2017

Value Capture Theory: A Strategic Management Review

Joshua S. Gans
Michael D Ryall

Strategic Management Journal
Strat. Mgmt. J., **38**: 17–41 (2017)
Published online EarlyView in Wiley Online Library (wileyonlinelibrary.com) DOI: 10.1002/smj.2592
Received 30 April 2014; *Final revision received 27 February 2016*

VALUE CAPTURE THEORY: A STRATEGIC MANAGEMENT REVIEW

JOSHUA GANS and MICHAEL D. RYALL*
Rotman School of Management, University of Toronto, Toronto, Ontario, Canada

Research summary: *This article provides the first review of a growing line of scholarly work in strategy that we refer to as "value capture theory." The common thread in this work is its use of cooperative game theory as a general, mathematical foundation upon which to build a deep understanding of firm performance in market settings. Our review: (1) describes the primary elements of the theory; (2) highlights important blindspots that it resolves with respect to existing theoretical approaches; (3) calls attention to several of its novel insights; and (4) summarizes a myriad of applications and empirical studies that have appeared in recent years using value capture theory.*

Managerial summary: *Traditionally, theoretical claims in strategic management have been supported by informal, qualitative reasoning. Recently, however, a new line of theoretical work based upon mathematical methods, known as "value capture theory," has been gaining in popularity. This article reviews the recent advances in this line with a particular emphasis upon a number of its important insights, several of which challenge longstanding propositions from the traditional line. For managers, the formal nature of value capture theory is well-aligned with data-driven analyses of strategic situations.*

INTRODUCTION

Understanding persistent heterogeneity in firm performance is, perhaps, *the* central objective in the field of strategy.[1] The seminal contribution of Brandenburger and Stuart (1996) initiated a unique stream of work designed to deepen our understanding of this phenomenon through the development of a mathematical theory of value creation and capture under competition. Since then, this literature has quietly and steadily grown to the point where it now represents a substantial body of work. Due to its mathematical nature, the collection of findings build upon each other, thereby creating a coherent, interlocking set of theoretical claims. These claims reveal subtleties of competition that were not previously apparent, pushing strategy scholars to rethink some fundamental ideas about value capture under competition. What is more, recent advances in empirical methods have opened the door to empirical investigation of these claims. Finally, it is important to note that the theory of interest is entirely published in strategic management journals. This has not been an exercise of locating technical results outside the field and importing them into strategy via reinterpretation and analogy. Rather, this work contains novel propositions of its own, with

Keywords: value creation; value capture; competition; cooperative game theory
*Correspondence to: Michael D. Ryall, 105 St. George St., Toronto, ON M5S 3E6, Canada. E-mail: mike.ryall@rotman.utoronto.ca

[1] As evidence for this assertion, we note that an adequate review of the existing empirical literature on this topic would require a separate article of its own. For example, a partial list of relevant empirical analyses on the extent to which above-average profits persist includes: Cubbin and Geroski (1987), Dosi, Lechevalier, and Secchi (2010), Jacobsen (1988), Knott (2003), McGahan and Porter (2003), Mueller (1977, 1986), Roberts (2001), Rumelt (1991), Waring (1996), Wiggins and Ruefli (2002), and Madsen and Walker (2015).

implications extending well beyond the boundaries of strategy.

Thus, we presently find ourselves in a situation in which a stream of literature has developed to the point that a general, scholarly audience is likely to find the insights it offers of substantial interest, yet it is early enough in the process that these insights are not broadly disseminated. In other words, the time seems just right for a review of this stream aimed at a general audience. The purpose of this article is to provide such a review. Our specific goals include identifying the major issues this research is intended to address, explaining the mathematical framework upon which it is built, highlighting the central theoretical insights thus far obtained, describing notable applications (both theoretical and empirical), and concluding with some thoughts about open issues and future directions.

The scope of our review includes papers (1) published in management journals that (2) present novel mathematical propositions using cooperative game theory. By adopting this unifying theme, the scope of our article is manageable, permits a coherent discussion, and allows us to treat key ideas with a reasonable degree of depth. Moreover, by restricting attention to management publications, we provide a sense of the novelty and substance of the stream within our own field.[2] Our intended audience includes those looking for an overview of the central concepts and latest results associated with this line of work as well as those considering entering and contributing to it. Therefore, this article is constructed to be self-contained within the content constraint of a single journal article—the formalism, relevant assumptions, associated interpretations, and important results to date are all covered, albeit concisely, with citations to more in-depth sources for those interested in pursuing them.

This being a literature review, it is also worth pointing out that the issues elaborated below were the product of a discovery process: theorists built models that seemed sensible; based on these models, formal propositions were proven; mathematical claims led to insights about value capture under competition in the real world; these insights led to a grasp of the kinds of issues that the dominant paradigms missed. This article proceeds in reverse order: we summarize the (now) apparent issues; then, we present the formalism and trace the line of results derived from it; this should spark insights in the reader regarding value capture under competition; which, cumulatively, should lead to a deeper understanding not only of this line of work but also to new questions for future inquiry.

A MOTIVATING EXAMPLE

To provide context for our survey, it is useful to begin not with the literature but with an example that motivates value capture theory and the contributions it can make. The example is intended to be illustrative rather than the model of a specific strategic situation.

Status quo configuration

Any formal model of an industry usually starts with a benchmark or what we will call a status quo configuration upon which changes can be compared. For our example, suppose there is an industry that consists of value networks composed of some finite numbers of agents. By "value network" we mean a collection of agents connected to one another via chains of transactions that, taken together, ultimately result in the production of economic value. The network includes all agents involved in the production of value, from the most upstream resource providers all the way down through the final consumer. An example of a value network would be the agents engaged in the Apple mobile computing ecosystem, including those who provide platform elements, mobile carriers who provide access to communications and the Internet, app developers, the content providers and the consumers and end-users themselves. The Apple value network competes with others built around Samsung, Google, etc.

Suppose there are exactly two such networks, **A** and **B**. Further, assume that—adding up the utility enjoyed by end-users and subtracting all of the economic costs associated with creating and delivering the products to those end-users (including costs relating to the transactions themselves)—Network **A** creates economic value of \$200 million and Network **B** creates economic value of \$180 million. This is the situation depicted in the first panel of Figure 1 (Configuration #1).

[2] Unfortunately, this scope does imply, on the one hand, passing over a number of thought-provoking, qualitative papers in strategy that adopt ideas from cooperative game theory and, on the other hand, ignoring potentially interesting papers published in other fields (most notably, economics). Stuart (2001) provides an early precursor to this article.

DOI: 10.1002/smj

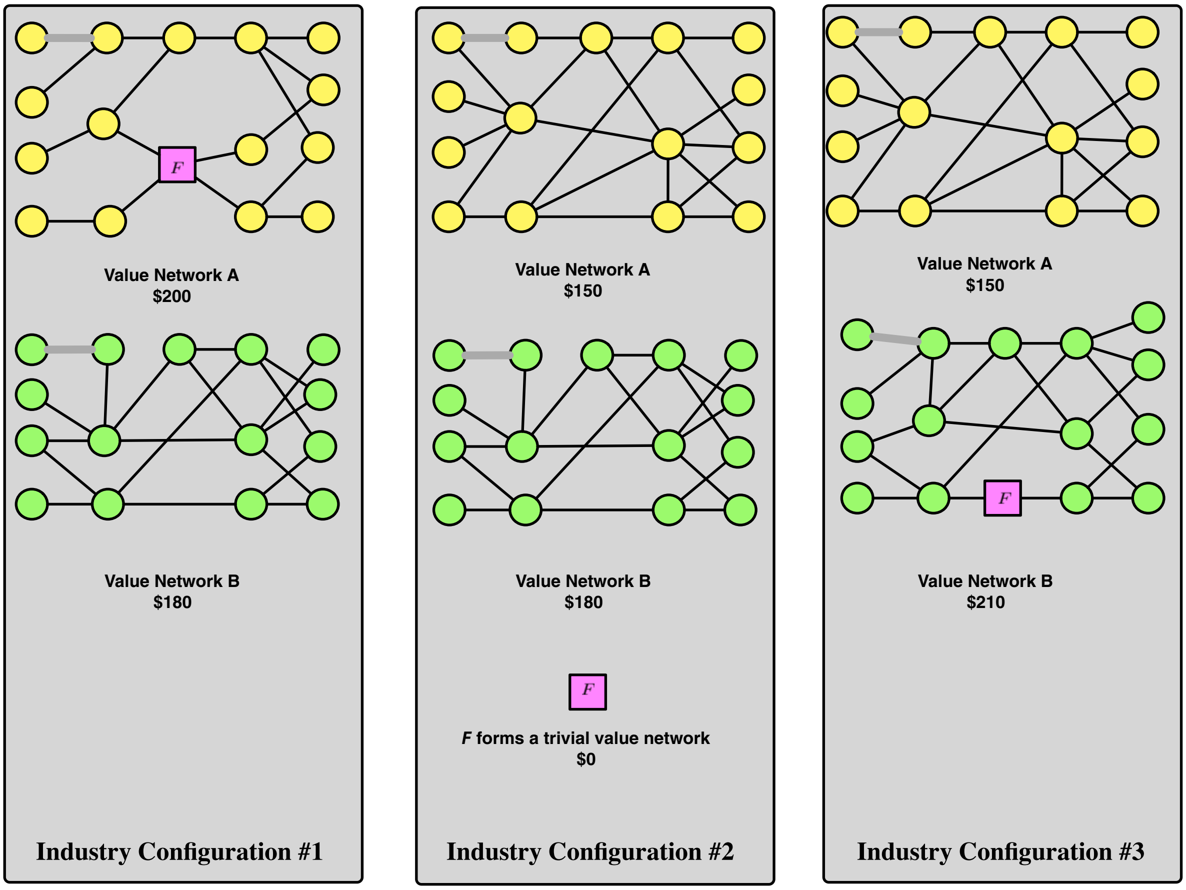

Figure 1. Three industry configurations

Assume this configuration is actually the one observed operating in the world and, hence, refer to it as the "Status Quo." In aggregate, the industry produces \$380 million in economic value. Finally, suppose our interest is in understanding the performance of a focal firm F, which is presently operating in Network **A**. For instance, to continue our mobile example, F may be Adobe, which provides specialized Flash software.

Although this setup is "simple" in the sense that it contains only two value networks, it is worth pointing out a number of simplifying assumptions that were *not* imposed. We did not limit the number of agents (beyond being finite), exclude transactions costs, restrict consideration to two-sided markets, exclude buyers, represent buyers as a linear demand curve, impose special functional forms on production functions, exclude positive externalities, and so on. This level of generality is a feature of the mathematical approach we describe in more detail below.

Competitive upper bound

In the second panel of Figure 1 (Configuration #2), we indicate the chains of transactions that would arise were *all* agents in the market to *forgo transactions with* F; for instance, deciding not to support Adobe's software. In our example, the value produced by Network **A** drops to \$150 million, while Network **B** continues to create \$180 million in value. Of course, F can create no value on its own. In this case, the aggregate value produced in the industry would drop from \$180 to \$330 million, a decrease of \$50 million. This difference—the incremental value produced as a result of F's presence in the market—is called its *added value* (to the market as a whole), symbolically, $av_F = 50$. It is easy to see that, in a noncoercive market setting, whatever amount of the \$380 million F ultimately captures, it cannot be more than \$50 million. Why? Because were F to insist upon a share greater than this, the other agents in the market could implement

Configuration #2, which would yield them more value than producing $380 million and giving *F* more than $50 million.

By now, the notion of added-value and its role as a limiter of value capture is well known within strategy (see, Brandenburger and Nalebuff, 1996; Gans, 2005). Even so, there are a few points worth highlighting. First, as shown in the figure, Configuration #2 contemplates a different collection of transactions to those at work in Configuration #1. Computing added value goes beyond the simple removal of transactions with a given agent, to include a complete assessment of the new collection of transactions that would actually arise in their place. Second, it illustrates a general feature of added value. An agent's added value can be computed with respect to any group (the quantity of value produced when the group includes that agent less the quantity produced without it. When the added value being computed is with respect to a status quo value network that *includes* the agent, that added value represents an *upper bound* on its ability to capture value. As we see from the figure, *F*'s added values to Network **A** and to the market as a whole coincide at $50.[3]

Finally, added value provides an important insight into the nature of competition. Competition in markets means vying for transaction partners. To engage in the transactions that produce $380 million, other transactions must be forgone. In particular, in order to ensure that the transactions depicted in Configuration #1 actually arise, *F* must offer enough value to all of the agents with whom it transacts to persuade them not to engage in the transactions depicted in Configuration #2. The latter transactions provide *F*'s transaction partners with a valuable alternative to dealing with *F*—an alternative against which *F* must, in effect, compete.

Competitive lower bound

Now consider the third panel of Figure 1, Configuration #3. Here, *F* deals with Network **B**, with a resultant increase in the value *it* produces from $180 to $210 million. Thus, *F*'s added value with respect to Network **B** is $30 million. Note the implication: the agents in Network **B** are willing to give *F* a share of up to $30 million (at which point they are indifferent to *F*'s inclusion) in order to effect the transactions depicted in Configuration #3. This illustrates an often-overlooked aspect of competition: *F*'s transaction partners must offer it enough value to ensure that *F* remains in Network **A** rather than abandoning *them* to transact with Network **B**. How much must they offer *F* to remain? At least $30 million.

It is important to observe that the focus here is entirely on the quantity of economic value a particular agent captures in return for the part it plays in the aggregate creation of value within the industry. Thus, there is no discussion of prices or margins or even a description of the products that generate revenue or induce costs. *Specifying prices, quantities, and costs is not the goal of this sort of analysis*.[4] Rather, this method places the productive powers of agents, the competition between them to transact with one another, and the implications for value capture front-and-center.

Notice that competition operates symmetrically. To induce the other agents in Network **A** to forgo alternative deals in favor of transacting with it, *F* must offer them a sufficient share of the value created by their joint economic activities. However, those very same transaction partners must, simultaneously, offer *F* a sufficient quantity of value to ensure that it does not act upon the alternative transactions that it has available. Here we see a flip side to added value: when the added value being computed is with respect to a status quo value network that *does not include* the agent, that added value represents a *lower bound* on the value the agent must capture. Thus, *competition need not hurt* an agent—it may play out in its favor. Value capture theory sorts out both sides of competition.

Thus, without imposing too many restrictions (and computing only four numbers), we expect that the quantity of value captured by *F* will be between $30 and $50 million. The former is a floor imposed by competition for *F*, and the latter is a ceiling imposed by competition for *F*'s transaction partners. The analytic work in the literature we review follows this thrust. In general, there are many configurations to consider, many ways to add interesting structure, and many insights available by so doing. Let us consider some now.

---

[3] The two need not be equal, though an agent's added value to the market as a whole is never greater than its added value to a distinct value network.

[4] A feature of cooperative game theory is consistent with the call by Wernerfelt (1984) and noted by Lippman and Rumelt (2003).

DOI: 10.1002/smj

SOME BLIND SPOTS IN PREVAILING THEORY

The modern era of strategy scholarship dates to the famous collection of papers appearing in the 1980s, all of which offer explanations for the phenomenon of persistent performance heterogeneity among firms. These include Porter (1979, 1980), which initiated the "industry positioning" stream (IP); Wernerfelt (1984), the paper that sparked the "resource-based view" line of work (RBV); and Williamson (1981), an early paper advocating his "transactions cost" approach to organizational analysis (TCE).[5] In the years following, strategy scholars expanded these original ideas in a number of fruitful directions, thereby forming a commonly accepted collection of explanations for persistent differences in firm performance. Our goal in this section is to identify several substantive oversights, or "blind spots," shared by these mainstream explanations. Blind spots are eliminated by turning attention toward them. By so doing, we see the issues surrounding our extant body of theory more completely which, in turn, opens the door to greater generalization and unification.

Of more immediate relevance for our purposes is that the blind spots discussed below were revealed by the methods reviewed by this article. Thus, our presentation reverses the historical process of analysis-reflection-insight by which these blind spots were revealed. Instead, we begin with the blind spots, explained in plain language. Our hope is that, upon reflection, these explanations will trigger insights into both the existence and potential importance of the identified oversights. Then, once we do elaborate the technical details of the method, our further hope is that these insights will carry over, allowing readers to grasp how the mathematics of value capture reveal the identified blind spots in an explicit fashion.

Before proceeding, we warn that any generalization of large, complex streams of scholarly work will, by its nature, blur their finer and more subtle points. Our intent is not to gloss over these subtleties. Rather, we wish to provide a persuasive illustration of the way in which different approaches (in this case, mathematical) can expand our explanatory horizons. Many scholars have weighed in on both sides of the argument regarding the efficacy of mathematical theory in management research. Again, our purpose here is not to engage those arguments. Rather, we simply point out that, for better or worse (depending upon one's philosophical sensibilities), a unique feature distinguishing the value capture stream from others in strategy is that its claims are formulated using a single, overarching mathematical framework. The result is a collection of rigorously derived findings that build upon one another to create a coherent whole, one with the potential to unify and refine prevailing theory.[6]

The free-entry blind spot

Because strategy is about the performance of firms in free markets, competition is always a central concern. In proposing their explanations for firm performance heterogeneity, traditional theories share the premise (often implicit) that wherever positive profits are to be found, competitors are attracted like a swarm of locusts to a luscious crop of wheat. Thus, "competition" is typically conceived of as an ever-present, corrosive force directed toward those enjoying positive economic profits.[7] The entrants-as-a-swarm-of-locusts conception is one inherited from neoclassical economics. As Makowski and Ostroy (2001) explain (p. 484): "Neoclassical economists borrowed from their classical predecessors the view that, in a production economy, perfect competition is the simple, inescapable conclusion of free entry. And with free entry comes zero profits." This notion of how competition works continues to run deep, both in strategy and beyond.[8]

The founders of modern strategy were interested in explaining what seemed obvious by casual empiricism—that persistent performance differences within industries was more the rule than the exception. Indeed, strategy presently sports a remarkable body of empirical work directly at odds with the claim that competition inexorably drives

[5] The rapid coalescence of these ideas into a separate field, business strategy, was surely facilitated by the disinterest toward this phenomenon apparent in economics at that time. Economics has since come around; see, e.g., Syverson (2011) and Bloom *et al.* (2013).

[6] In this sense, the work surveyed here answers the challenge posed by Oxley *et al.* (2010).

[7] See, for example, the comments by Porter (1980: 5 and elsewhere).

[8] See, e.g., any strategy textbook for first-year MBAs or, in economics, historic work such as Schumpeter (1934) to more modern pieces like Hopenhayn (1992).

DOI: 10.1002/smj

industry profits to zero.[9] However, to supervene this claim, one must find a way to escape the neoclassical logic. One reasonable escape path is to propose a class of *barriers* thought to interrupt the progression induced by free entry toward zero profits. This is precisely the approach adopted by each of the major schools: IP proposes barriers to industry mobility, the RBV proposes barriers to resource mobility, and TCE proposes barriers to perfect information. It is important to note that, even though it avoids the essential conclusion of the standard neoclassical model, this "barrier view" retains its essential premise—that entry is an ever-present threat whenever positive profits are present.[10]

The challenge we now pose is: Why not discard the neoclassical premise altogether? Suppose the world's stock of economically productive resources (raw materials, capital, human beings, time, etc.) is finite. Of all the assumptions one might entertain, the finitude of resources strikes us as supremely reasonable—an immediate, universal feature of our shared economic experience. Yet, this axiomatic premise has serious implications. In such a world, logic does *not* dictate that the moment some resource becomes a source positive value capture for its owner, other agents will reallocate their own resources against it in a tsunami of competition. Rather, economic logic implies that agents allocate *away* from less profitable to more profitable settings, and resource finitude implies that such settings need not be exhausted—*in a finite world, economic profit may well be the rule rather than the exception*.

The importance of this point cannot be overemphasized in the context of strategy: global resource finitude implies that special barriers to competition are neither necessary nor sufficient for the existence of persistent economic profit. To be sure, there may be settings in which such barriers are a proximate cause of profit persistence.[11] That said, a body of theory focused only upon this aspect of competition must, in itself, be incomplete. This carries with it the empirical implication that variation in the strength of special barriers is unlikely to be a robust predictor of variation in firm performance. As we show below, the finite agent/resource assumption is explicitly embedded in the value capture model, even when scaling up to global-economy-level scope.

The competitive-determinism blind spot

Theorists have an aversion to models prone to ambiguous conclusions. For example, the two market models with which most strategy scholars are familiar are Cournot and Bertrand, the workhorse tools of industrial organization economics. No small part of the popularity of these models arises from their tractability, in particular, their ability to provide exact, point-estimate solutions for profits once their parameters are fixed. In what follows, we show that the value capture model suggests that competition is properly construed as placing *bounds* on the amount of value an agent may capture without fully determining it. Taking the premises of the model seriously, this is a "feature not a bug" (in the parlance of our time). It says that, in general, competition defines a precise *interval* within which an agent's value capture lies; where within that interval actual capture lands is due to factors other than competition.[12] As we discuss in greater detail below, the theory points toward a new conception of "competitive intensity," as well as the existence and possible importance of "persuasive" resources, the express purpose of which is the capture of more value *within* the range permitted by competition.[13]

The product-price blind spot

Harkening back to modern strategy's founding corpus, Wernerfelt (1984) argues that, in order

[9] To cite several compelling instances: Cubbin and Geroski (1987), Dosi (2007), Jacobsen (1988), Knott (2003), Madsen and Leiblein (2015), McGahan and Porter (1999), Mueller (1977, 1986), Roberts (2001), and Villalonga (2004).

[10] Gans, MacDonald, and Ryall (2008) elaborate on this idea. As an anonymous reviewer pointed out, the free-entry, "swarm of locusts" idea appears most prominently and explicitly in RBV scholarship.

[11] By "special" we mean explicit barriers to the allocation or reallocation of economic resources.

[12] On the issue of generality, it is worth reminding readers of Kreps and Scheinkman (1983), who illustrate the sense in which the Cournot results are consistent with a Bertrand game with capacity commitments. Similarly, Stuart (2005) shows that Cournot can also be thought of as a special case of a cooperative game with capacity precommitments. In other words, both Cournot and Bertrand can be thought of as special cases of a cooperative game. A related line of research is pursued in Byford (2010).

[13] Here, it is useful to note that empirical research in strategy indicates that the firm effect is significant in explaining performance heterogeneity (see Dosi, 2007, for a review). The possible importance of a "persuasive" category of resources opens up a new area of exploration for empirical scholars interested in explaining these effects (e.g., in the spirit of Roberts and Dowling, 2002; Villalonga, 2004).

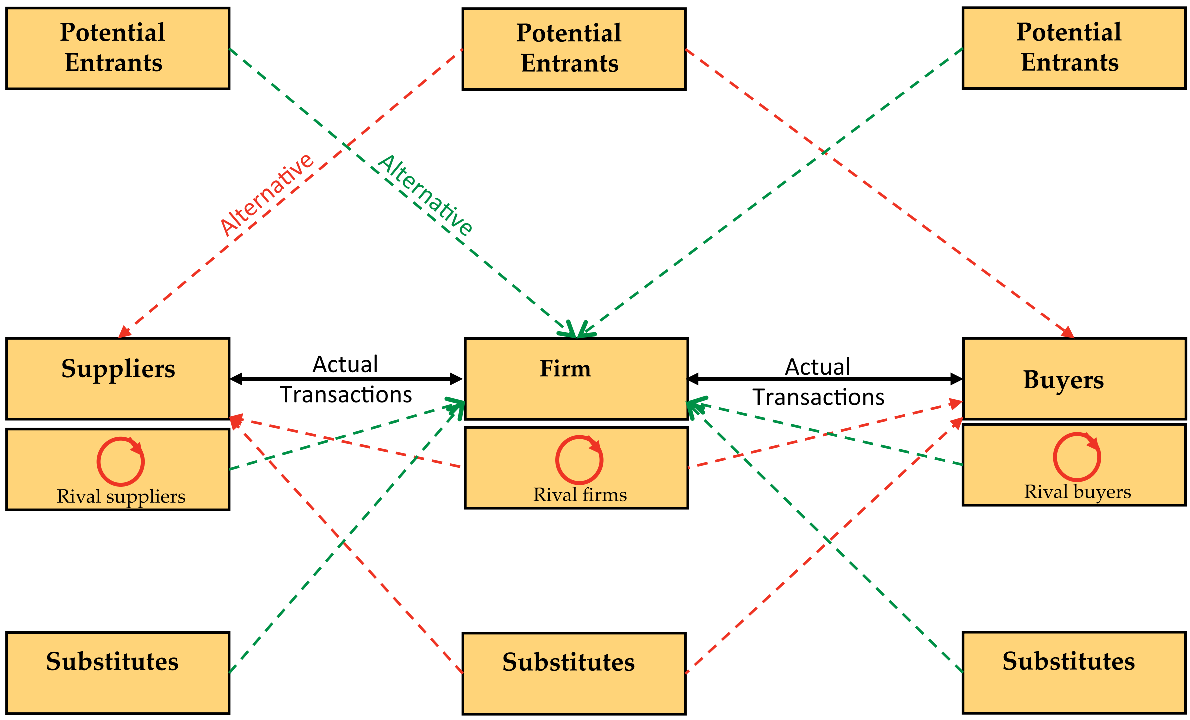

Figure 2. A balanced, but tangled, view of the one "force" of competition

to develop a deeper understanding of persistent performance heterogeneity, analysis must shift from the pervasive price-product orientation of industrial organization economics to a more strategy-useful agent-profit perspective. The massive, resource-based view literature, which arose from this observation, extends the issue by positing that persistent performance heterogeneity is due to the existence of unpriced resources (a result of "market failure"). The blind spot here is not in arguing that the product-price unit of analysis is problematic but, instead, in not taking the agent-appropriation perspective sufficiently seriously. When the agent is the unit of analysis, the quantity of value it captures is, in essence, a price. It is the price of engaging that agent—including its resources and capabilities, themselves priced or not—in the activities required to produce value. Market failure at one level does not imply failure at the other. The value capture model shifts away from products and prices toward agents, the myriad value creation opportunities available to them, the value actually produced, and what it takes in terms of value capture to engage them in the associated productive activities.

The web-of-transactions blind spot

The central insight of Porter (1979, 1980) is that understanding competition requires a broad view of who counts as a competitor. In his famous "Five Forces" model, Porter (1979) argues that the set of active market participants—those agents whose decisions are aimed at appropriating a share of the value produced within an industry—must include buyers, suppliers, rivals, even *potential* rivals and the producers of substitute goods outside the industry. Extending this insight to its logical conclusion, Porter's (1979) model itself is incomplete: a firm's appropriation ultimately depends upon the Five Forces groups, as well as *suppliers of suppliers*, *substitute buyers in other markets*, *potential entrants into distribution channels*, and so on and so on. One gets a sense of the larger picture from Figure 2, though this too falls short. In any given industry, value is created and appropriated through complex webs of transactions pursued by multiple layers of active, intelligent agents. Returning attention to market models like Cournot and Bertrand, the problem is that the only active agents (i.e., the only agents whose decisions are taken into account) are typically the firm and its direct rivals. They too are incomplete, often making implicit, ad hoc assumptions that leave price-setting power in the hands of the firms, with all other market participants playing a passive role.

THE BIFORM MODEL

The biform model of Brandenburger and Stuart (2007) is presently the state-of-the-art model for analytical work in strategy on value capture under competition. The "biform" model is so-called because it synthesizes two distinct areas of game theory—noncooperative (NGT) and cooperative (CGT) —into a single model. The great strength of NGT is its ability to capture strategic behavior; i.e., situations in which agent actions interact in the generation of outcomes and, therefore, in which assessing the behavior of others is important. The power of CGT is in its ability to analyze value creation and capture in markets, especially in settings where agents' dealings do not follow some predefined process. In a biform model, each type of theory plays a role consistent with its strengths, thereby creating a complementary whole.

Biform models consist of two stages. The first stage, NGT, includes all the agents whose actions have a significant impact upon the outcomes for a focal firm. In an industry, this typically includes direct rivals as well as, e.g., the agents in the other categories of Porter's Five Forces framework (buyers, suppliers, etc.). In the NGT stage, agents take actions designed to "prepare the competitive battlefield" to their advantage—that is, to structure their competitive environments in a way that permits them to capture the most value possible. Such actions include initiatives like marketing projects, capacity decisions, new product or technology introductions, mergers, recruiting policies, market entry, and so on. The joint actions taken by the agents in the first stage induce a particular market environment in the second stage, CGT. Here, agents engage in organizing and executing the "free-form" deals that result in the production and capture of value. Essentially, the value captured in second-stage competition serves as the the payoff to the strategic action taken in the first. We begin by elaborating the second (cooperative game) stage.

COOPERATIVE GAME STAGE

A *cooperative game* consists of a pair (N, v) where $N \equiv \{1, \ldots, n\}$ indexes the $n < \infty$ *agents* participating in the market, and v, referred to as the *characteristic function*, is a map from each subset of agents to the quantity of economic value producible by the agents within it. The scope of this setup is broad: it can be used for small interactions (e.g., n equal two to a few, such as firms competing for an acquisition or contemplating a strategic alliance for technology development), all the way up to those involving large numbers of agents (e.g., the global economy). Given a group of agents $G \subseteq N$, $v(G)$ denotes the amount of value the agents in G can create on their own (i.e., when the members of G limit their transactions to one another only). Thus, the starting point for the CGT stage is an elaboration of how much value each of the various groups of agents could create in this way were they to decide to do so. Out of all the possibilities so elaborated, some deals do occur, some value is produced, and some share of it (possibly zero) is captured by each of the agents in N. A *distribution of value*, denoted $\pi \equiv (\pi_1, \ldots, \pi_n)$, is a list indicating how much each agent captures in return for its productive activities (i.e., agent i captures π_i).

Again, the central idea behind this setup is to show how the feasible productive opportunities in a market shape a firm's ability to capture value. This requires some linkage between v, which describes the former, and π, which describes the latter. The first step is to determine how much aggregate value is actually produced within the market. The answer is straightforward: $v(N)$ is, by definition, the aggregate value generated by the agents in N. In strategy applications, $v(N)$ represents the actual, aggregate quantity of economic value that will be created (in the case of a theoretical prediction) or was created (in the case of an empirical analysis). It is a key "observable" associated with the model. The strategy literature commonly uses the special label $V \equiv v(N)$ to emphasize this distinction. The majority of of the $v(G)$s for groups other than N never actually materialize. These unrealized possibilities *are constitutive of competition* and, as such, shape the distribution of value.[14]

[14] When the value created within a market arises as the aggregate of productive activities by distinct groups, then the economic value created by each of those particular groups is, in fact, realized.

DOI: 10.1002/smj

Note that V and the $v(G)$s are real numbers. The usual interpretation is that $v(G)$ is the economic value created by the transactions that would actually arise were transactions restricted to the agents in G. However, depending upon the application, they can also represent cash, expected value, net present value, etc. Frequently, authors assume that v is *superadditive*: the union of two disjoint groups produces at least as much as the sum of what the groups produce independently. The rationale for this is that the transactions associated with the independent group values are still feasible even though new, cross-group transactions may arise under the union of the two groups.[15]

How do the value creation opportunities shape value capture? Typically, two assumptions are made. The first is a *feasibility* assumption:

$$\sum_{i \in N} \pi_i = V, \tag{1}$$

i.e., the amount of value captured equals the amount produced. The second is a competitive *consistency* assumption: for every group $G \subset N$,

$$\sum_{i \in G} \pi_i \geq v(G), \tag{2}$$

i.e., the aggregate value captured by the agents in G must be at least as large as the value they could produce on their own. Suppose, in return for their contributions to the production of V, the agents in G face deals such that $\sum_{i \in G} \pi_i < v(G)$. Then, the agents in G could eschew those deals and, instead, undertake to produce $v(G)$ under terms that would make each and every one of them strictly better off.[16] In the strategy literature, a distribution of value that meets the feasibility and consistency conditions is said to be *competitive*. The set of all such distributions is referred to as the *core*.[17]

In what sense are distributions that meet Equations (1) and (2) "competitive"? The answer lies in thinking of V as the aggregate value arising from a *specific set of actual transactions*. Think of these transactions, the value they create, and the shares of value captured by the agents involved as observables (i.e., as data or potential data). The $v(G)$s, then, are the feasible values produceable through alternative transactions and groupings. In this context, $v(G)$ should reflect the actual economic value that would be produced via some other specific set of transactions, including any costs associated with organizing them (e.g., switching costs). Then, the alternative transactions compete with those anticipated in the production of V. Or, as viewed from the agent perspective, the members of G provide competition *for one another* and *against their respective transaction partners* in the creation of V.

To interpret competitive distributions in a strategy context, bring Figure 2 to mind. The firm engages in specific transactions with a specific network of suppliers and buyers resulting in the creation of actual economic value. At the same time, rivals, potential entrants, and producers of substitute products offer the firm's customers and suppliers productive alternatives. These agents compete against the firm and for its partners. Simultaneously, rivals, potential entrants, and substitutes operate at every level of the value chain, thereby providing competition for the firm and against its partners. The characteristic function v provides a summary of all these competitive alternatives. This leads to the following insight.

Insight 1

There is only *one* force of competition, not five or some other number. Competition is implied by a tension between having to neutralize all the competing alternatives in Equation (2) using the limited value produced in Equation (1). These conditions impose value capture consequences on every agent in the market. That is, while its value capture implications may vary from agent to agent, the set of competitive consistency conditions from which those implications arise is the same for every agent. All agents face a tension between Equations (2) and

[15] In most strategy applications, this assumption is innocuous—although it does rule out cases in which some subset of agents creates negative externalities with others (and, additionally, in which economic activities cannot be organized in such a way as to neutralize them).

[16] Some authors refer to this as the "stability" condition; the idea being that aggregate production of V is assured only if this condition is, by necessity, met.

[17] This term is imported from economics, which, like most economics terminology associated with CGT, is not especially helpful or clarifying. There are many other approaches to analyzing the solutions to cooperative games in strategy settings. For example, de Fontenay and Gans (2008) use the Shapley value, which has the feature of providing point-estimates in place of core intervals. As we explain below, however, the identification of intervals is a useful feature with respect to understanding one's strategic situation.

DOI: 10.1002/smj

(1). From the firm's point of view, the effects of this tension run in two directions: one in its favor (providing it more options), the other against it (providing others more options). Thus, the CGT conceptualization treats competition as a single force, with symmetric structure and two effects, one good and one bad, on each agent.

Insight 2

Competition *bounds* an agent's value capture possibilities. Mathematically, Equations (1) and (2) imply an *interval* of value capture possibilities for an agent. That is, for every agent i, competition determines an interval, with bounds $\pi_i^{\min} \leq \pi_i^{\max}$, such that π_i is part of a competitive distribution of value if and only if it is contained in $\left[\pi_i^{\min},\ \pi_i^{\max}\right]$. This is contrary to the point-estimate intuition inherited from familiar models like Cournot and Bertrand. Competition *for* an agent pushes up $\pi_i^{\min}$, while competition *against* it drives down $\pi_i^{\max}$. By way of analogy to bilateral trade, $\pi_i^{\max}$ is the market's "willingness-to-pay" for agent i's involvement in the creation of V. Similarly, $\pi_i^{\min}$ is, loosely, i's "willingness-to-sell" that involvement to the market. Both values are pinned down by competition—implicit in the feasible alternatives to producing V available to market participants.

Insight 3

Competitive intensity with respect to agent i should be conceptualized with respect to the length of $\left[\pi_i^{\min},\ \pi_i^{\max}\right]$. At its most intense, $\pi_i^{\min} = \pi_i^{\max}$. While the usual conception of extreme competitive intensity (i.e., "perfect competition") may hold ($\pi_i^{\min} = \pi_i^{\max} = 0$), so may competitive intensity at the other end of the spectrum ($\pi_i^{\min} = \pi_i^{\max} = V$), and everything in between (Montez, Ruiz-Aliseda, and Ryall, 2015). Thus, extreme competitive intensity is neither good nor bad per se. At the same time, the model admits situations in which $\pi_i^{\min} = 0$ and $\pi_i^{\max} = V$; i.e., competition plays *no* role in determining what the agent captures. If these distinctions are features of real-world industries, they have substantial implications for strategy scholars and practitioners alike.

Insight 4

Value capture depends upon two classes of resources—competitive and persuasive. Which is class is most important varies with competitive intensity. The value captured by firm i is the sum of its competitively guaranteed minimum plus some portion of its feasible interval, the latter obtained via extra-competitive means. Formally,

$$\pi_i = \pi_i^{\min} + \alpha_i \left(\pi_i^{\max} - \pi_i^{\min}\right), \tag{3}$$

where $\alpha_i \in [0, 1]$ is called i's *appropriation factor*. The α_i parameter summarizes the effect of all super-competitive determinants of firm i's ability to capture value—i.e., factors that induce others to part with value beyond the implications of Condition (2).[18] When competitive intensity is slack, the control of superior "persuasive resources" is the key to value capture (Ryall, 2013). From a positive perspective, the prediction here is that high-performing firms in low-intensity industries either control superior persuasive resources or enjoy institutional advantages (e.g., bidding norms) vis-à-vis the other agents. At the other end of the spectrum, when intensity is at its most extreme, persuasive resources play no role. Accordingly, empirical analyses that fail to account for variation in intensity/control of superior persuasive resources are unlikely to provide consistently good explanations of firm performance heterogeneity.[19]

NONCOOPERATIVE GAME STAGE

Ultimately, strategy scholars and practitioners alike are interested in understanding how firm strategies affect value capture. To investigate this issue, Brandenburger and Stuart (2007) propose linking the competitive (CGT) stage to a preceding strategic (NGT) stage by way of a "biform game." The goal is to admit a model in which the agents vie with one another to alter the competitive landscape (i.e., the second-stage cooperative game) to their advantages. For example, firm strategies may involve investments in cost-reducing process technologies, capacity expansion, new market development, personnel practices, attempts to adjust corporate culture

[18] When used prospectively, the appropriation factor can be thought of as a subjective estimate of an agent's own ability to persuade its transaction partners to part with value, by all means other than pointing out the implications of v.

[19] We are presently unaware of any explicit investigations into the competition/persuasion resource dichotomy. However, the theory is consistent with phenomena such as the investments in the acquisition of partner information prior to price negotiations, as described by Madsen and Leiblein (2015).

DOI: 10.1002/smj

and so on. Simultaneously, distributors may adopt new technologies, end users may hire negotiation experts, suppliers may implement new information systems, etc. All of these activities, presumably the result of each agent's business strategy, interact to determine the actual value creation opportunities available in the marketplace.

Formally, each agent $i \in N$ begins with a set of *feasible actions*, A_i, with typical element a_i, an *action*. Action sets can be finite or infinite, scaler or multidimensional, depending upon the application. Since the first stage is a noncooperative game, it can take strategic or extensive form. The latter is appropriate for investigating situations in which move timing and information issues are important. To keep it simple, we will describe the general strategic form: each agent chooses its action without knowledge of the other agents' choices (i.e., simultaneous-moves). The result is a list containing each agent's action, referred to as an *action profile* and denoted $a \equiv (a_1, \ldots, a_n)$. The set of all such profiles is denoted A.

The idea is that the joint actions of the market participants interact to influence both the characteristic function, v, as well as each agent's *appropriation factor*, $\alpha_i \in [0, 1]$, ultimately determining each agent's appropriation, Equation (3). Thus, the parameters describing value capture in the second stage can be identified by action profile superscripts. For example, two different action profiles a and a' induce cooperative games (N, v^a) and $(N, v^{a'})$ as well as appropriation factors α_i^a and $\alpha_i^{a'}$ for agent i, etc. Then, each agent's payoff in the second stage is assessed according to Equation (3). Thus, the value captured in the market stage is the payoff to an agent's action in the strategic stage, given the effects of everyone's actions. Formally, agent i's value capture given action profile a can be written

$$\pi_i(a) = \pi_i^{\min}(a) + \alpha_i^a \left(\pi_i^{\max}(a) - \pi_i^{\min}(a)\right), \quad (4)$$

where $\pi_i^{\max}(a)$ and $\pi_i^{\min}(a)$ are the bounds implied by v^a and α_i^a is the net effect on i's persuasive effectiveness implied by the actions in a.

In closing this section, it is worth pointing out that the biform setup opens strategy analysis to the entire toolbox of NGT. For example, most strategy applications to date have used Nash, or Bayesian Nash, as a solution concept in making theoretical claims in specific settings. However, those interested in strategic settings of bounded rationality might employ subjective equilibrium (Kalai and Lehrer, 1995; Ryall, 2003, for a strategy application) or ambiguity-averse Nash equilibrium (Ellsberg, 1961; Ryall, 2009, for a strategy application). On issues of persistent performance heterogeneity, the model can be expanded into a repeated game. In addition, methods used in studies of cheap talk, information asymmetry, and behavioral economics can all be adapted to this framework. All of these powerful modeling techniques can be brought to bear to examine how firms in a wide variety of settings behave and how those behaviors affect value capture under competition.

For example, consider application of subjective equilibrium in a simple setting. A subjective equilibrium arises when (1) each agent chooses a strategy optimizing its expected payoffs given its subjective beliefs and (2) actual expected payoffs induced by these strategy choices are consistent with subjective expectations. In other words, the consequences of agent actions are consistent with their subjective beliefs along the path of play—but may be wrong in critically important ways off the equilibrium path.

To see this, suppose there are n symmetric firms facing n buyers. Each firm can produce one unit of product at cost c. Buyers value one unit of product at $u = x + c$ where $x > 0$. Here, each firm is essentially in a pure bargaining contest with its buyer: any split of x between the buyer and seller is possible. Assume the firms begin with $\alpha_i = 0$. However, in the NGT stage, each firm chooses either to invest in bargaining training or not. The cost of the training is $0 < z < x$, the consequence of which is to shift the firm' α_i from 0 to 1. The unique Nash equilibrium for this game is for every firm to invest in the persuasive resource and receive a CGT-stage payoff of $(x - z) > 0$. Suppose, however, that the firms believe (incorrectly) that they are in a traditional "perfectly competitive" market—that is, in which $\pi_i^{\max} = \pi_i^{\min} = 0$. In such a situation, persuasive resources are useless. Based upon this belief, the subjectively optimal choice is not to invest in bargaining. Then, in the CGT stage, $\alpha_i = 0$ for all firms and, as a result, actual value captured is $\pi_i = 0$. Of course, this is exactly as expected. This is a subjective equilibrium: poor performance is attributed to tough competition when, in fact, it is due to zero investment in persuasive resources—and no investment will occur as long as markets are believed to be tough.

COMPARISON TO OTHER MODELING APPROACHES

Many readers for whom the preceding formalism is novel will already be quite familiar with the standard models of NGT (e.g., Cournot, Bertrand, Stackelberg, and variations). These models are workhorse theoretical tools in industrial organization economics. Thus, a perfectly legitimate question is: Why should strategy go in the direction of biform games rather than follow the standard adopted by economics? Essentially, this article taken as a whole is intended to provide a compelling answer to that question. Even so, this is a good point at which to provide a couple of more specific answers to it.

First, optimal model selection is highly context-dependent—which approach is "best" depends upon the questions being asked. That said, our hope is that the preceding discussion illustrates why the biform model is ideal when the questions being asked are general ones about value creation and capture in market settings. In the standard market models of NGT, a raft of assumptions must be made, including whether managers set prices or quantities, the precise order of moves, what managers know when they move, specific functional forms for costs and demand, etc. If these details are available to the investigator, great. Most of the time, however, economic activities in real markets often involve messy, arm's-length wheelings and dealings that are incompatible with this minute level of specificity. The CGT framework, in contrast, invokes relatively weak premises. These grant the analyst considerable freedom of interpretation with respect to the details of value creation being represented by the formalism. As a result, the findings derived via CGT tend to be applicable to a broad spectrum of settings of interest to strategy scholars.

Second, there are instances of papers that provide formal translations between the various standard models used to analyze competition. The most famous of these is Kreps and Scheinkman (1983). In their model, firms choose capacities ahead of a Bertrand price-setting game. The main result is that, under certain rationing rules, the equilibrium of the game gives rise to the same equilibrium result in prices and quantities as would arise under a suitably-chosen Cournot game. In other words, they show the sense in which Cournot can be thought of as a special case of a Bertrand model with capacity choices. Stuart (2005) demonstrates that Cournot can also be thought of as a special case of a biform game with capacity choices. The mapping from simultaneous-move Bertrand games to CGT games that result in equivalent value-capture outcomes is relatively trivial. Hence, Cournot and Bertrand are often special cases of a more general cooperative game.[20]

Choice of solution concept is another issue. CGT sports a number of different solution concepts—the core, stable set, strong epsilon-core, Shapley value, kernel, nucleolus, and Nash bargaining. In strategy, the core is typically the solution concept of choice. Why the core? Because, as already noted, it imposes a weak set of competitive consistency conditions that seem reasonable in market settings. Moreover, we argue throughout, the identification for each agent of an *interval* of appropriation consistent with these conditions (i.e., rather than a point-estimate) is actually a *desirable* feature—one that generates several important insights about the effects of competition on value capture.

There are examples of biform game applications outside of strategy that use other solution concepts. For example, the property rights theory of the firm by Hart and Moore (1990) uses the Shapley value. The Shapley value is sensible in this setting because there are no competitive externalities requiring resolution outside of the game. This is not the case in many strategy contexts (de Fontenay and Gans, 2005, 2014). Also, some recent industrial organization papers doing structural estimation of vertical relations use the Nash bargaining solution for the CGT stage (Collard-Wexler, Gowrisankaran, and Lee, 2014). This also poses difficulties for dealing with competitive externalities (that arise when competing firms are prevented, say by antitrust laws, from negotiating directly with one another), but it has proven to be tractable for the study of vertical relations when downstream firms operate in distinct markets.

The primary issue raised against the core is that the primitives of value creation may be such such that there are no π_is satisfying conditions (1) and (2). This is of sufficient concern to Lippman and Rumelt (2003) that, at the same time they argue eloquently in favor of using CGT as a foundation for strategy theory, they also express reservations

[20] Byford (2010) provides a variant on the biform model in which the second stage is a CGT, but with linear prices as assumed by Kreps and Scheinkman.

about using the core as a solution concept. The source of the problem in such situations is always the same: V is too meager to be shared in a way that prevents every group G from acting on its alternative, $v(G)$.[21] The frequency of such situations in the real world is an open empirical question.[22] Moreover, Stuart (1997) demonstrates that the core does, in fact, exist for a wide range of situations relevant to strategy research. In our judgment, the overall benefits of using the core to develop strategy theory, particularly in terms of the insights it provides into value capture under competition, presently outweigh these concerns.

GENERAL PRINCIPLES

The extant theoretical literature using cooperative or biform games to investigate issues in business strategy can be categorized into two substreams. The first takes characteristic functions as given, operating at the general, abstract level to develop general principles with respect to value creation, competition, "persuasive" resources, and firm performance. The second builds up the characteristic function from primitives that are relevant to a particular issue. For example, the analysis may be specific to two-sided markets, productive networks, or differentiated products. These studies begin with specific assumptions about agents and their roles, buyer preferences, production technologies, costs, capacities, etc. The primitive assumptions are then used to construct the implied characteristic functions that arise from the agents' various strategic choices (i.e., in a biform game). In this section, we focus upon the historical development of the former and highlight its most significant findings.

The first paper to use CGT to examine value capture in strategy is Brandenburger and Stuart (1996). This paper introduces the strategy audience to the formal notion of *agent added value*.[23] Informally, an agent's added value is the difference between the aggregate value created in a market in which the agent participates and the amount of value that could be created were that agent to be removed from the market (alternatively, were all the other agents to shun transactions with that agent). Formally, the added value of agent i is:

$$av_i \equiv V - v\left(N_{-i}\right), \tag{5}$$

where N_{-i} denotes the set of agents other than i.

Principle of the added value (GP1)

From Equation (5), it follows that added value places an upper bound on what an agent can appropriate (Brandenburger and Stuart, 1996): If $\pi_i > av_i$, then satisfaction of Equation (1) implies $\sum_{j \in N_{-i}} < v\left(N_{-i}\right)$, which violates Equation (2). In words: *positive added value is a necessary condition for value capture*. Here, competition works against i in the sense that, the agents on the other side of the transactions with i that contribute to V must be induced to freely accept these and not some other set of transactions from which i is excluded—in this case, the ones involved in the production of $v(N_{-i})$. The added value of i is an upper bound on the other agents' "willingness-to-pay" for i's involvement in the production of V. When $\pi_i = av_i$, agent i is said to be a *full appropriator*.

Principle of adding up (GP2)

Brandenburger and Stuart (1996) clearly state that added value is a necessary but insufficient condition for value capture (p. 14). Nevertheless, their analysis proceeds (from p. 15) under the premise that achieving positive added value "is the path to value appropriation" (i.e., rather than being one step *on* that path). They then illustrate several examples in which every agent captures exactly its added value—no more, no less. These examples are built on an interesting condition known as the "adding-up" property: *if the agent added values themselves sum to V, then every agent captures*

[21] Put another way, for any collection of nontrivial $v(G)$s, there is a minimum V below which there are no distributions of value satisfying core conditions (1) and (2).

[22] In the real world, there are many ways competition may be softened to admit existence. For example, if antitrust law prohibits the agents in G from transacting, then $v(G) = 0$. Alternatively, if the agents in G are unaware of their productive possibility then, again, $v(G) = 0$.

[23] This idea arises in economics at least as early as Edgeworth (1881). More recently, economists Makowski and Ostroy, in a substantial line of joint work, demonstrate the importance of understanding that, in a general equilibrium model, agents and the value they capture are the mathematical dual to solutions based upon products and prices (see, e.g., Makowski and Ostroy, 1995). They refer to added value as "agent marginal product." Brandenburger and Stuart (1996) cite this work as having had in important influence on their paper (p. 23). See Makowski and Ostroy (2001) for a wide-ranging, highly accessible review.

DOI: 10.1002/smj

precisely its added value. Formally, if $\sum_{i \in N} av_i = V$ then, for all i, $\pi_i^{\min} = \pi_i^{\max} = av_i$.

The preceding result corresponds to what general equilibrium theorists Makowski and Ostroy (1995, 2001) define as "perfect competition." Strategy scholars who are used to identifying "perfect competition" with the zero profit world induced by free entry may find this definition jarring: here, perfect competition implies everyone is a full appropriator. Note that this definition includes the traditional meaning of perfect competition as a special case: under free entry, added values are driven to zero and, hence, so is value capture (agents are, trivially, full appropriators). Now, competition is "perfect" in the sense that it squeezes everyone's core interval to a single point (equal to the added value).

Principle of the minimum residual (GP3)

A grasp of the preceding, counterintuitive conceptualization of perfect competition invites scholars to stop thinking of competitive intensity as a force of pure badness. GP2 shows that, under the right circumstances, competitive intensity actually guarantees full appropriation. This finding naturally leads one to wonder about cases in which competition is not so intense as to induce trivial intervals. MacDonald and Ryall (2004) note the absence of attention in strategy research on the good side of competition—namely, its potential effect of guaranteeing positive value capture ($\pi^{min} > 0$). They define agent i's *minimum residual*: $mr_i = V - \sum_{j \in N_{-i}} av_j$. It follows that *every agent must capture an amount of value at least as great as its minimum residual*. Formally, $\pi_i \geq mr_i$. This is implied by the straightforward logic of GP1: since added value is an upper bound to value capture, if the sum of everyone else's added values (i.e., besides i's) is less than V, then i must receive (at least) the difference. While GP3 is simple to derive, its interpretation is subtle. Agent i's minimum residual is an indirect consequence of the $n-1$ alternatives to exclude precisely one of the other market participants from the transactions that produce V. When other participants add meager value compared to i, implicit competition *for i* is high, guaranteeing i positive value capture.

Principle of the essential tension (GP4)

The preceding are all based upon added value, an intuitive and technically simple idea. However, added value ignores the lion's share of the feasible opportunities to create value elaborated in competitive consistency condition (2). The agent added values depend upon the n groups of the form N_{-i}. More generally, however, there are $2^n - 1$ possible groups that can be formed from the n agents in N. If the number of agents in a market is even moderate, added value is unlikely to account for much of competition's overall impact on value capture. Having established GP3, MacDonald and Ryall (2004) go on to examine the effect of *all* the alternatives on an agent's competitive minimum. Specifically, they seek to characterize the conditions under which competition alone guarantees that an agent *must* capture a positive share of value. In service of this goal, they define an agent's *minimum value* and use it in their main proposition.

Agent i's minimum value, denoted mv_i, is the solution to the following optimization problem: minimize the aggregate amount of value required to make sufficient payments to agents other than i such that i receives zero and, yet, cannot induce any group to deviate from the production of V. Formally, mv_i is the solution to the following linear program:

$$mv_i \equiv \min_{\pi \in \mathbb{R}_+^n} \left\{ \sum_{j \in N} \pi_j | \pi_i = 0 \text{ and } \sum_{k \in G} \pi_k \geq v\left(G_{+i}\right) \text{ for all } G \subset N_{-i} \right\}, \tag{6}$$

where G_{+i} is the group G with i added. The important thing to note is that, the more productive i is with the various groups, the more the aggregate value required to pay off the other agents such that i—facing zero value capture—cannot organize a deal with the members of *some* that makes everyone in it (including i) strictly better off than what they would get by participating in the production of V. Put differently, if "the market" proposes that highly productive agent i receive a "price" of zero economic value in return for its participation in creating V, the other agents will need to receive large quantities of value if i is to be prevented from inducing them to deviate to some other activities.

Once the concept of minimum value is grasped, it is easy to see that $V < mv_i$ if and only if $\pi_i > 0$; i.e., competition guarantees strictly positive value capture when a particular tension arises between an agent's *potential* to create value versus the *actual* quantity of value it creates. Note that mv_i depends only upon the groups in which i is a member. Thus,

DOI: 10.1002/smj

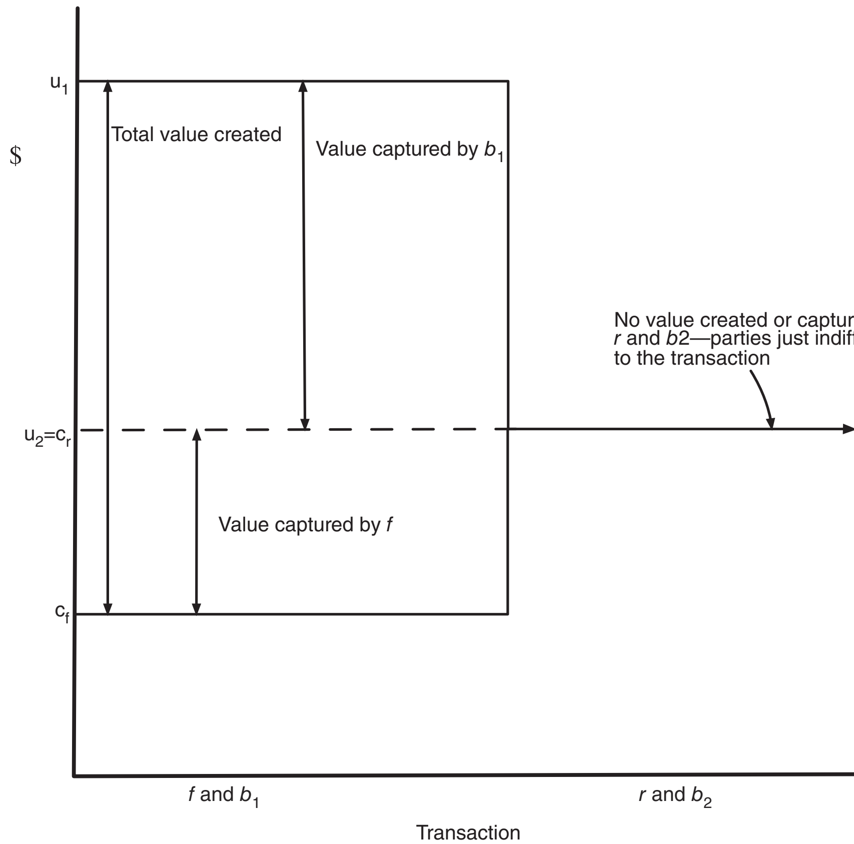

Figure 3. Perfect competition

mv_i is an important measure of competition *for* i. MacDonald and Ryall (2006) explore whether the entry of any new agent (direct competitor, supplier, employee, buyer, etc.) affects the firm's competitively guaranteed value capture—'under what conditions is this guarantee created, destroyed, or maintained? In particular, they identify which categories of value-creating groups—those that include the entrant, exclude it, include only the firm and the entrant, and so on—matter for the creation, destruction, or maintenance of the guarantee.

A new example ending in a paradox

Before moving on to the last general principle in this section, it may be instructive to illustrate these various results with a series of related examples in the simplest context possible—a small, two-sided market. To start, assume there are two firms, f and r (for "rival"). Each firm has one unit of capacity with which to make a product at constant marginal cost c_f and c_r, respectively, with $c_f < c_r$. There are two buyers, 1 and 2, who each wish to consume exactly one unit of product. Suppose each buyer views the firms as indistinguishable: u_i denotes the utility enjoyed by buyer i from consuming one unit of either firm's product. At the same time, assume the buyers are heterogenous in their utilities such that $u_1 > u_2 = c_r$. Since the buyers view the firms' products as identical, which buyer buys from which firm is not a critical issue: the overall value created in this market should be $V = u_1 + u_2 - c_f - c_r$ or, since $u_2 = c_r$, $V = u_1 - c_f$. Assume the agents' outside options are normalized to zero.[24] See Figure 3.

The first step is to compute added values. It should be easy to see that removing r or buyer 2 from the market has no effect on the aggregate value produced. Therefore, $av_r = av_2 = 0$. In the absence of f, r and buyer 1 transact to produce $v(\{r, 1\}) = u_1 - c_r > 0$. Therefore, f's added value is $av_f = c_r - c_f$ (the incremental effect of its superior

[24] Setting $v(\{i\}) = 0$ is without loss of generality.

DOI: 10.1002/smj

Table 1. Implications of added values

	Firm f	Firm r	Buyer 1	Buyer 2
Added value	$c_r - c_f$	0	$u_1 - u_2$	0
Implication for $\pi_i^{\max}$	$\pi_f^{\max} \leq c_r - c_f$	$\pi_r^{\max} = 0$	$\pi_1^{\max} \leq u_1 - u_2$	$\pi_2^{\max} = 0$
Implication for $\pi_i^{\min}$	None	$\pi_r^{\min} = 0$	None	$\pi_2^{\min} = 0$

Table 2. Implications of adding up

	Firm f	Firm r	Buyer 1	Buyer 2
Added value	$c_r - c_f$	0	$u_1 - u_2$	0
Implication for $\pi_i^{\max}$	$\pi_f^{\max} = c_r - c_f$	$\pi_r^{\max} = 0$	$\pi_1^{\max} = u_1 - u_2$	$\pi_2^{\max} = 0$
Implication for $\pi_i^{\min}$	$\pi_f^{\min} = c_r - c_f$	$\pi_r^{\min} = 0$	$\pi_1^{\min} = u_1 - u_2$	$\pi_2^{\min} = 0$

cost position). Similarly, removing buyer 1 results in a transaction between f and buyer 2, which implies $av_1 = u_1 - u_2$ (the incremental effect of the greater value it places on these products). The implications of GP1 are summarized in Table 1.

As it turns out, there is much more we can say. A moment of reflection indicates that GP2 applies:

$$\sum_{i \in N} av_i = c_r - c_f + 0 + u_1 - u_2 + 0$$
$$= u_1 - c_f.$$

Since $u_1 - c_f = V$, we know that each agent must receive *exactly* its added value. Competition is fully determinative; there is no room for extracompetitive value appropriation.

When adding-up is satisfied, all minimum residuals equal their corresponding added values. Since an agent can get no less than its minimum residual and no more than its added value, it must get exactly this amount. Here, competition for each agent (as measured by minimum residual) is perfectly balanced with competition for their transaction partners (as measured by added value), thereby guaranteeing that every agent is a full appropriator (see Table 2).

GP3 also holds. To see this, begin by imposing $\pi_f = 0$ and ask how much value is required to ensure that f cannot offer the buyers a profitable side deal. The answer is straightforward. If $\pi_f = 0$, buyer 1 must capture at least $\pi_1 = u_1 - c_f$. Otherwise, competitive consistency fails since it would be the case that $\pi_f + \pi_1 < v(\{f, 1\}) = u_1 - c_f$. Similarly, buyer 2 must capture at least $\pi_2 = u_2 - c_f > 0$. Happily for f, $V = u_1 - c_f$ is not sufficient to make both payments, which is what is required to neutralize f's ability to upset the production of V in the event that $\pi_f = 0$: if the buyer transacting with f insists on $\pi_f = 0$, f can always make a deal with the other buyer that makes both it and the other buyer strictly better off. Therefore, we know that $\pi_f^{\min} > 0$—competition for f is too intense for it to capture no value (which we already knew, but now see as a result of the essential tension). The same reasoning applies to buyer 1. Finally, it is not hard to see that $mv_r = mv_2 = V$. There is exactly enough value available to assure that r and buyer 2 receive no economic value.

The paradox

Assume everything is as before, with the exception that now the buyers are identical—each values one unit of consumption from either firm at $u > c_r$. Now, $V = 2u - c_f - c_r$, arising from two transactions, one each between a buyer and a firm. Focus attention of f. The competitive distribution in which f captures its maximum is the one in which the firms are full appropriators. That is, $\pi_f^{\max} = av_f = u - c_f$. The worst f can do is to capture $\pi_f^{\min} = c_r - c_f$. The reason is that r's buyer competes with f's buyer to do a deal with f because it is the efficient producer. Thus, if f's buyer insists on leaving it with less than $c_r - c_f$, r's buyer can agree to a deal that makes both strictly better off.

Now, suppose f is not happy with 50 percent market share and, being the efficient producer, decides to "own" the market by expanding capacity to two units. How much better off is f following such a move? Both buyers transact with f to produce $V' = 2(u - c_f)$. If f is removed from the market,

DOI: 10.1002/smj

one buyer will still be able to transact with r. Therefore, its added value is $av'_f = 2\left(u - c_f\right) - \left(u - c_r\right) = u - c_f + c_r$. Note well that $av'_f > av_f$.

In this situation, there is nothing preventing the buyers from being full appropriators: feasibility and competitive consistency conditions are all satisfied in such a scenario. Therefore, $\pi_f^{\min'} = 0$. Moreover, the buyers must each appropriate $u - c_r$. Any amount less than this allows the buyer to cut a deal with r that makes both better off. Therefore, $\pi_f^{\max'} = 2\left(u - c_f\right) - 2\left(u - c_r\right) = 2\left(c_r - c_f\right)$.[25] Suppose f's efficiency advantage is small relative to the quality of its product—specifically, that $2(c_f - c_r) < (u - c_f)$ – then, $\pi_f^{\max'} < \pi_f^{\max}$.

In other words, reasonable parameters can be found for this example such that f unambiguously (1) increases its added value through the capacity expansion and (2) worsens its situation with decreases in *both* its minimum and maximum. This outcome is so counterintuitive that it has been used by some in economics to argue against the use of the core as a solution concept.[26] As we now show, the result is actually quite sensible once the shifting effects of competition in this example are properly understood.

Principle of competitive intensity (GP5)

Montez *et al.* (2015) are motivated by Insights 1–3 of the earlier section. In particular, they investigate increasingly complex notions of "competitive intensity," each of which is consistent with the thrust of these insights. As it turns out, the simplest of these fully explains why the preceding result, though counterintuitive, is economically intelligible.

To see this, begin by defining a *focal value partition*, denoted $\mathscr{P}^*$, for a cooperative game. Mathematically, $\mathscr{P}^*$ is a collection of subsets of agents with the following properties: (1) $\mathscr{P}^*$ is a partition of N and (2) $\sum_{G^* \in \mathscr{P}^*} v\left(G^*\right) = V$.[27] The groups in $\mathscr{P}^*$ are called *value networks*.[28] This structure is meant to represent industries containing distinct value chains—each value network is a set of agents linked to one another through their transactions, with the value created by the industry being equal to the aggregate of the value created by each network. For any particular game, there may be many candidates for $\mathscr{P}^*$ (i.e., many ways to allocate the agents into groups such that the group values add up to V). The label "focal" is intended to convey the idea that $\mathscr{P}^*$ corresponds to a real situation; e.g., the actual transactions an empiricist would identify upon observation of the industry.

To understand the simplest notion of competitive intensity and grasp its implications for value capture, assume an industry is decomposable into two or more value networks. The notion of competitive intensity is based upon added value. Symbolically, if G^* is a value network in which firm f is not a member, then f's *added value to* G^* is given by $av_f\left(G^*\right) = v\left(G^*_{+f}\right) - v\left(G^*\right)$. If $av_f(G^*) > 0$, then the value network G^* competes *for* f—the members of that network would like f to join because, in that way, more value will be produced. Competition of this kind raises f's minimum. Indeed, $\pi_f \geq av_f(G^*)$. The reason is that the agents in a value network must, in aggregate, capture exactly the value produced by that network; i.e., $\sum_{i \in G^*} \pi_i = v\left(G^*\right)$. If, in addition, $\pi_f < av_f(G^*)$, then the group comprised of G^* and f violates the competitive consistency condition.

An agent's "first-order competitive intensity" (FOCI) is defined as the maximum value added to a value network outside its own. The general principle is that *every agent must capture value at least as large as its first-order competitive intensity*. The economic intuition is immediate from the preceding logic. Consistent with Insight 1, the logic applies symmetrically to all the other agents in f's own value network. The greater the competitive intensity for f's transaction partners in that network, the less value is left over for f —its maximum is driven down. This is consistent with Insight 2. As competitive intensity for f and its transaction partners increases, f's competitive interval narrows, consistent with Insight 3.

Our earlier, "paradoxical," example is fully explained by this simple notion of competitive intensity. Let us see how, by assuming the parameters are such that the nonintuitive outcome arises; i.e., $2(c_f - c_r) < (u - c_f)$. Consider the focal value partition prior to the capacity expansion. Suppose

[25] Firm r is still in the market. Obviously, the persistent lack of revenue must induce exit. However, it is sufficient that r's capacity utilization drops to less than 100% (i.e., and possibly greater than 0%).

[26] See, e.g., Postlewaite and Rosenthal (1974).

[27] A partition of N is a collection of disjoint groups, the union of which equals N.

[28] The asterisks on the groups indicate their status as value networks.

DOI: 10.1002/smj

f sells to buyer 1 and r to buyer 2. Then, the value networks can be denoted $F^* = \{f, b_1\}$ and $R^* = \{r, b_2\}$. These represent simple transactions, with $v(F^*) = u - c_f$ and $v(R^*) = u - c_r$. This implies $V = v(F^*) + v(R^*)$. Therefore, $\mathscr{P}^* = \{F^*, R^*\}$ fulfills the requirements for a value network. We loosely conceive of buyer 2 creating competition *for f* because it provides an opportunity to create more value than does transacting with r. FOCI quantifies this: the added value of f to R^* is $c_r - c_f$ (in the group containing f, r, and b_2, b_2 simply deals with f rather than r). Therefore, $\pi_f \geq c_r - c_f$. At the same time, buyers 1 and 2 are interchangeable. As a result, there is no FOCI for b_1: inserting b_1 to R^* does not create any additional value. Thus, $\pi_f \leq u - c_f$.

How do the economics change when f expands capacity? First, both buyers transact with f. We have $F^{*\prime} = \{f, b_1, b_2\}$ and $R^{*\prime} = \{r\}$ with $v(F^{*\prime}) = 2(u - c_f)$, $v(R^{*\prime}) = 0$ and $V = v(F^{*\prime}) + v(R^{*\prime})$. Now, having enticed b_2 into its own transaction network, competition for f *disappears*. Simultaneously, the empty capacity in $R^{*\prime}$ *creates* competition for both of f's buyers.[29] The loss of competition for f drops its minimum and the appearance of competition for f's buyers drops its maximum. The FOCI for the post-capacity expansion of f is zero, while the FOCI for *each* buyer is $u - c_r$.

Beyond the economic insights provided, there are a couple of additional points worth mentioning with respect to this novel conceptualization of competitive intensity. First, the definition of FOCI presents an alternative solution concept to the core. Second, unlike the core, it always exists. Third, the concept is simple—it relies on a very small number of groups, as compared to the core. Moreover, it is set up to facilitate reference to actual transactions (i.e., as embodied in the focal value networks). These features make it much friendlier to empirical analysis than findings that rely on the entire set of core constraints. Finally, it provides a basis for related extensions, several of which are investigated (e.g., a notion of indirect competition called second-order competitive intensity).

[29] Note that the effect would have been the same had the example been such that several other buyers preferred r's product and remained with it. The key to the outcome is that b_2's switch leaves r with one unit of slack capacity.

APPLICATIONS TO SPECIAL SETTINGS

In this section, we turn our attention to papers that approach their analysis from the bottom up. That is, whereas the work cited in the previous section begins with an arbitrary characteristic function (and, hence, provides abstract but general findings about the nature of value creation and capture in markets), the following work starts with specific assumptions that are relevant to a specific setting of interest, and then analyzes performance in light of the resulting characteristic function.[30]

The application papers reviewed below are a significant complement to those focused on the development of general principles—from them, a deeper understanding emerges of the theoretical links from variation in agent resources, skills, strategies, and contexts to variation in the force of competition and effectiveness of persuasive abilities (and, hence, in performance itself). In the following discussion, the connections made from the specific findings of these papers to the general principles elaborated above are our own, with the intention of illustrating how the latter illuminate the former and, thus, of conveying a sense of the overall coherence of this line of work.

A demand-based perspective

One of the important messages in the value capture literature, that arises implicitly from the formalism, is that end-users are integral to the value creation and capture story: without them, there is no value creation. While it is possible to construct a biform game that shunts buyers into the background (à la the demand curve in a Cournot model), there is nothing in the formalism that requires it. Indeed, the modeler is encouraged to treat buyers on an equal footing, in terms of model inclusion, as any other kind of agent. Thus, the approach is a natural choice for Adner and Zemsky (2006), who are motivated to study demand-side influences on sustained competitive advantage.

Adner and Zemsky (2006) set up a model in which consumers vary in terms of their preference for quality as well as the extent to which they have decreasing marginal utility (DMU) for performance. Firm heterogeneity is captured on four

[30] In some papers, the characteristic function is explicitly derived from the more primitive assumptions. In others, it is left implicit to operate behind the scenes.

DOI: 10.1002/smj

dimensions: (1) rate of technology acquisition, (2) technology lead (or lag), (3) a fixed quality component, and (4) marginal cost of production. Value capture is assumed to be proportional to value added.

The analysis goes on to relate competitive advantage to the interaction between firm strategies and the degree of DMU inherent in consumer preferences for product performance. For example, they show that the ability of a challenger to unseat an incumbent through innovation depends subtly upon a firm's cost position and the degree of DMU. If the new technology comes with a cost advantage, higher DMU speeds adoption. However, if the challenger has a cost disadvantage, higher DMU retards adoption.

The paper also explores GP1 under the effects of different types of resources (process, product, innovation), the effects of imitation, multisegment competition, and strategic diversity. Throughout, the focal issue is on the interaction between firm strategies and the features of consumer preferences (primarily DMU). Because the model assumes value capture is proportional to added value, competitive intensity and the distinctions between competitive versus persuasive sources of performance advantage are not a feature of the analysis.

Value capture in productive social networks

A persistent empirical finding with respect to productive social networks is that network brokers—agents who intermediate transactions between two or more parties—tend to enjoy higher levels of value capture than agents with similar characteristics but less central network positions (see Burt, 2000, for a review of empirical findings). A question that naturally arises is whether the superior performance is due to a broker's network position, or whether the agent's position is due to its superior productive characteristics. In the latter case, network centrality and superior performance are due to the common cause of the agent's ability to create superior value.

Ryall and Sorenson (2007) build a biform game in which second-stage value creation possibilities depend upon a collection of possible projects which, themselves, require subsets of connected agents for completion. For example, an action movie requires a set of agents—director, actors, producers, special effects artists, and so on—with complementary abilities to come together for its completion. The key feature of the model is that agents must be connected within a productive network in order to work on a project. The formation of network relationships occurs in the first stage of the biform game, in which agents make and accept (or reject) offers to work with one another.

The paper demonstrates that brokers can, in fact, enjoy guaranteed value capture purely by virtue of the competition for them generated by their network position, viz. the Principle of the Essential Tension. However, it also illustrates the insight that attaining such a position is problematic: If an individual really has no effect on value creation per se, the agents who *are* required for value creation have no interest in accepting relationships that have the end-effect of installing that individual into a value-soaking position of network scarcity.

Vertical integration

De Fontenay and Gans (2008) examine firms' decisions to outsource functions and transact under separate ownership. In their analysis, the NGT part of the biform game is a market where real assets are traded while the CGT part of the game is based on the Shapley value. While it is well known that consolidation of ownership can allow those owners to capture more value, this must be traded off against any value creation that might arise form outsourcing (that is, vertical disintegration or separation of ownership). What de Fontenay and Gans (2008) show is that, consistent with our earlier blind spot regarding the harm generated by competition, it matters *who* you outsource to.

For instance, if you outsource and use that to create competition, that improves value capture in the CGT part of the game relative to outsourcing to an existing firm that keeps competition (at least structurally) the same. But, precisely because of this, the NGT part of the game that involves a transaction with the outsourcing firm (which acquires assets to provide services), involves several parties that care about the degree of competition that results. The paper shows that the NGT incentives dominate the CGT outcomes and, in equilibrium, a firm will end up outsourcing to an established firm (keeping competition structurally the same) rather than an independent firm that creates more competition. Thus, the biform structure of value capture theory—by forcing us to understand the full consequences of an outsourcing relationship—can turn traditional

DOI: 10.1002/smj

intuition about the competition-creating effects of outsourcing on its head.

Strategic factor markets

The RBV claims that sustained performance advantage arises from control of resources and capabilities that are rare, valuable, and inimitable. Barney (1986) and Makadok and Barney (2001) further assert that a source of inimitability is a superior understanding of resource value. To frame the idea in the logic of this paper: Competition for the transaction partners of a superior performer may fail to appear because the agents required to provide such competition simply do not understand the value creation possibilities of their resources in that context. Heterogeneous expectations, then, become the basis for "strategic factor markets" (SFM) —markets in which value-creating resources are demanded and supplied as a result of heterogeneous expectations regarding their value.

In elaborating the free-entry blind spot, we explained why material finitude implies that resource mobility barriers are neither necessary nor sufficient for sustained performance advantage.[31] Adegbesan (2008) explores the substance of this implication explicitly in the context of strategic factor markets. The setup is an assignment model: a cooperative game representation of a two-sided market in which sellers offer units of an undifferentiated product to buyers. While the sellers' products are identical, buyers do not value them identically.[32] The interpretation here is that the products are resources that offer heterogeneous complementarities to the firms on the buy-side of the market. Thus, performance heterogeneity does not arise as a result of heterogeneous expectations but, rather, of heterogeneous resource complementarities. Note well the implicit premise that the features of the buy-side firms that create the complementarities are, themselves, in finite supply.

Adegbesan (2008) demonstrates a number of interesting findings. First, the analysis shows that the resource acquirers with greater complementarities are guaranteed greater value capture (via the Principle of the Essential Tension). Second, the paper highlights the traditionally overlooked fact that a portion—potentially significant—of superior performance may be due to persuasive capabilities (relating to Insight 4 and to the Principal of Competitive Intensity). This latter issue is in the context of SFMs themselves. At the top end, of course, the importance of persuasive capabilities is capped according to the Principle of the Added Value. Here, added value is also computed in context of the SFM and arises as a consequence of complementarities.

Value chain frictions

Chatain and Zemsky (2011) study what happens with respect to value creation and capture when frictions arise in industry value chains. The issue arises as a natural consequence of conditions (1) and (2) for a competitive distribution of value. The number of possible value-creating groups that one must, theoretically, consider is exponentially increasing in the number of agents, n. In most real-world situations, search and switching costs may prevent agents from being aware of or, equivalently, being unable to act on a large portion of these possibilities.

The model examines two firms competing for a single buyer, in which each firm fails to meet the buyer with some probability. This quantifies the firm-specific "friction" inherent in the market. The analysis proceeds to examine the effects of friction in the context of each of the five "forces" in Porter (1979). An important, cumulative insight arising from these findings is that Porter's Five Forces cannot be considered in isolation. For example, changes in persuasive capabilities affect threshold friction levels that induce entry. This message is consistent with our earlier Insight 1: there is only *one*, highly interactive force of competition. Changing any one of the values actually created, relative persuasive advantage or the value of productive alternatives, has the potential to alter the balance of competition throughout the system in subtle and counterintuitive ways.

Obloj and Zemsky (2015) examine competitive advantage in the context of buyer–supplier relationships in which the integrity of the supplier plays a role in determining which deals are consummated, how much value is created, and how that value is distributed. The novel aspect of this setup is that it embeds a contracting problem in the first stage: contracts are offered in the first (NGT) stage, and value creation and capture are determined in the second

[31] See also Pacheco-de Almeida and Zemsky (2007: 651–652) for a discussion of other issues that have been raised with respect to the theoretical underpinnings of the RBV.

[32] The canonical example is a housing market in which all the sellers offer houses that, for all relevant purposes, are identical. However, buyers may have a different willingness-to-pay due to their individual preferences for house features.

DOI: 10.1002/smj

(CGT) stage. Value creation in stage two depends upon the relative efficiencies of supplier production technologies, the contracts offered, the actual deals struck (implied by the characteristic function), and the gaming proclivities of the suppliers. Actual value created depends only upon supplier efforts devoted to productive actions. Buyers would like to contract on these actions, but cannot. Instead, they can only contract on an observable that, while correlated with productive effort, can be gamed when suppliers engage in certain unproductive tasks. Suppliers are of two types—gamers and honest—with the former more inclined to divert effort to unproductive activities. This sets the stage for an interesting moral hazard problem integrated into a biform game.[33]

This paper has a number of interesting results. For example, dishonest suppliers who yet have a sufficient efficiency advantage may have greater added value than honest suppliers and, hence, be selected by buyers. By implication, it is possible for inefficiency rears its head in stage two whenever a deal is consummated with a dishonest supplier. More significantly, they show that an increase in integrity on the part of *either* type of supplier results in: (1) greater value creation *regardless of which supplier is ultimately chosen* for the deal, (2) greater value capture for the buyer, (3) greater value capture for the supplier with greater integrity, when used, and (4) less value captured by the supplier with no change in integrity, when used.

Apart from these specific results (and more), this analysis is significant for at least two additional reasons. First, it demonstrates the idea that the characteristic function represents the *actual* value producible by agents in various combinations, "implicitly accounting for limitations implied by information and agency considerations, transactions costs, configuration of productive resources, barriers to technology transfer, institutional structure, regulation, and so on" (MacDonald and Ryall, 2004: 1324). Second, the model identifies a novel category persuasive resource—reputation for honesty—and examines its effect on value creation and capture. In so doing, it also reminds us to cast our nets wide regarding the possible sources of value capture above and beyond the competitive minimum.

[33] There is no adverse selection here—buyers know the supplier type (e.g., due to unmodeled reputation effects).

EMPIRICAL PROGRESS

Value capture theory has recently led to empirical work that utilizes both the theory's insights and its methodological innovations. Here, we review these papers, highlighting how value capture theory has given researchers a way of endogenizing elements—particularly regarding multiagent interactions—that were previously held exogenous when formulating hypotheses. This has led both to cleaner hypothesis specification and new ways of looking at the identification of causal relationships in explanations of firm performance.

Consider, first, the papers that use value capture theory for hypothesis formulation. One example is Marx, Gans, and Hsu (2014), who examine the commercialization choices of startup firms in the speech recognition industry. In their model, startups face a choice between competing with established firms or cooperating with them. However, what makes this a complex interaction is that the choice is exercised over time and startups can pivot between different commercialization strategies. Their model uses value capture theory by explicitly modeling the amount of value created through a cooperative choice relative to a competitive choice, taking into account the fact that those choices may impact on future value creation and also the value that each party can guarantee to appropriate. This allows them to hypothesize how technological characteristics impact the sequence of observed commercialization choices of startups. For instance, for disruptive technologies that are initially of low value to incumbents and are difficult to integrate, competition will be preferred. However, following the resolution of uncertainty over the future trajectory of the technology, a switch to cooperation may be observed. As that uncertainty is actually resolved to the empirical observer, Marx *et al.* (2014) can actually test whether those pivots occurred—and whether they related to the technology being commercialized in the fashion anticipated by the theory. They found that the data were, indeed, consistent with the theory.

Chatain (2011) uses value capture theory to investigate an important new area regarding the resource-based view of the firm: namely, how does competition itself impact on the relationship between capabilities and performance? This is a subtle issue. As noted earlier, while improved capabilities improve value creation (in general), they may not improve a firm's ability to appropriate

DOI: 10.1002/smj

that value precisely because the terms of competition change. Chatain (2011) distinguishes between certain actions that improve total value creation between a firm (in his application law firms) and its clients but in a way that is specific to a client (for instance, knowledge of a law firm of third-party contract specifics or the details of manufacturing operations) and those that are more generic. For the former, the very action reduces the set of agents that can compete effectively for the client's business. He finds that if a client requires new services, then those are more likely to be provided by the client's existing law firm when client-specific knowledge is an important component of those new services. By contrast, the aggregate level of expertise in the firm does not spur a similar bias in a client's choices for new services. Finally, the agglomeration of services is related positively to the client-specific relationships of a firm and also negatively to the client-specific relationships of rivals. Using data on U.S. law firms, Chatain (2011) confirms the predicted hypotheses from value capture theory in this context.

One of the primary empirical challenges posed by the theory—especially since it represents one of the theory's essential features—is how to work with the inequality constraints posed by competitive consistency conditions (2). Chatain (2013) breaks new methodological ground in this stream by implementing value capture estimates based upon an estimation approach that admits the use of inequality constraints. This approach was introduced to economics by Fox (2008) and first applied in strategy by Mindruta (2013) in her study of the drivers of researcher–firm matching. Applying this new methodology to the two-sided market between law firms and their corporate clients, Chatain (2013) uses law firm quality rankings to construct implicit willingness-to-pay on the part of firms that, in turn, implies inequalities, Equation (2). He finds that law firms with broader client scope appropriate less and, moreover, at an equal cost per lawyer as those with narrrower scope—results that run counter to the prevailing wisdom of practitioners in this industry.

Also ambitious is Grennan (2014). His paper examines negotiations between suppliers and buyers (usually hospitals) of medical devices. Grennan (2014) is able to observe total value created computed from observed prices and market shares for coronary stents. Importantly, the prices are negotiated prior to quantity being chosen, with doctors doing that depending on demand and relative costs. This allows Grennan (2014) to identify key elasticities that drive willingness to pay for these stents because some of the later variation in demand are not observed or anticipated when prices are negotiated (e.g., a large study taking place in a hospital). Grennan (2014) also observes the supplier costs of stents based on their type. He then uses the Nash bargaining solution, simultaneously solved and applied across the market, to generate a structural equation for the determinants of supplier appropriation. As he shows, this is equivalent to the value capture model based on the core in the context of his setting. Critically, this allows him to decompose the determinants of value capture by suppliers into their competitive options (namely, computed added value) and also bargaining ability. He finds that 79 percent of the variation in appropriation across suppliers is explained by bargaining ability. Moreover, this ability is itself correlated with firm identity. It differs across hospitals, improving over time as hospitals learn how to negotiate with suppliers.[34] This suggests the long-anticipated efficacy of trying to consider firm organization and performance as embedded in their competitive environments.

Bennett (2013) explores the sources and effects of persuasive resources in an analysis rich with institutional detail. He exploits variation in the organizational design of car dealerships to examine the impact of this on subsequent bargaining outcomes in a biform game. In some car sales negotiations, a salesperson owns the entire deal (a parallel process), while in others deals are closed by a, presumably, more expert salesperson (a serial process). Recall that the price eventually negotiated will be, under value capture theory, a function of competitive options for the seller and the seller's relative bargaining ability. Because bargaining ability is likely to be higher when an expert salesperson is brought in, then it is also the case that the impact of competitive options on the final price is lower in this organizational form. Bennett shows, consistent with the theory, that as the inventory of cars diminishes (a proxy for competition or scarcity), prices increase under the serial sales process by relatively less. In addition, using the college education of buyers as a proxy for Internet search saaviness, Bennett

[34] Adegbesan and Higgins (2011) also find a similar impact on bargaining ability in their study of biotech–pharmaceutical company alliances, although their analysis is not structural in nature, nor is it causally identified in the way Grennan (2014) is able to do in his setting.

DOI: 10.1002/smj

finds that in this situation the serial process performs worse than the parallel one; i.e., the costs of that more drawn-out process are not translated into pricing benefits.

In summary, the empirical work based on value capture theory is young but already yielding valuable insights. In particular, the broad predictions of the theory do appear in the data in a variety of applications. Moreover, some empirical work has demonstrated that competition alone cannot explain all of the variation in firm performance and that some idiosyncratic factors that drive relative bargaining power can be very significant. It will be interesting to see how that assists in focusing out attention on the drivers of those factors, as well as how far the theory can be pushed to assist in quantifiable structural modeling.

FUTURE DIRECTIONS

While the cooperative game theory has a long history in economics, it is relatively recently that this methodology began being applied to the central questions in strategic management. That theory reviewed here grounds the issue of firm performance in both a market setting and in a setting where resource constraints are endogenous to overall economic activity. This allows the researcher to fully characterize how the force of competition provides bounds on a firm's appropriability and how these bounds relate to value capture by other firms, customers, and suppliers. This approach not only provides insights into the robustness of predictions stemming from more conventional models of market competition but also provides entirely new insights that are not available from those models. Here, we have highlighted areas where the theory provides both nuance and challenge to existing preconceptions.

At the same time, we must stress that these insights have just begun to be broadly tested by our empirically motivated colleagues. Indeed, as excited as we are by the promise posed by the recent flurry of empirical work, the surface has only been scratched. In our view, one of the most important issues is further effort designed to identify and measure the impact of "persuasive" resources. That is, how do firm strategies affect the value capture Equation (3)? What, exactly, are "persuasive" resources? Do they operate as predicted? Are they more or less likely to convey durable advantages to their owners? Of course, these questions are closely related to competitive intensity. The theory predicts that persuasive resources are important when competitive intensity is low, competitive resources more so when intensity is high. Initial empirical work (e.g., Bennett, 2013) is promising. However, the extent to which competitive intensity varies across markets and firms is an open question. If the typical firm faces strong intensity (and, hence, a narrow competitive interval), then claims about persuasive resources are a theoretical curiosity without real-world relevance. In either event, the answer will provide a deeper understanding of persistent performance heterogeneity among related firms.

Another important opportunity, as highlighted by Grennan (2014), is the ability to use value capture theory to bring structural modeling to strategy—in a fashion distinct from the present approaches used in the empirical industrial organization (based on dynamic, NGT). Combined with a refined understanding of how firms' actions affect their value capture equations, structural models along the line of Grennan (2014) permit researchers to explore the drivers of firm performance in a market setting. Moreover, they permit "what-if" assessment of business policy counterfactuals. The formal approach requires precision, but with that precision comes a more subtle set of hypotheses that can be tested and examined.

Finally, the theory is far from settled. The empiricist's ability to classify resources to capabilities into persuasive versus competitive categories will be greatly facilitated by further theoretical work on this issue. This suggests a need for more theory that avails itself of the full biform formalism (in contrast to much of the theoretical work to date, which focuses primarily upon the second-stage, CGT part of the model). While the early focus on understanding this less familiar side of the theory is justifiable, the pressing questions for strategy transcend it. Critically, we must understand how the implications of value capture drive actual firm behavior—how do managers formulate strategies to capture value given their competitive environments? What *should* they do? In thinking seriously about managerial action in the real world, issues of bounded rationality inevitably arise. Many tools have been developed for the study of bounded rationality using NGT. Here we have in mind evolutionary game theory, subjective rationality, game theoretical learning, and behavioral economics. All of these are available to biform game analysis. Hence, we are

DOI: 10.1002/smj

optimistic that future strands of this work will take such issues seriously and explore their implications for strategy. In summation, there is much, much more to be done.

ACKNOWLEDGEMENTS

Thanks to Adam Brandenburger, Martin Byford, Olivier Chatain, Catherine de Fontenay, Glen MacDonald, and Gus Stuart for detailed discussions over the years. We also wish to express our gratitude to the editorial team of the special issue and the excellent comments and suggestions provided by three anonymous referees. Responsibility for all views expressed remains our own.

REFERENCES

Adegbesan JA. 2008. On the origins of competitive advantage: strategic factor markets and heterogeneous resource complementarity. *Academy of Management Review* **34**(3): 463–475.

Adegbesan J, Higgins M. 2011. The intraalliance division of value created through collaboration. *Strategic Management Journal* **32**(2): 187–211.

Adner R, Zemsky P. 2006. A demand-based perspective on sustainable competitive advantage. *Strategic Management Journal* **27**(3): 215–239.

Barney J. 1986. Strategic factor markets: expectations, luck, and business strategy. *Management science* **32**(10): 1231–1241.

Bennett V. 2013. Organization and bargaining: sales process choice at auto dealerships. *Management Science* **59**(9): 2013–2018.

Bloom N, Eifert B, Mahajan A, McKenzie D, Roberts J. 2013. Does management matter? Evidence from India. *Quarterly Journal of Economics* **128**(1): 1–51.

Brandenburger A, Nalebuff B. 1996. *Co-opetition*. Doubleday: New York.

Brandenburger AM, Stuart HW. 1996. Value-based business strategy. *Journal of Economics & Management Strategy* **5**(1): 5–24.

Brandenburger A, Stuart H. 2007. Biform games. *Management Science* **53**(4): 537–549.

Burt R. 2000. The network structure of social capital. *Research in Organizational Behavior* **22**: 345–423.

Byford MC. 2010. A constrained approach to price formation. unpublished Working paper, University of Melbourne, Melbourne, Vic, Australia.

Chatain O. 2011. Value creation, competition, and performance in buyersupplier relationships. *Strategic Management Journal* **32**(1): 76–102.

Chatain O. (2013). Estimating value creation from revealed preferences: application to value-based strategy. Unpublished Working paper, Paris, France.

Chatain O, Zemsky PB. 2011. Value creation and value capture with frictions. *Strategic Management Journal* **32**: 1206–1231.

Collard-Wexler A, Gowrisankaran G, Lee R. 2014. Nash-in-Nash bargaining: A microfoundation for applied work. Working paper 20641, National Bureau of Economic Research, Cambridge, MA.

Cubbin J, Geroski P. 1987. The convergence of profits in the long run: inter-firm and inter-industry comparisons. *Journal of Industrial Economics* **35**(4): 427–442.

Dosi G. 2007. Statistical regularities in the evolution of industries. A guide through some evidence and challenges for the theory. In *Perspectives on Innovation*. Laboratory of Economics and Management, Sant'Anna School of Advanced Studies: Pisa, Italy; 153–186.

Dosi G, Lechevalier S, Secchi A. 2010. Introduction: interfirm heterogeneity—nature, sources and consequences for industrial dynamics. *Industrial and Corporate Change* **19**(6): 1867–1890.

Edgeworth FY. 1881. *Mathematical Psychics: An Essay on the Application of Mathematics to the Moral Sciences*. C. Kegan Paul & Co.: London, UK.

Ellsberg D. 1961. Risk, ambiguity, and the savage axioms. *Quarterly Journal of Economics* **75**(4): 643–669.

de Fontenay CC, Gans JS. 2005. Vertical integration in the presence of upstream competition. *RAND Journal of Economics* **36**(3): 544–572.

de Fontenay CC, Gans JS. 2008. A bargaining perspective on strategic outsourcing and supply competition. *Strategic Management Journal* **29**(8): 841–857.

de Fontenay CC, Gans JS. 2014. Bilateral bargaining with externalities. *Journal of Industrial Economics* **62**(4): 756–788.

Fox JT. 2008. Estimating matching games with transfers. Working paper 14382, National Bureau of Economic Research, Cambridge, MA.

Gans JS. 2005. *Core Economics for Managers*. Thomson: Melbourne, Australia.

Gans J, MacDonald G, Ryall M. 2008. The two sides of competition and their implications for strategy. Unpublished Working paper, University of Toronto, Toronto, ON, Canada.

Grennan M. 2014. Bargaining ability and competitive advantage: empirical evidence from medical devices. *Management Science* **60**(12): 3011–3025.

Hart O, Moore J. 1990. Property rights and the nature of the firm. *Journal of Political Economy* **98**(6): 1119–1158.

Hopenhayn H. 1992. Entry, exit, and firm dynamics in long run equilibrium. *Econometrica* **60**(5): 1127–1150.

Jacobsen R. 1988. The persistence of abnormal returns. *Strategic Management Journal* **9**(5): 415–430.

Kalai E, Lehrer E. 1995. Subjective games and equilibria. *Games and Economic Behavior* **8**(1): 123–163.

Knott A. 2003. Persistent heterogeneity and sustainable innovation. *Strategic Management Journal* **24**(8): 687–705.

Kreps DM, Scheinkman JA. 1983. Quantity precommitment and Bertrand competition yield cournot outcomes. *Bell Journal of Economics* **14**(2): 326–337.

Lippman SA, Rumelt RP. 2003. A bargaining perspective on resource advantage. *Strategic Management Journal* **24**(11): 1069–1086.

MacDonald G, Ryall MD. 2004. How do value creation and competition determine whether a firm appropriates value? *Management Science* **50**(10): 1319–1333.

MacDonald G, Ryall MD. 2006. Do new competitors, new customers, new suppliers, … sustain, destroy or create competitive advantage? *Strategic Management Journal* (Forthcoming).

Madsen T, Leiblein M. 2015. What factors affect the persistence of an innovation advantage? *Journal of Management Studies* **52**(8): 1097–1127.

Madsen T, Walker G. 2015. Competitive heterogeneity, cohorts, and persistent advantage. *Strategic Management Journal* (Forthcoming).

Makadok R, Barney JB. 2001. Strategic factor market intelligence: an application of information economics to strategy formulation and competitor intelligence. *Management Science* **47**(12): 1621–1638.

Makowski L, Ostroy JM. 1995. Appropriation and efficiency: a revision of the first theorem of welfare economics. *American Economic Review* **85**(4): 808–827.

Makowski L, Ostroy JM. 2001. Perfect competition and the creativity of the market. *Journal of Economic Literature* **39**(2): 479–535.

Marx M, Gans JS, Hsu D. 2014. Dynamic commercialization strategies for disruptive technologies: evidence from the Speech Recognition Industry. *Management Science* **60**(12): 3103–3123.

McGahan A, Porter M. 1999. The persistence of shocks to profitability. *Review of Economics and Statistics* **81**(1): 143–153.

McGahan A, Porter M. 2003. The emergence and sustainability of abnormal profits. *Strategic Organization* **1**(1): 79–108.

Mindruta D. 2013. Value creation in university-firm research collaborations: a matching approach. *Strategic Management Journal* **34**(6): 644–665.

Montez J, Ruiz-Aliseda F, Ryall MD. (2015). Quantifying the negative – and positive – effects of competitive intensity on firm performance.

Mueller D. 1977. The persistence of profits above the norm. *Economica* **44**(176): 369–380.

Mueller D. 1986. *Profits in the Long Run*. Cambridge University Press: Cambridge, UK.

Obloj T, Zemsky P. 2015. Value creation and value capture under moral hazard: exploring the microfoundations of buyer–supplier relationships. *Strategic Management Journal* **36**: 1146–1163.

Oxley JE, Rivkin JW, Ryall MD, The Strategy Research Initiative. 2010. The strategy research initiative: recognizing and encouraging high-quality research in strategy. *Strategic Organization* **8**(4): 377–386.

Pacheco-de Almeida G, Zemsky P. 2007. The timing of resource development and sustainable competitive advantage. *Management science* **53**(4): 651–666.

Porter ME. 1979. How competitive forces shape strategy. *Harvard Business Review* **57**(2): 137–145.

Porter ME. 1980. *Competitive Strategy: Techniques for Analysing Industries and Competitors*. Free Press: New York.

Postlewaite A, Rosenthal RW. 1974. Disadvantageous syndicates. *Journal of Economic Theory* **9**: 324–326.

Roberts P. 2001. Innovation and firm level persistent profitability: a Schumpeterian framework. *Managerial and Decision Economics* **22**(4–5): 239–250.

Roberts P, Dowling G. 2002. Corporate reputation and sustained superior financial performance. *Strategic Management Journal* **23**(12): 1077–1093.

Rumelt R. 1991. How much does industry matter? *Strategic Management Journal* **12**(3): 167–185.

Ryall MD. 2003. Subjective rationality, self-confirming equilibrium, and corporate strategy. *Management Science* **49**(7): 936–949.

Ryall MD. 2009. Causal ambiguity, complexity, and capability-based advantage. *Management Science* **55**(3): 389–403.

Ryall MD. 2013. The new dynamics of competition. *Harvard Business Review* **91**(6): 80–87.

Ryall MD, Sorenson O. 2007. Brokers and competitive advantage. *Management Science* **53**(4): 566–583.

Schumpeter J. 1934. *The Theory of Eeconomic Development: An Inquiry into Profits, Capital, Credit, Interest, and the Business Cycle* (10th edn). Transaction Publishers: New Brunswick, NJ.

Stuart J. 1997. The supplier-firm-buyer game and its m-sided generalization. *Mathematical Social Sciences* **34**(1): 21–27.

Stuart HW Jr. 2001. Cooperative games and business strategy. In *Game Theory and Business Applications*, Chetterjee K, Samuelson WF (eds). Springer US: New York; 189–213.

Stuart HW. 2005. Biform analysis of inventory competition. *Manufacturing & Service Operations Management* **7**(4): 347–359.

Syverson C. 2011. What determines productivity ? *Journal of Economic Literature* **49**(2): 326–365.

Villalonga B. 2004. Intangible resources, Tobin's q, and sustainability of performance differences. *Journal of Economic Behavior & Organization* **54**(2): 205–230.

Waring G. 1996. Industry differences in the persistence of firm-specific returns. *American Economic Review* **86**(5): 1253–1265.

Wernerfelt B. 1984. A resource-based view of the firm. *Strategic Management Journal* **5**(2): 171–180.

Wiggins R, Ruefli T. 2002. Sustained competitive advantage: temporal dynamics and the incidence and persistence of superior economic performance. *Organization Science* **13**(1): 81–105.

Williamson OE. 1981. The economics of organization: the transaction cost approach. *American Journal of Sociology* **87**(3): 548–577.