no artificial ingredients? Takepart. Retrieved from http://www.takepart.com/article/2016/03/08/kraft-mac-and-cheese-artificial-ingredients
Best, R. (2000). *Market-based management: Strategies for growing customer value and profitability* (2nd ed.). Saddle River, NJ: Prentice-Hall.
Bhaskar, V. (2007). *The kinked demand curve* (Working paper). University College London. Retrieved from http://www.ucl.ac.uk/~uctpvbh/kinked.pdf
Brandenburger, A., & Stuart, H. (1996). Value-based business strategy. *Journal of Economics, Management and Strategy*, *5*, 5–24.
Brandenburger, A., & Stuart, H. (2007). Biform games. *Management Science*, *53*(4), 537–549.
Chatain, O. (2014). How do strategic factor markets respond to rivalry in the product market? *Strategic Management Journal*, *35*(13), 1952–1971.
Costa, L. A., Cool, K., & Dierickx, I. (2013). The competitive implications of the deployment of unique resources. *Strategic Management Journal*, *34*(4), 445–463.
Fisher, F. (1989). Games economists play: A noncooperative view. *RAND Journal of Economics*, *20*(1), 113–124.
Ghemawat, P. (1991). *Commitment: The dynamic of strategy*. New York, NY: Free Press.
Ghemawat, P. (2006). *Strategy and the business landscape* (2nd ed.). Reading, MA: Addison Wesley.
Ghemawat, P., & Rivkin, J. (2014). *Competitive advantage. Core curriculum reading 8105*. Cambridge, MA: Harvard Business School Press.
Ghiselin, M. (1989). *Intellectual compromise: The bottom line*. St. Paul, MN: Paragon House.
Hitt, M., Ireland, R., & Hoskisson, R. (2007). *Strategic management, competitiveness and globalization: Concepts and cases* (7th ed.). Mason, OH: Thomson-Southwestern.
Hoopes, D., Madsen, T., & Walker, G. (2003). Why is there a resource-based view? Toward a theory of competitive heterogeneity. *Strategic Management Journal*, *24*, 889–902.
Jensen, M. (2001). Value maximization, stakeholder theory, and the corporate objective function. *Journal of Applied Corporate Finance*, *14*, 8–21.
Lieberman, M. (2010). *Competitive advantage: Theoretical challenges and insights from value-price-cost approaches.* Presentation at the Academy of Management Annual Meeting, Montreal, Session 348.
Lippman, S., & Rumelt, R. (1982). Uncertain imitability: An analysis of interfirm differences in efficiency under competition. *Bell Journal of Economics*, *13*, 418–438.
MacDonald, G., & Ryall, M. (2004). How do value creation and competition determine whether a firm appropriates value? *Management Science*, *50*, 1319–1333.
Makadok, R. (2010). The interaction effect of rivalry restraint and competitive advantage on profit: Why the whole is less than the sum of the parts. *Management Science*, *56*, 356–372.
Makadok, R. (2011). The four theories of profit and their joint effects. *Journal of Management*, *37*, 1316–1334.
McGrath, R. (2013). *The end of competitive advantage: How to keep your strategy moving as fast as your business*. Cambridge, MA: Harvard Business Review Press.
Oxley, J., Rivkin, J., & Ryall, M. (2010). The strategy research initiative: Recognizing and encouraging high-quality research in strategy. *Strategic Organization*, *8*, 377–386.
Peteraf, M., & Barney, J. (2003). Unraveling the resource-based tangle. *Managerial and Decision Economics*, *24*, 309–323.
Popper, K. (1966). *The open society and its enemies, Volume 2 The high tide of prophecy: Hegel, Marx, and the Aftermath* (5th ed.). Princeton, NJ: Princeton University Press.
Porter, M. (1985). *Competitive advantage*. New York, NY: Free Press.
Radner, R. (2003). Viscous demand. *Journal of Economic Theory*, *112*, 189–231.
Reinganum, J. (1984). Practical implications of game theoretic models of R&D. *American Economic Review Papers and Proceedings*, *74*, 61–66.
Rogerson, W. (1997). Intertemporal cost allocation and managerial investment incentives: A theory explaining the use of economic value added as a performance measure. *Journal of Political Economy*, *105*, 770–795.
Shaked, I., Michel, A., & Leroy, P. (1997, October 1). *Creating value through EVA: Myth or reality? Strategy + Business, Fourth Quarter.* Retrieved from http://www.strategy-business.com/article/12756
Siggelkow, N., & Wibbens, P. D. (2015). *A ladder of competitive advantage*. Academy of Management Proceedings, *2015*(1), 10511.
Simon, H. (1996). *The sciences of the artificial* (3rd ed.). Cambridge, MA: MIT Press.
Stuart, H. (2015). Value gaps and profitability. *Strategy Science*, *1*, 56–70.
Taylor, A. (2014, July 24). Ford's epic gamble: The inside story. *Fortune.* Retrieved from http://fortune.com/2014/07/24/f-150-fords-epic-gamble/
Tirole, J. (1988). *The theory of industrial organization*. Cambridge, MA: MIT Press.
Train, K. (2003). *Discrete choice methods with simulation*. Cambridge, UK: Cambridge University Press.

Wertenbroch, K., & Skiera, B. (2002). Measuring consumers' willingness to pay at the point of purchase. *Journal of Marketing Research*, *39*, 228–241.

Williamson, O. (1991). Strategizing, economizing, and economic organization. *Strategic Management Journal*, *12*, 75–94.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Postrel S. Transaction surplus superiority in canonical market segments: Using the profit map to guide positioning and investment choices across price-rivalry regimes. *Strat Mgmt J*. 2018;39:1573–1602. https://doi.org/10.1002/smj.2769