Laboratory of Economics and Management
Sant'Anna School of Advanced Studies

Piazza dei Martiri della Libertà, 33 - I-56127 PISA (Italy)
Tel. +39-050-883-341 Fax +39-050-883-344
Email: lem@sssup.it Web Page: http://lem.sssup.it

LEM

Working Paper Series

The Dynamics of Wealth, Profit and Sustainable Advantage

Michael G. Jacobides[1], Sidney G. Winter[2], Stefan M. Kassberger[3]

[1] *London Business School and Advanced Institute for Management Research*
[2] *The Wharton School of the University of Pennsylvania*
[3] *University of Ulm, Germany*

2010/22 **December 2010**

ISSN (online) 2284-0400

The Dynamics of Wealth, Profit and Sustainable Advantage

Michael G. Jacobides*

London Business School and Advanced Institute for Management Research
Sussex Place, Regent's Park, London NW1 4SA, United Kingdom
mjacobides@london.edu; Tel (+ 44 20) 7000 8716; fax (+ 44 20) 7000 8701

Sidney G. Winter

The Wharton School of the University of Pennsylvania
Suite 2000, SH-DH, 3620 Locust Walk, Philadelphia PA 19104-6370, USA
winter@wharton.upenn.edu; Tel (215) 898 4140; fax (215) 898 0401

Stefan M. Kassberger

University of Ulm, Germany
Helmholtzstrasse 18, 89069 Ulm, Germany
stefan.kassberger@uni-ulm.de, Tel (+49) 731 50 23518

December 4, 2010

Working Paper – London Business School

The Dynamics of Wealth, Profit and Sustainable Advantage

Abstract

This paper shows how idiosyncratic resources can be the basis of sustained profitability and persistent heterogeneity under *competitive* conditions: Generic inputs purchased in the market become idiosyncratic resources by investments in customization. Analytically, we show how heterogeneous firms co-exist in equilibrium. Computationally, we show that sustainable profits can emerge without "monopolistic" imperfections. We consider how capability heterogeneity, resource customization cost and ease of expansion interact to drive short-run and sustainable profits. Results illustrate that, in an industry evolution context, sustainable profits may represent a small part of total wealth creation, and that changes in factors shaping a sectors' evolutionary trajectory may be more important than changes in factors that determine profits' ultimate sustainability, thus calling into question the familiar emphasis on "sustainable advantage."

Keywords: Competitive advantage; sustainability; heterogeneity, industry evolution, rent appropriation.

The broad question of how profitability emerges and evolves has long been a central concern in both strategic management and industrial organization research. Recognizing that this broad question presents diverse challenges, researchers have focused on narrower areas. In the process of focusing attention and dividing labor to generate useful insights, existing research has followed a number of trajectories, each yielding important progress. The view of the big picture, however, remains incomplete. The fragmentation of effort has left some important areas unexplored.

In the field of strategic management, the widely accepted "Resource-Based View" (RBV) starts from the hypothesis that firms are different largely because they draw on different resources, which makes them differentially effective (Barney, 1986; Wernerfelt, 1984). With increasing clarity, superior returns at the firm level have become identified with persistent scarcity rents on superior resources. The characteristics that resources must have if they are to yield such persistent returns have been extensively discussed; less attention has been devoted to the factors determining who gets the rents (but, see Coff 1999). This research has concluded that fairly stringent criteria must be fulfilled if firms are to become ("abnormally") profitable in equilibrium (Barney 1991; Peteraf & Barney, 2003). It also sheds light on some of the dynamic processes through which firms might establish the preconditions for capturing persistent returns in equilibrium – e.g., by investing in assets subject to "time compression diseconomies" (Dierickx & Cool, 1989). Almost without exception, however, this theoretical research has focused on how, *in equilibrium,* after the hand of selection has eroded temporary profits, some firms can still be more profitable than others. Guided by the concept of "sustainable advantage," the inquiry has been predicated on the goal of understanding the conditions under which profits can be maintained in equilibrium, and identifying resource characteristics and firm strategies that could be associated with this result.[1]

This still leaves a substantial gap, which this paper aspires to fill. There are two closely related issues to address. First, we extend and clarify the existing analysis of how firm heterogeneity and profitability are linked. The more successful firms tend to grow, and as they do so they typically make continuing investments in their idiosyncratic resources and capabilities. Such investments are the source of the economically significant distinction between generic *inputs,* which any firm can hire or purchase in the marketplace, and the idiosyncratic

[1] In this paper we regard "sustained advantage" as necessary but not sufficient for "sustained profitability," with the question of value appropriation intervening between the two. Our analysis treats both issues of advantage, in the sense of net value creation, and profitability, or value appropriation.

resources that a particular firm uses to do business in its own particular way. The persistence in equilibrium of the *returns on those investments* is, we claim, what "sustained advantage" is about. Sustained advantage translates to sustained abnormal profitability when the conditions of rent appropriation favor the firm over other stakeholders (Coff 1999; Brandenburger and Stuart 1996). Second, we seek to rebalance the theoretical picture of "how firms make money" by giving appropriate emphasis to profitability *out of equilibrium*. We claim that the "appropriate emphasis" is, in fact, a great deal of emphasis. We explicate the theoretical logic of this claim and point to the relevant empirical questions. This second theme is closely connected to our first, because an assessment of the relative importance of transient (out-of-equilibrium) and sustainable (equilibrium) returns must rest on a clear conceptualization of the latter.

These two key issues point to a broader question involving both terminology and techniques of analysis. We argue that the term "wealth creation" provides a more accurate and effective orientation for strategic issues than the term "profit" or any of its variants, and that the simplest analytical approach that is fit for the purpose of exploring wealth creation is the analysis of the net present value (NPV) of cash flows. And this demands a fully dynamic analysis.

To achieve these goals, we frame the analysis with a stylized model of industry evolution. The model includes a market in which firms compete for generic inputs. Firms also compete in the market for the final product, facing a price-sensitive demand, and behave as price-takers. They have varying degrees of capability and make idiosyncratic investments to convert generic inputs into resources, thereby implementing their distinctive approaches to the competitive struggle. Capacity levels are fixed in the short term (a single period), but firms respond to profit signals by adjusting their capacity over time. The extension of capacity by more capable firms is the central mechanism driving (Darwinian) selection and shaping market structure over time.

We derive some limited conclusions from our model by analytical methods, and a much richer set of implications by numerical computations. In the computational experiments, we vary in particular the degree of heterogeneity and the cost of the "customizing" investments that give resources their firm-specific character. We examine how these parameter changes affect wealth creation as well as firm profits and their distribution, and the prices and quantities of both resources and final goods. This analysis is supplemented by a robustness check exploring the influence of the pace of dynamic adjustment and also of oligopolistic rather than strictly competitive behavior.

The results of our modeling and computations provide a clear view of what might reasonably be called the "strong RBV" case – the case where firms successfully appropriate the rents generated by idiosyncratic resources that they do not (and cannot) own. Further, consistent with the efficiency-based logic of the RBV (Peteraf & Barney 2003), the relevant mechanisms operate even when input and output markets are fully competitive. We believe, therefore, that our analysis clarifies the logic of the RBV at a basic level, and does so in a manner fully consistent with its important intuitions relating to competition, firm heterogeneity, and financial performance. This clarification is accomplished via a strong dose of simplifying assumptions; we certainly recognize that there is a much wider area of reality where RBV logic operates but is veiled by a variety of complexities.

We also propose, however, an important qualification to the RBV emphasis on sustainability. We suggest that, when viewed in relation to the total wealth created in the course of industry evolution, sustainable profits may be relatively unimportant. We therefore point to the need for better understanding and measurement of wealth creation and its relation to the dynamic factors that drive industry adjustment paths.

The following section addresses the theoretical background of the investigation, with particular reference to the central concept of "profit." We then describe in turn the motivation for the model, its detailed workings, and the partial analytical solutions that can be derived. Next we turn to the description of the range of situations considered in the illustrative calculations. The final three sections present the results, their interpretations, and the broader implications. In the latter discussion, we address in particular the relationship of our notion of customization cost to previous work by Lippman & Rumelt (1982) and Sutton (1991).

THEORETICAL BACKGROUND: PROFITS, RENTS AND RESOURCES

For all of its centrality, the conceptualization and explanation of profitability remains a complex and challenging region of intellectual territory (Lippman & Rumelt, 2003a). The question "what is profit, and how does it emerge?" receives multiple answers, even within the discipline of economics, and a cacophony of opinion emerges when one looks at the uses of the concept in accounting, finance, and strategic management. This persistent diversity is perhaps attributable to the fact that clarity on what "profit" means is typically needed only "locally," i.e., for the purposes of a particular investigation, and a locally satisfactory solution can often be framed without addressing questions that would arise in a broader context. Unfortunately, for many key analytical tasks in the field of strategic management, such "local" resolutions tend to be inadequate. We argue, more pointedly, that a wealth-creation orientation provides the best intuitive guidance for strategic analysis. The correct top-level question is, "what course of action will increase the wealth of the firm's principals?" To address this question satisfactorily, a broad view of the relevant mechanisms is needed – broader, in particular, than the textbook-level economic analysis of "profit" typically provides.

While this is not the place for a full review of the conceptual complexities and the literature addressing them, we do need to locate our own approach within this broad field of discussion. We begin by noting three important distinctions pertaining to the concept of profit. ("Profits", for the moment, is taken to mean "quantities that someone has termed 'profits'".). First, we can distinguish between equilibrium and disequilibrium conditions, and the different sorts of returns that accrue under each. Second, we can distinguish between profit that emerges when firms use distinctive "informational" goods, which are "non-rivalrous" (such as concepts, ideas or methods)[2] and those where firms are largely homogeneous in informational terms and can draw on inputs that are rivalrous in use and also scarce (such as location, capital, or labor). And third, we can distinguish between competitive and non-competitive situations (the latter focusing on how individual firms can influence prices and thus obtain profits).

The last distinction deserves particular mention, since the analysis of *non-competitive* situations has played a big, even dominant, role in discussions of business strategy. Michael Porter (1980) famously adapted for strategy analysis the structure-conduct-performance paradigm of industrial organization (IO) economics, turning profitability from a symptom of threats to the public interest into an indicator of business success. Subsequent analyses, often with game-theoretic tools, have explored many details of this basic framework, sometimes with great subtlety. Recently, the methods of cooperative game theory have been applied to understand how the returns from an economic opportunity are divided among self-interested players who must cooperate to create value (Dixit & Nalebuff 1991; Brandenburger & Stuart 1996; Lippman & Rumelt 2003b; McDonald & Ryall 2004).

Although the insights gained from the non-competitive branch of profitability analysis have been substantial, they are not central to the present paper. Here, we follow the lead of the

[2] Institutional devices such as patents create property rights in information and thus a form of scarcity that is artificial (in the sense that it is dependent on the institutional context, which is in substantial part the product of intentional human design efforts.) We consider the implications of this later in this section.

many authors who have seen business competition more in terms of a struggle to achieve competitive advantages that derive from superior efficiency (in one sense or another). In Williamson's (1999) terms, our analysis relates to "economizing" as opposed to "strategizing" (Peteraf & Barney, 2003; Teece, Pisano, & Shuen, 1997). The division of opinion on this point in the strategic management literature has important antecedents in the IO literature (Demsetz, 1973; Schmalensee, 1985). While prices remain close to the center of our story, as they obviously must in any analysis of profitability, we adopt here the presumption that is standard in competitive economic analysis: the *individual* firm does not consider itself to have influence over the prices at which it transacts. Collectively, of course, firms do have such influence, as our analysis describes. That being said, our research does consider one form of oligopolistic behavior as an extension of our main analysis. We consider an alternative setup where firms restrain their output expansion as the sector becomes more concentrated, requiring an expectation of higher rates of return to induce them to expand further. We consider some numerical examples that suggest that profits in such an oligopolistic case do not dwarf competitive profits from the wealth-creation standpoint, partly because a substantially concentrated structure takes time to appear.[3]

Our focus is on how profit emerges in such a competitive setting. First, we consider the *equilibrium* profits under competition (and consider how different types of resources or production factors relate to them); and then expand the discussion to include *disequilibrium* profits, focusing in particular on the profits that emerge as a new industry gradually converges toward equilibrium.

Equilibrium Profits Under Competition

In the competitive branch of the literature, the case of competitive equilibrium returns on ordinary resources is by far the most studied and best understood. The familiar tendency is for the profits of the firm itself, i.e., *apart from any ownership stake it may have in inputs*, to be zero. Theoretically, a sufficient condition for this result is that production occurs under constant returns to scale and productive knowledge is a non-excludable public good, i.e., fully imitable. The assumption of constant returns to scale can itself be derived from an axiomatic analysis of

[3] Note that our approach differs from Makadok (2001; 2010) in that our focus on profits takes into account industry dynamics and we do not consider one individual firm at a time. We are in agreement with Makadok's (2010) suggestion that an efficient firm that choosing its output might under-exploit its efficiency differential, while it tries to avoid price erosion. We differ in our analysis on the basis of profitability in terms of "competitive advantage" based on resource ownership (vs use or, in particular, customization.)

production that includes the axiom of additivity (Koopmans, 1957; Mass-Collel, Green, & Whinston, 1995).

Of course, the conclusion of a zero return evaporates if the firm is considered inseparable from some input; for example, if the farmer and the farmer's land together constitute an enterprise, "the farm". In this case, the equilibrium land rent on the (scarce) farm land obviously accrues to the farm-firm (along with farmer wages, perhaps), but there is an equally obvious conceptual question as to whether, or to what extent, such a return is properly called a "profit". This issue continues to create some confusion in the context of the RBV. The phrase "a return in excess of opportunity cost" may *sound* like it describes a profit, but this is misleading. The fact that some resource is producing a return in excess of its current opportunity cost certainly cannot be taken as evidence, or even a suggestion, that somebody is prospering as a result. (See (Lippman & Rumelt, 2003a) for a discussion and apologia on these issues, including a critique of "opportunity cost".) Even more misleading is the practice of referring to accounting net income as profit, since that "profit" includes the normal return to equity investors.

Finally, another possible explanation for sustained profitability concerns the dynamics of *time compression dis-economies* (e.g., Dierickx & Cool, 1989; Pacheco de Almeida & Zemsky, 2006). In this case, firms engage in a protracted course of investment to create a resource or capability that subsequently yields a flow of returns. Again, however, it is not entirely clear whether this return should be should be regarded as a profit or as a rent to a non-tradable resource created along the investment trajectory. This puzzle lies very close to the investment dynamics issues that we subsequently explore.

Transient Profits and Competitive Adjustment Paths

In addition to considering what can lead to profits in long-run equilibrium, we should consider the profits that arise under disequilibrium conditions. Such conditions can arise for various reasons and persist for various periods; we focus here on the quantitatively and historically important case of the appearance of a new industry. The familiar patterns of industry evolution testify to the size and persistence of such disequilibria. (Gort & Klepper, 1982; Klepper, 1997; Klepper & Graddy, 1990; Utterback & Abernathy, 1975; Klepper & Simons, 2000).

Historical examples of industry evolution present a rich array of phenomena, and there is a corresponding diversity in the conceptual and theoretical schemes that describe and explain it. We focus here on a simple story that is at the heart of a number of models in the evolutionary economics literature. In this story, individual firms create production methods by processes of path-dependent learning. Although an underlying fund of public knowledge may provide a common starting point for this process across all firms, the contingencies of path-dependent learning in complex situations make the operational methods significantly different. Once they have achieved their idiosyncratic solutions to the production problem, firms settle on their methods permanently and seek to scale them up when this is profitable. A production method is replicable by the firm that created it (constant returns), but not fully imitable.[4] An "industry" of firms facing the same input and output markets then becomes the arena for the process of economic natural selection. If several firms have overall efficiency levels that are effectively identical and superior to those of other firms, then these efficient firms survive in the long run and the ultimate equilibrium is the familiar competitive one with zero profits.

In equilibrium, there would typically be a set of identically efficient firms with zero profits, just like the textbook equilibrium. Note, however, that this evolutionary story differs significantly from its textbook counterpart. The most important implication is that we have to consider the speed of the selection process. Since the selection process takes time, the efficient firms are typically making profits along the way. Therefore, in terms of the wealth of firm owners, as assessed in terms of the net present value of cash flows, the adjustment process may be more important than the ultimate equilibrium – and clearly *is* more important if that equilibrium is of the traditional zero-profit kind. The "wealth creation" in this process is often largely attributable to the re-investment of cash flows obtained in the early stages, which creates a legacy of real assets generating valuable product – but such assets yield only "normal returns" when prices are ultimately at competitive levels and asset valuations are computed in the economist's standard way, as present values of those very returns. There is another important difference from the equilibrium analysis of the preceding section: the transient returns (as the industry moves toward equilibrium) are typically not entirely rents to ordinary resources; they reflect in part the superior informational resources of more effective firms, i.e., the superior achievements in the initial learning process.[5]

[4] The reason for inimitability, i.e., whether natural or artificial, is not consequential for this analysis as long as there is nothing that can be done to change the situation. Firms may have patents or simply protect their secrets. In many cases they may protect their advantages simply by denying "access to the template," i.e., by not letting others closely inspect what they are doing. (Nelson and Winter 1982; Rivkin, 2001).

[5] It is optional to attribute such superiority to the efforts of a particularly competent entrepreneur, engineer or inventor; what matters is the heterogeneity and the gradual character of the selection process.

Since our interest now shifts to the returns realized as the industry moves from initial disequilibrium, we need to consider how long this might take, and why. There are several possible answers. First, capacity expansion is checked by the limited availability of internal funds – given that capital markets are imperfect and lenders are likely to be skeptical of high debt-equity ratios, especially when the nature of the superiority enjoyed by a particular firm cannot be readily ascertained. Second, irreversibility of investment, in any degree, restrains the pace of selection. It may well be the case that several firms that exist at the inception of a sector are not viable in the long run (cf. Jovanovic, 1982; Lippman & Rumelt, 1982), as historical examples confirm. To the extent that the relevant investments are sunk or only slowly depreciating, the capacity created by such inferior firms hangs over the market and checks the profitability and growth of the superior firms. Indeed, it is entirely possible for non-viable firms not only to survive for a long time, but to be profitable overall during their temporary existence., having made a positive contribution to the process of wealth creation. Third, consistent with evolutionary theory, replication may not be costless (Winter & Szulanski, 2002). The services of the resources of the existing enterprise at a point of time are typically needed as inputs to the replication process (Rubin, 1973), and this creates adjustment costs and time-compression diseconomies that check the rate of growth (Dierickx and Cool, 1989).

If only one or a few firms are particularly efficient, their increasingly dominant role should eventually lead them to exert their power over prices. The hand of selection may create oligopoly or monopoly. In that case the assumption of price-taking becomes increasingly dubious as time passes, and the mode of analysis should in principle jump to the non-competitive branch. While we do not address these dynamics, we do consider briefly how the results would change under one version of oligopolistic behavior. Our results indicate that such a shift of assumptions need not dramatically change the picture.

THE MODEL: MOTIVATION AND CONCEPTUAL ISSUES

Antecedents

In specific respects, the industry model developed here follows the modeling tradition of evolutionary economics (Dosi, Marsili, Orsenigo & Salvatore, 1995; Malerba, Nelson, Orsenigo, & Winter, 1999; Nelson & Winter, 1982; Winter, 1964, 1971), and extends the analytical framework used by Jacobides (2000; 2008). It is a "temporary equilibrium" model in which an individual period represents a Marshallian "short run." In each such period, the interaction of the firms determines key features of their shared environment at that time, particularly prices. The environment then shapes the change of the industry state to the next period, by determining the values of firm state variables for the following period, one firm at a time. In particular, prices determine profits, and profits in turn are a major influence on firm investment behavior. The equations determining investment are understood as behavioral rules that are responsive to the firm's prevailing state and the realized values of prices; we do not consider firms to be endowed with the extensive structural knowledge that would permit them to anticipate (even in probabilistic terms) the long-run evolution of the industry correctly.[6] The iteration of the single-period equilibrium and change processes generates the dynamic path of the industry.

The industry's development is an evolutionary process in which selection plays a key role. Firms are heterogeneous (except in one set of scenarios included for reference purposes). Firms operate under long-run constant returns to scale, and they are capable of replicating their distinctive practices as they grow. In the long run, the more efficient firms expand and put competitive pressure on the less efficient ones, causing them to shrink.

We assume that firms behave as price takers both in the short run and in their investment decisions, so the expansion of the more successful firms is not checked by voluntary restraint derived from perception of the market power associated with a large market share. Given the assumptions of price-taking and constant returns to scale, a legitimate interpretation of the model is that each recognized "firm" actually corresponds to a "firm type" involving many atomistic firms with identical characteristics and behavior. This (optional) interpretation may serve to alleviate any discomfort with the idea that a firm with a substantial market share does not recognize its market power. (The role of market power is considered as a modelling extension, reported below, and we also explore the robustness of the results to a change in the number of model firms.)

As noted previously, in any model with constant returns to scale, price-taking behavior and otherwise standard assumptions, economic profits are zero in long-run equilibrium, and this conclusion holds independent of the number of firms. When our specific assumptions are indeed "standard," the analytical solution for this zero-profit equilibrium is straightforward. Given this

[6] This is the key point that distinguishes this type of model from full rationality (rational expectations) models of industry evolution such as those put forward by Jovanovic (1982) and others. For a rare attempt to anchor a full rationality analysis in the details of a real situation, see Spence and Porter (1982), and the comment by Winter.

familiar foregone conclusion, the long-run equilibrium analysis holds little interest *per se* if the inquiry is motivated by questions about wealth creation and/or the nature of competitive advantage. The illumination that our analysis casts on the latter issues, which are indeed our concern, derives partly from its explicit dynamics, partly from our attention to present value outcomes as opposed to single-period profit measures, and partly from a key departure from standard assumptions (described immediately below). Numerical computation of the dynamic paths of the model, with comparisons across time and across parameter settings, is our primary tool in the dynamic analysis.

Modeling (Idiosyncratic) Resources

The most distinctive feature of this model is its treatment of idiosyncratic resources. To motivate this treatment, we take note of some intrinsic and formidable difficulties that lie in the way of any effort to elucidate the insights of resource-based theory in a formal economic model. Most importantly, there is the question of the mechanism(s) by which a firm's resources are "semi-permanently attached" to it (Wernerfelt, 1984). Ownership by the firm is one important form of attachment, and it is one that is relatively easy to understand. It is clear, however, that the RBV would be quite impoverished if the only recognized forms of attachment were ones that cannot apply to human assets. Not only are human assets a critical part of the mix underpinning capabilities, but associated issues of tacit knowledge and social complexity are often identified as important sources of "inimitability." (Coff, 1997, 1999). It is likewise clear that the actual mechanisms of attachment are quite diverse. If one thinks, for example, of the mechanisms that keep an R&D scientist attached to her job in a particular firm, it is quite clear that these include some quite subtle considerations affecting job satisfaction, personal or family mobility, and the character of the potential alternative offers.

Second, the theory emphasizes resource heterogeneity, or, in other formulations, the idiosyncratic or nearly firm-specific character of resources. This immediately warns of complications affecting the market-level relationships of resources of similar type. Consider, for example, the category of retail locations in a major city – plainly heterogeneous, clearly in the "same market" in some sense, obviously carrying different prices. How does one characterize the price relationships in a simple way appropriate to a formal model, capturing the implications of diverse site attributes? Third, a firm typically draws on the services of numerous resources, and the market processes affecting their valuations outside the firm are complex, interwoven and shaped by idiosyncratic history – consider, for example, the shared experiences that underpin the effectiveness of a team. Lastly, partly because of some of the foregoing considerations, it is quite common that actual market quotations on particular resources are rare events. In the absence of competitive price quotations that pin down the opportunity cost, how do we get any clue as to whether the resource owner or the employer is capturing the rents?

Given these challenges and the hazardous analytical terrain they present, our modeling choices are guided by the objective of making the key economic issues stand out as clearly as possible in a scheme that is only as complex as is needed to accomplish that goal. We assume that there is a single resource, and the level of that resource is the key state variable characterizing a firm in a time period. Seeking to rebalance the picture presented by the many familiar models in which such a central state variable would be an owned asset called "capital" or "capacity," we make the contrasting assumption that the resource is not owned. Rather, it is hired period by period at a price that is endogenously determined in the market, and the obvious interpretation is a trained employee. There is a second input, which we conceive as being merely "supportive" to the focal one. A possible image is that the focal resource is talented professionals of some sort, and the other input is simply the office space – facilities and equipment of a generic type, not owned by the firm, but available every period at a market price in which the industry faces a supply curve for the input. (Because the supporting input lacks any attachment to the individual firm beyond a single period, it would not be consistent with RBV thinking to call it a "resource".)

The nature of "attachment" and its consequences is modeled after the notion of firm-specific training familiar in labor economics (Becker, 1993 [1964]). A firm makes a costly investment to customize the resource, i.e., to convert the generic resource to its own idiosyncratic type, and the benefits of this investment are realizable only within that firm. Thus the conversion does not create outside opportunities and attendant bargaining power that could be used to extract a significant fraction of the benefit the investment produces. On the other hand, by conceding a *trivial* fraction of the rent (modeled as zero), the firm can make its offer superior to the market alternative indefinitely, thus assuring the attachment and avoiding the need to repeat the conversion investment with a resource newly acquired from the marketplace. A "hold-up problem" is sometimes alleged here: once the investment is sunk, the worker can bargain for a piece of the return. We, however, are assuming a situation where a large number of similar workers are potentially available for the training. Credible commitments by the firm can block the hold-up. The formal results depend on the specifics of the game (Brandenburger & Stuart, 1996; McDonald & Ryall, 2004). If the situation involved a single employee with idiosyncratic traits, or a group engaged in collective bargaining, a different result could certainly ensue. (See Coff (1997, 1999, 2010) for a helpful analysis of relevant considerations and firm strategies.)

We abstract from the problems of finite lives and turnover, and make the attachment potentially permanent. We simply assume that the customization investment is a one-time cost that remains effective as long as the resource is available to the firm. A key implication is that the costs of customization investments are borne only when the firm is growing.[7]

Considered at the level of an individual employee or piece of equipment, the assumption that customizing investments are infinitely durable is extreme. The assumption serves, however, to lay bare a logic that still operates as long as these investments, made by a firm as it grows, do not have to be *fully* repeated every time a piece of equipment wears out or an individual leaves the firm.[8]

Firm-level Attributes: Capabilities, Firm Choices and Competitive Interaction

Heterogeneity in production methods is formally represented in a simple, familiar way, by a firm-specific multiplicative constant affecting the firm's production function. Thus, firms are (allowed to be) differentially capable, capability being defined as the efficiency with which a firm turns a set of (customized) resources into outputs. We conceive of these capability differences as arising from the independent exploratory search efforts of firms at the earlier stage of industry history.

Thus the model distinguishes among *inputs* (traded in the market), idiosyncratic *resources* (unique to the firm) and *capabilities*. It also suggests capabilities' link to resources: to replicate its distinctive method at a larger scale, a firm needs to make the customization investment for additional units of the focal resource. It also needs an amount of the supporting input that is appropriate to the chosen scale. Although the supporting input is hired for a single period, we assume that the commitment to its level must be made before prices and actual production levels are known. Thus there is a maximum amount of output that a firm can offer the market in a

[7] There are potential complications (with real world significance) if a firm is fluctuating in its scale of operations and thus has changing needs for its idiosyncratic resources. Can a firm, for example, rehire its former workers after a business cycle slump or a seasonal lull? Our experiments do not actually involve such contingencies, so we can sidestep this issue with a minor technical assumption.

[8] At the expense of some complication, turnover could be admitted to the model while retaining the idea that some of the "training" costs have a one-off character. Isolated, sequential replacements of machines and personnel can be made with the support of a context that is largely unchanged, which greatly reduces the cost of *continuing* to operate an organization effectively at a given scale, compared to the cost of initially achieving that effectiveness.

period, determined by the available stock of the resource and the level of the supporting input. What is open to influence by the market is the quantity of the resource that is actually hired, and the resulting output level. We assume that, in setting the amount of the supporting input, the firm anticipates full utilization of the available resource and chooses the support quantity that would be optimal for that. In terms of the example given, the amount of office space hired for the period is the amount that would be long-run optimal given the present size of its corps of professionals. The utilization of those professionals, but not the office rent, depends on the contracts that come in that period.

A point of contrast with many evolutionary models is that we represent the firm as responding optimally to the price situation that prevails within the individual time period. (Collectively, firm choices interact in the market with resource supply and output demand conditions to determine, jointly, market prices and the specifics of all individual firm decisions.) While this formulation is in some tension with the reliance on a behavioral approach that is generally advocated in evolutionary economics, the contrast is not as sharp as it might appear. First, from the evolutionary viewpoint, a (myopically) “optimal” rule is just another rule or routine, and deliberate optimization is just another candidate process, occasionally a realistic one, that can generate such rules (Nelson & Winter, 1982:. 126-28). Second, prevailing market prices do provide a type of guidance for behavior that is realistically available to firms -- in sharp contrast to knowledge of the future, which is not realistically available.

MODEL: ANALYTICAL STRUCTURE

Firm Decisions and Market Equilibrium in the Short Run

We begin with the development of the more conventional portion of the model framework. The level of the resource is represented by R. Output, Q, is produced according to a production function of the Cobb-Douglas form,

$$Q = a \cdot R^b \cdot K^{1-b} \tag{1}$$

Here, R is the amount of the focal resource and K is the quantity of the supporting input. Firms have different values of the efficiency parameter, a, but the same value of b. For the time being, we analyze the “representative firm” and suppress the firm indexes, as well as the time index.

At the start of a period, a firm has level R of the resource available, reflecting its past decisions to customize units of the generic input. To support this level of the resource, the firm commits to an optimal level K of the supporting input. Denoting by W the price of one unit of R

and by v the price of one unit of K, the amount of the supporting input that is optimal given R satisfies

$$W \cdot R \cdot \frac{1}{K \cdot v} = \frac{b}{1-b} \tag{2}$$

This is the familiar condition for cost minimization when the production function is Cobb-Douglas. We assume that the price v is given exogenously, and without loss of generality that its value is one – the unit of K is "a dollar's worth". This yields

$$K = W \cdot R \cdot \frac{1-b}{b}. \tag{3}$$

With up to R units of the resource available, and expecting to pay W for units actually employed, the firm will purchase/rent this amount of the supporting input.

With K and R determined, the firm faces in period t problem of choosing an optimal level X of actual utilization of the resource:

$$\begin{aligned} &\textit{Maximize } \Pi(t) \\ &\textit{subject to } 0 \le X \le R(t) \end{aligned} \tag{4}$$

where

$$\Pi(t) = P(t) \cdot a \cdot X^b \cdot K(t)^{1-b} - W(t) \cdot X - 1 \cdot K(t) \tag{5}$$

The final constant does not, of course, affect the maximizing choice, but it is by definition part of the maximand. We call the optimal X in this problem $R^*(t)$, and the resulting optimized level of operating profit is $\Pi^*(t)$ The optimization yields the short run input demand curve, relating the firm's resource use to prevailing prices and the extent of input usage within that period,

$$R^*(t) = min\left\{ \left(a \cdot b \cdot \frac{P(t)}{W(t)} \right)^{\frac{1}{1-b}} \cdot K(t), R(t) \right\} \tag{6}$$

On the output side, the firms face the market demand curve

$$Q = D \cdot P^{-\varepsilon_D} \tag{7}$$

Market clearing in the short run involves equating supply and demand in the two markets, one for the resource and one for the final product. At this point, clarity is served by making firm-specificity explicit:

$$S \cdot (W - W_0)^{\varepsilon_S} = \sum_j R_j^*(t) \tag{8a}$$

$$\sum_j a_j \cdot R_j^*(t)^b \cdot K_j(t)^{1-b} = D \cdot P^{-\varepsilon_D} \qquad (8b)$$

After substituting in from (6), equations (8a-b) become two equations in the two prices *P* and *W*; the solutions are *P(t)* and *W(t)*.

In sum, the logic of each period starts from the firms' available amounts of their specific resources *R,* which are a consequence of their past decisions. They plan to complement their resources with long-run optimal levels of the supporting input. The short-run market process produces a joint determination of actual levels R^* and the market prices *P* and *W*. The two equations above then determine the input and output prices for the period and also individual firm resource levels, hence outputs and profits.

Model Dynamics

The linkage between periods is provided by assumptions regarding expectations and decision rules affecting the choice of *R(t+1)*. For every additional unit of the resource that the firm wishes to have available in *t+1*, beyond the amount available in *t*, the firm incurs a cost in the amount *C*, a parameter we call "customization cost." Such costs are incurred only in the anticipation that the investment will prove profitable, but of course the future is unknown when the investment is made. Firms need to decide whether it would be profitable to expand; and weigh that against their required rate of return *i* when they do so. All firms face the same discount and required rate of return; and they do not differ in terms of the efficiency with which they raise capital or engage in customization (in contrast to Kato, 2010).

First, firms need to assess whether the expansion will be profitable. In the real world, the powerful dynamics attending early-stage industry evolution create a setting where volatile expectations can become a source of instability. Although numerous historical examples testify to this possibility, we consider this set of issues to be extraneous to the present inquiry and have put it aside. Our assumptions about expectations are chosen to yield orderly, stable adjustment paths, are highly consistent with our long run equilibrium analysis, and have the virtue of simplicity.

We assume that firms assess the trend in variable profit per unit of the resource, extrapolate that trend for a single period, and assume constancy after that. This assessment tends to be optimistic with respect to the long run, since profitability is declining. By contrast, the one-period extrapolation tends to be pessimistic because the rate of the decline is itself diminishing.

As time passes, the pace of change drops, and both of these biases decline in magnitude until, at the threshold of equilibrium, the estimates are essentially accurate. The threshold of equilibrium is precisely where it matters that they be accurate, from the viewpoint of industry-level outcomes, since those expectations shape the last investment decisions as the system closes in on equilibrium.

The one-period extrapolation is derived from an "adaptive expectations" mechanism that assessed the trend in "variable profit per unit of the resource", denoted by $V_j(t)$ – i.e., operating profit with the fixed cost of the supporting input added back in, divided by $R^*(t)$.

$$V_j(t) = \left[\Pi^*_j(t) + 1 \cdot K_j(t)\right] / R^*_j(t) = P(t) \cdot a_j \cdot \left[R^*_j(t) / K_j(t)\right]^{b-1} - W(t) \qquad (9)$$

This quantity is always non-negative by virtue of the simultaneity of profit maximization and market clearing. Our experiments show that it provides a good basis for the expectations governing the investment decisions. In effect, the firm calculates the expected change of variable profit per unit as a weighted average of recent percentage changes, with geometrically declining weights. The details of this are more cumbersome than the concept, and are consigned to an appendix.

To establish the profitability test for an incremental investment in R_j, which we denote by $T_j(t)$, we take into account variable profit, but also the fixed costs per unit, the customization cost, and time discounting, as well as the expectation of percentage changes in unit profit, which we term z_j [9] As is described further in the appendix, firms extrapolate the recent changes one period forward to estimate changes in future profitability. Specifically, the investment profitability test given the data of time t is based on the sign of the expression

$$T_j(t) = \frac{(1 + z_j)}{i} \cdot V_j(t) - \frac{(1 - b) \cdot W(t)}{i \cdot b} - C \qquad (10)$$

That is, $T_j(t)$ is the excess of the present value of the future rents per unit earned per unit of the resource over the cost of the customizing investment. Note that $W(t)$ enters this expression explicitly by virtue of its role in determining the level of the supporting input (equation (3)), and we assume that the realized period t value is extrapolated for this purpose.[10] A positive value of $T_j(t)$ indicates an attractive investment in present-value terms, given our specific assumption as

[9] The interest rate matters in our model because it describes the returns available on alternative investments opportunities available to the firm or its investors.
[10] In general, an extrapolation scheme for W would be appropriate here. However, in our computations we take W to be a constant, and choose here to avoid the additional complexity introduced by such an extrapolation.

to how firms evaluate the returns to be expected in a future that is unknown at the time the investment is made.

When the test $T_j(t)$ is passed, the next question is how many new units of the resource should be customized. We assume that this question is answered by reference to the amount of cash available to cover the customization investment; the level of investment is determined by the ratio of available cash to customization cost per unit of resource. As the customization cost is considered to be incurred at the start of a period, contemporaneously with the commitment to *K* for that period, the net cash flow of period *t* is

$$\Pi_j^n(t) = \Pi_j^*(t) - C \cdot max\left(0, R_j(t) - R_j(t-1)\right) \tag{11}$$

The investment in customization is irreversible, so no cash is released when the *R* level declines. For reasons discussed subsequently, we assume that only a fraction of this total is applied to customization. To determine the actual number of resource units customized, we divide the dollar investment by the customization cost per unit, *C*.

$$R_j(t+1) = R_j(t) + \begin{cases} max\left(0,\ \theta \cdot \dfrac{\Pi_j^n(t)}{C}\right), & T_j(t) > 0 \\ 0, & T_j(t) \le 0, R_j^*(t) = R_j(t) \end{cases} . \tag{12}$$

$0 \le \theta \le 1$ is a parameter controlling the fraction of the cash flow available for investment.

On the other hand, if the investment test is not passed, and also a firm did not even use up all of its customized resource, i.e. $R_j^*(t) < R_j(t)$, it is assumed that a fraction λ of the excess resource availability is permanently lost to the firm in the next period. (This might reflect, for example, changing expectations of resource owners about the prospects of further work with the particular firm, or the dwindling of commitments to its particular locale.) In that case,

$$R_j(t+1) = R_j(t) - \lambda \cdot (R_j(t) - R_j^*(t)) \tag{13}$$

Figure 1 summarizes the model; and Table 1 provides a glossary of the variables used.

Insert Figure 1 and Table 1 around here

ANALYTIC RESULTS

Before turning to computational experiments with the model, we examine special cases in which the character of the long-run equilibrium can be explored analytically. Generally speaking, we have exact analytical results only for the cases in which all firms are identical. However, when customization cost C is zero, we know that the only firms that survive in the long run are ones that all have the highest efficiency level, regardless of the initial heterogeneity, for selection always entails homogeneity in the long run. By contrast, when C is positive, firms with differing efficiencies can coexist in the long run, as we will demonstrate. The mix of firms that co-exist in the long run is, however, a path-dependent, historical phenomenon. Our analysis can illuminate the extent of the potential indeterminacy in the mix of survivors, but no general conclusions beyond that are possible.

Equilibrium with Identical Firms

The analytic results we can derive relate to firm behaviour as the sector converges to long-run equilibrium. We consider the case of homogeneous firms, and begin the analytical exercise by supposing hypothetically (contrary to the dynamic model) that the total availability of the resource to the industry is fixed, and the level of the supporting input is appropriately matched to it. On that assumption, the return to the resource is a Ricardian rent in the strictest sense. What would be the equilibrium level of that rent, per unit of the resource, and how would that rent vary with the assumed total availability? As may be obvious, the smaller the postulated level of resource availability, the smaller the level of production and the higher the final good price – and the higher the rent per unit assigned to the resource. Now compare that rent with what the actual supply curve of the resource would yield as the supply price at the assumed resource level. The difference between the two is the true rent per unit that accrues to the firms when resource use is at the assumed level.

Following this intuitive logic, the rent schedule construction is straightforward. Given a value of R, we consult the supply curve of the resource to identify the corresponding W. That value, in conjunction with the exogenously given price of the supporting input, determines optimal input proportions between R and K. With those proportions known, the production function tells us the output quantity Q that corresponds to the initial R. Consulting the demand function, or rather its inverse, we find the output price P. The value of output and its cost (excluding rent) are now determined, the difference between the two is the total rent, and that quantity divided by R is the unit rent value for that R.

Following this path analytically, we first consider the supply-price curve $s(R)$

$$W = s(R) \tag{14}$$

Invoking equations (3) and (1) above, we find that long-run quantity is related to R by

$$Q = A(R) \cdot R \tag{15a}$$

where

$$A(R) = a \cdot \left(s(R) \cdot \frac{1-b}{b} \right)^{1-b} \tag{15b}$$

Note that a relationship of the form of (15a) would prevail for any production function displaying constant returns, but the particular Cobb-Douglas specification is obviously involved in (15b).

Let the demand-price function be *h(Q)*:

$$P = h(Q) \tag{16}$$

Then the desired result, the rent per unit (W_V), is obtained as

$$W_V = \left(Q \cdot h(Q) - W \cdot R - 1 \cdot W \cdot R \cdot \frac{1-b}{b} \right) / R = \frac{Q \cdot h(Q)}{R} - W - W \cdot \frac{1-b}{b} \tag{17a}$$

Here, equation (3) has been invoked again, while (15a) and (15b) are implicit. This gives us the rent value corresponding to a given *R*.

If there were *nothing* to halt firms from expanding, this expression would be zero in the ultimate equilibrium. In this case, we could substitute zero at the left hand-side, and thus define the equilibrium relationship between *Q,* and *R.* In general, however, there is a cost to expanding; and this cost is a function of the customization cost *C* and of the interest rate *i.* The prospect of a perpetual rent stream in excess of $i \cdot C$ is enough to motivate the investment *C* in creating another unit of the resource. So, barring any other influences on firm growth (such as deliberate output restraint), we can match the benefit from the use of resource to the cost of creating it by customizing generic input. We can do so analytically (provided all firms are identical). So, invoking (15a), *R* conveniently divides out, and (17a) can be re-written as

$$W_V = A(R) \cdot h(A(R) \cdot R) - s(R) / b \tag{17b}$$

For given values of interest rate *i* and customization cost C, the equation

$$A(R) \cdot h(A(R) \cdot R) - s(R) / b = i \cdot C \tag{18}$$

thus determines the equilibrium value of *R* (and hence all other variables) in the qualified sense described above – specifically, this expresses the relevant marginal condition for capacity *expansion*. Note the particularly simple form that appears if *s(R)* is a constant, with the result that *A(R)* is as well.[11]

Co-existence of Differentially Capable Firms in Equilibrium

While the rent schedule provides an analytical solution for identical firms, it cannot tell us where the system will equilibrate when firms do not have homogeneous capabilities. However, it can provide a different and quite valuable insight, by helping us understand when differentially efficient firms *might* co-exist in equilibrium. To do so, we have to consider a situation with any two firms with unequal capability (where $a > a'$) and see the conditions under which firm a would not want to expand, and a' does not want to shrink, if it is in the sector already and has sunk the customization cost investment for its current scale. The analysis described above still holds as it relates to the most efficient firm, with capability *a*. What we know for sure is that this firm cannot be in equilibrium if its prevailing rent rate exceeds the value that would give it an incentive to expand its capacity, investing in customizing the resource. We also know that none of the less efficient firms will have an incentive to invest when the most efficient firm is getting close to zero incentive. The open question is when a less efficient firm has an incentive to reduce output, which would have the effect of increasing the profitability of the most efficient firm and re-inspiring it to grow. In effect there is a locus in the *(W , P)* space where equilibrium might occur, if it is a case where the asymptotic *R* value of the most efficient firm is approached (gradually) from below. This locus can be obtained by modifying (18), replacing *h( )* by P and *s( )* by *W*, and taking the *a* value corresponding to the most efficient firm:

$$a \cdot \left(W \cdot \frac{1-b}{b} \right)^{1-b} \cdot P - \frac{W}{b} = i \cdot C \tag{19}$$

A less efficient firm that has already sunk an investment in capability will not have the incentive to contract as long as it can simply cover its continuing costs, perhaps with room to spare. This requires that its efficiency parameter, *a'*, satisfies

$$a' \cdot \left(W \cdot \frac{1-b}{b} \right)^{1-b} \cdot P - \frac{W}{b} \geq 0 \tag{20}$$

[11] Standard comparative statics techniques can be employed on Equation (18) – but subject to an important qualification. If an equilibrium achieved by the dynamic system is disrupted by a decrease in *C*, then (18) will hold in the new equilibrium at a higher *R*. The standard techniques will therefore be useful in characterizing it. But suppose the disruption is from an *increase* in *C*. Then, since the customizing investments are already sunk, *R* does not decrease as (18) would imply. In that case and equilibrium is established at a higher level of activity than would have occurred if the higher *C* value had been prevailing all along.

Moving W/b to the right side in (19) and (20) and dividing yields the "coexistence condition" for a', namely

$$\frac{a'}{a} > \frac{W}{W + i \cdot b \cdot C} \tag{21}$$

We can further simplify this expression. Given that we have expressed C as an absolute cost value, it is easy to scale it as a fraction of the prevailing price for the resource W, and set c as the "cost to invest in a unit of resource as a fraction of the cost of the resource itself" (which is a more intuitive measure). Since $C = W \cdot c$, W then conveniently drops out from this fraction,[12] yielding

$$\frac{a'}{a} > \frac{1}{1 + i \cdot b \cdot c} \tag{22}$$

This equation presents an important result. It shows that given the efficiency level of the most efficient firm, survival prospects for less efficient firms tend to be an increasing function of the customization cost, the interest rate, and the resource elasticity of output. In other words, it shows that firms of somewhat inferior skills *might well persist, even in long-run equilibrium* (as long as they happened to appear early on, and made the appropriate sunk investments, and were not too inefficient). This provides an evolutionary explanation for heterogeneity in capabilities even in long-run "selection equilibrium". It also shows that the extent of dispersion that can be sustained in equilibrium is a function of the extent to which it is difficult to convert "generic inputs" into "idiosyncratic resources," as well as of the interest rate, and of the relative role of these "idiosyncratic" resources in total output. This provides a new twist to the intuition that in "commoditized" sectors, differentially capable firms cannot co-exist in equilibrium and no profit can be had: the explanation here suggests that the lack of *input customization,* as opposed to lack of product differentiation, can account for competitive intensity.

Note that this coexistence condition does not involve the relative sizes of the differentially capable firms. Thus, the equilibrium logic does not offer a reason why the most efficient firm should not turn out to be the smallest surviving firm in the industry. While there are good reasons to think that result unlikely, they are historical, path-dependent reasons, not equilibrium

[12] This result is somewhat over-simplified, as W is endogenous to our model, whereas C has been set as a fixed cost, not dependent on price. Thus, c cannot in general be thought of as a true constant. Obviously, (22) is exact when the resource supply is infinitely elastic, so W is a given constant. And it is a good approximation when s(R) is nearly constant in the relevant region. Our computations assume it is exactly constant.

reasons.[13] In a real-world industry evolution setting, one might well imagine that a very late entrant could have the best methods – but large incumbents have sunk their customization investments, and they are not going to pull back just out of respect for the young winner of the "efficiency Olympics."

While the foregoing provides an interesting link between RBV theorizing and equilibrium analysis, it does not give anything like a full picture of how *C,* or capability dispersion, affects not only the *coexistence* of firms, but also profits, resource prices, and quantities. More important, this calculus leaves the important issues of *paths of adjustment, wealth creation* and the industry dynamics we discussed earlier entirely in the dark. For us to consider them, we need to shift from closed-form solution to numerical investigation, creating an appropriate experimental design to explore the implications of the model, and then return to examine the theoretical implications.

EXPERIMENTAL DESIGN

Setup and Parameterization

For calculation of illustrative adjustment paths our computational techniques parallel those used in (Jacobides 2000; 2008). A common background is posited for these calculations, which we interpret as a highly stylized representation of the processes of industry evolution. For guidance in the choice of parameter values characterizing this background, we draw on the extensive literature that has studied the characteristic historical patterns that mark these processes (Gort & Klepper, 1982; Klepper, 1997; Klepper & Graddy, 1990; Utterback & Abernathy, 1975). In these studies, primary attention has been given to technological change, entry and exit, and the evolution of industrial structure. What we draw from them is, first, a general sense of relevant timeframes. Major industries show significant evolutionary change over periods of decades. Dramatic episodes lasting a decade or less are common in smaller sectors, and within the lengthier trajectories of large industrial sectors. Consider, for example, the expansion of the U.S. telephone industry from about 100 firms in 1894, when the Bell patents expired, to over 9000 organizations in 1902 (Barnett, 1995); or consider, more recently, the rapid evolution of the disk

[13] Note that, in contrast with some models of industry evolution (e.g., Kato, 2010), and in keeping with micro-econometric productivity studies (e.g. Chew et al, 1990), we do not assume that scale itself confers any absolute advantage in terms of either production, or investment / growth. Our analysis is thus in the spirit of Demsetz (1973), where efficiency begets scale rather than scale begetting efficiency (or profitability / concentration). Our formulation is consistent with long-run CRS as well as with short-run DRS (as observed in practice and expected in theory).

drive industry from 1975 to 1989 (Christensen, Suarez, & Utterback, 1998). Second, this literature documents the powerful role of firm heterogeneity in these stories, and recent contributions provide plausible interpretations of the origins of that heterogeneity in terms of firm-level learning and the historical descent of capabilities (Helfat & Lieberman, 2002; Levinthal, 1997). [14]

These various facts provide part of the empirical basis for the background conditions specified in our experiments. Regarding profitability and wealth, we are less well endowed with quantitative research. It has long been clear, however, that great fortunes are founded in the early years of great industries. We also know that outstanding stock market performance is strongly associated with "newness," and over an industry's history is often closely related to the rate of entry (Foster & Kaplan, 2001). The assumptions we now introduce represent our effort to characterize, in a highly stylized context, a situation consistent with the broad implications of the historical record. More systematic use of that record is a significant item on the agenda for future work.

In all cases, we posit an initial state representing an early stage in the process of industry evolution. Firms have already developed their production methods, but industry resource use is a modest fraction of the equilibrium values. The initial resource levels are constant across settings and scenarios. Initial resource use is the same across cases, and in the base case is roughly a third of its equilibrium value. Increasing C makes equilibrium output smaller and in this sense implies that the initial condition is closer to equilibrium. We compute the model outcomes on the basis of quarterly periods, which avoids computational stability issues that can arise when the system is changing quickly. We extend the computation for 250 quarterly periods, to allow the slower-moving equilibration processes the time to produce something close to their asymptotic results. (For example, in the scenarios with low but positive heterogeneity, it takes a long time for the modest efficiency differences to express themselves in firm-size differences.) In most cases, the asymptotic equilibrium is closely approximated by period 100; in a few cases it takes until about period 140. By period 250, all our scenarios are essentially at equilibrium.

---

[14] In the strategic management literature, the significance of firm heterogeneity in various characteristics and in performance has long been recognized (Nelson 1994; Rumelt, Schendel & Teece 1994). Regarding productivity differences in particular – the locus of heterogeneity in our computations – a substantial body of large-sample evidence has been developed in economics. See, for example, (Bloom & Van Reenen 2007), and references cited therein.

The profitability conditions of the final period are then extrapolated to infinity for the purposes of present-value calculations, providing assurance that there are no distortions arising from arbitrary truncation of the cash flow series. At reasonable discount rates, however, this adjustment is of negligible consequence. As we discuss below, this observation reflects an important qualification to the strategic significance of "sustained competitive advantage," if we take the shareholder's perspective and consider an industry evolution setting. For our discount rate we use a quarterly rate of 1.8245%, which compounds to an annual of 7.5%, and reasonably approximates the long-run real rate of return on U.S. equities. At that rate, the time zero present value of 100 dollars at period 101 (start of year 26) is $16.10.

Experimental Design: Customization Cost Settings, Heterogeneity Scenarios

We vary the experimental conditions in three different ways, summarized in Table 2. There are three values for customization cost *C*, of which the lowest is effectively zero and the largest produces an equilibrium rent of approximately the same magnitude as the price of the generic resource. Our "base case" is characterized by the intermediate *C* value. There is empirical grounding for the general plausibility of this value. To judge this plausibility, the appropriate focus is ratio comparisons of implied equilibrium profits (as disclosed by our computations) relative to the wage bill, or to revenue. For each of the three *C* values, there are 11 heterogeneity "scenarios". These range from no heterogeneity to a case where the efficiency ratio between the best and worst firms is three to one. The specific formulation is that there are 11 firms (or types) spread uniformly across an interval centered on $a = 1$, and ranging from $1- x/20$ to $1+x/20$, for $x = 0, 1, 2, \ldots, 10$. [15]

For this first set of experiments with the model, we chose some key parameters with an eye to the interpretability of results, and also to assure computational stability. The elasticity of demand is set to one, which means that revenue is a constant and thus not an active factor in the determination of the industry profitability picture. The supply curve of the generic version of the resource is chosen to be perfectly highly elastic, i.e., *W* is constant.[16]

[15] These are the values on the annual basis. For the quarterly values on which we compute, scale by 0.25.

[16] The Cobb-Douglas exponent on the resource is 0.8, reflecting our intention to give the resource the dominant role in the picture, relative to the generic "supporting" input. At low output prices, the Cobb-Douglas form implies that very high resource productivity is achieved at very low utilization rates of the supporting input. The result is that a firm with positive capacity – and all firms are always in that condition given that at most a fraction of the resource disappears each period – always chooses a positive output level. Thus inefficient firms "fade away" from the marketplace; they do not disappear abruptly. This is another formulation that is not recommended on the grounds of realism, but rather because it is convenient in an exploratory exercise that does not allow for continuing entry and thus has no mechanism to balance the effects of permanent exit on the industry dynamics.

Also highly relevant to the model results is our specific assumption concerning investment finance. When capability increase passes the profitability test (eqn. (11)), it is constrained to be a fraction of the level of net profit, i.e., the current operating profit less the previous amount of customization expenses (eqn. (12)). This simple formulation is intended to capture the reality of a feedback from profitability to growth as a factor in industry dynamics, not to model the investment decision realistically. Firms might obviously spend beyond the limits of internal finance, by borrowing. On the other hand, given reliance on internal financing, they would spend less if capability increase involved investment costs other than those of customizing the resource – e.g., if the ("supporting") plant and equipment had to be owned rather than rented, as is realistic and commonly assumed, but not central to our story. Or there might be other adjustment costs of some sort, or perhaps a dividend payout. The latter sorts of considerations lie behind our choice of 0.5 as the fraction of net profit invested. The important thing, again, is to acknowledge the dynamic linkage between profitability and investment. The specific assumption yields a plausible time scale of events, and it lies on an interpretable dividing line between those considerations that point to higher investment levels and those pointing in the opposite direction. We leave to the future the exploration of alternative assumptions either way. We note, however, that the effect of this assumption is particularly powerful when customization cost is negligible, as in that condition the only limit on firm growth, the financing constraint, is barely operative and strong disequilibrium disappears quickly. Since it is fundamental to our approach that customization cost is *not* zero in interesting and relevant cases, the distorting effect when C is nearly 0 does not present a flaw in our results.

Extension: Considering Imperfect Competition

As we emphasized in the introduction, the paper focuses primarily on *efficiency profits,* as opposed to profits related to output constraint and price manipulation. As a robustness check, however, we explored one alternative version assuming that firms take a more strategic approach to decisions about capacity expansion, and avoid expanding in a way that would destroy their margins. Specifically, we postulate that all firms in the industry attend to the industry structure and apply higher hurdle rates to investments as concentration increases. As a result, a higher concentration leads to less capacity and thus to greater output constraint. This formulation has the advantage of capturing the main consequence of imperfect competition, at least for some part of the wide range of behaviorally plausible conditions. Specifically, in our "oligopolistic behavior" setup we set the required rate of return a firm uses in its expansion calculus to be i' = $i \cdot (1 + 2 \cdot H)$, where H is the prevailing value of the Herfindahl index.[17] So, given that we set i to 7.5% on an annual basis, this means that for a heavily concentrated industry with a Herfindahl of 0.5, the expected return for *all* firms competing would be 15% – double the return in a totally competitive sector, and a monopolist (with $H=1$) would expect a return of 22.5% before they expand. We re-ran all of our scenarios in this "oligopolistic expansion" mode, and we report the differences for some key scenarios that illustrate the intuitions derived from these results. Table 3 provides a summary of parameters, permutations and robustness checks undertaken.

As a further robustness check, we investigated the question of whether a sample of 11 firms provides an adequate approximation of competitive conditions. We considered the impact of an increase in the number of firms (or representative firm categories) from 11 to 21 firms for each of our 11 scenarios. To maintain comparability, we adjusted the initial size of each firm, thus keeping the sector size to 32% of its equilibrium value. The issue explored here is analogous to the choice of an interval in a numerical integration computation – how small does the interval have to be before further reduction is of negligible consequence?

Insert Tables 2 and 3 about here

RESULTS

We rely primarily on graphical methods to present our results, since these convey a strong sense of the qualitative features that are the main interest. Since our model is deterministic, there is no need for multiple runs to provide a clear picture of the model's logic.[18] We rely particularly on a 3D display with the value for a particular variable shown over a plane; the two dimensions of the plane are the time period and the heterogeneity scenario (0 = no heterogeneity, 10 = maximum heterogeneity). While we run the model for 250 quarterly periods, almost all the action happens between periods 1 and 100, except in the settings with the highest customization

[17] We allow for compounding when adjusting to the quarterly basis.

[18] There are minor, quasi-random fluctuations visible in the model output. These are a reflection of its non-linear dynamics and particularly of the discontinuity in investment behavior created by the profitability test. The details and timing of these fluctuations would in some cases shift noticeably with a very small change in parameter values or initial conditions, but not to an extent that would affect our conclusions. The closest thing to an exception is the small effect seen in the zero heterogeneity condition, where the discreteness of the final expansion decision, and its sensitivity to timing, produces results barely visible in our charts. This happens because all firms are doing precisely the same thing.

cost. Our graphs provide the results of periods 1 to 100, and then rather than continuing smoothly with period 101 they show periods 246 to 250 – to show how far convergence has happened and allow us to consider the results "further out" in the future, without overly compressing the periods "where the action is". The results thus displayed include some key industry aggregates and descriptive statistics, as well as variables that actually enter the model logic. We describe first the results with our basic competitive model, and then consider the "oligopolistic expansion" extension noted above, as well as other robustness checks.

Industry Evolution: Profits, Prices, Quantities, Dispersion

The familiar logics of industry evolution are visible in common features of the computed industry trajectories; these form a qualitative picture that is the common background for the experimental effects that are of central interest. Figures 2a–2e show the trajectories, in the base case[19], for output price, total operating profit, the average capability in use, the variance of capability, and total resource use. The computations for average capability in use and variance use input share weights.

In the first two panels, the most obvious feature is the decline in price and profitability associated with the approach to equilibrium from an initial condition at far below equilibrium capacity. Growth decelerates as profits fall, and the transition to a near-equilibrium condition is less abrupt when firms are different and thus stop growing at different times. Note that the operating profit total is clearly approaching an asymptotic value that is above zero; this is the rent consequence of the customization cost and receives more extensive comment below. In panel 2c, the effects of heterogeneity and selection are portrayed. The average (share-weighted) resource productivity increases with heterogeneity, and to an increasing extent over time, as selection has time to enhance the market shares of the most efficient firms. Panel 2d shows the decline of the share-weighted variance of productivity as share becomes increasingly concentrated in the most productive firms. Note, however, that the variance does not appear to be approaching zero. This is a consequence of the long-term coexistence of firms of different efficiency, a possibility discussed above in abstract terms (see eqn. (21) and related discussion).

Insert Figures 2a to 2e about here

[19] Our baseline has a modest customization cost *C* of 10 which implies a quarterly interest cost of 0.18245. The price of a quarter's service of the input is 1.75, so customization increases input cost by 10.4% on a continuing basis. For a customization cost of 100, customization costs about as much as the market price of the generic input. We also ran our results with a very low customization cost (*C=1*) and have the full set of results available upon request.

These basic features are accompanied by others, in a reasonably straightforward way. Total resource use rises steeply in the initial phase, but then stabilizes as total output growth slows. While output grows and price declines continue as a result of relatively slow selection effects, resource use is effectively constant. This is because, with unitary demand elasticity, and price hanging close to average cost, the decline of output price exactly offsets the effect of rising productivity on input demand. Figure 2e shows the amount of total resource use, and confirms this. We leave for future research the analysis of how higher or lower elasticity of demand affects industry profits and resource prices.

Sustainable Advantage and Persistent Profits: Causes and Consequences

As implied by our theoretical discussion above, the profitability test for capability investment ultimately checks the growth of capacity and output, and produces an industry equilibrium situation in which price exceeds the continuing production costs for the most efficient firms. The effects of increasing customization cost are seem most clearly in the contrast between our *C=10* baseline case and the *C=100* case, which are shown in Figures 3a–3e. Comparing 3a to 2a, we see that the final value of the output price is increased by over 68% in all heterogeneity conditions.

Equilibrium output is correspondingly reduced; nevertheless, equilibrium is approached much less rapidly because of the slowing of adjustment associated with the higher customization cost. At equilibrium, the operating profit (or rent) is on the order 37–45% of revenue, as against 5–7% in the base case, with variation across scenarios. Operating profits shown in 3b are high, but recall that in making investment decisions, firms compare this rent with the value of $C \cdot i$ per unit of the resource, as discussed in the analytical exposition. At the margin of the capability investment, returns are close to "normal" in all cases.

Insert Figures 3a, 3b, 3c, 3d and 3e about here

Coexistence of Firms in Equilibrium, and Dynamic Paths

Also as analyzed above, positive customization costs promote the long-term coexistence of firms of different efficiencies. When such costs are negligible (*C=1*), the familiar selection logic is fully in effect. The output share of the most efficient firm in the highest heterogeneity scenario has reached 100% by period 91. In the corresponding scenario for the base case, the top *two* firms are headed for long-term survival, with output shares of 72% and 28%. And in the high customization cost case, it is *seven* firms with shares from 38% down to 4%.

The important role of customization costs in affecting long-term industry dynamics can also be confirmed by looking at the Herfindahl concentration indexes in the two cases – *C=10* and *C=100*, reported in Figures 4a and 4b respectively. In the former, the two most effective firms expand, taking the entire market. In the latter, concentration increases very gradually,, as even the most efficient firms are checked in their expansion. From about period 200, the index is essentially level at a value slightly below .23, implying a "numbers equivalent" of less than five firms. Thus, the asymptotic situation is one of relatively loose oligopoly – but under our assumption of price-taking, it is one in which firms exert no market power.

Wealth Creation: NPV of Cash Flows

We now turn to the important observation that, from the wealth creation viewpoint, sustained profitability plays a minor role in creating the net present value shown by the trajectory as a whole. For instance, consider the contribution of equilibrium profits (considered in all scenarios at period 250, which is a close approximation to the asymptotic result), as a proportion of the full NPV. Table 4 provides this fraction for *C*= 10 and 100 in its second row, where we extrapolate the equilibrium profits to infinity for the purposes of the NPV calculation. We provide the figures for scenario 5 (medium heterogeneity), across different settings. (Results are robust to scenario choice.) As we can see, the fraction is pretty negligible; even for *C*=100, it is 0.17% in terms of appropriately discounted figures.

Some might argue, as an objection to this comparison, that it is simply the result of discounting returns far off in the future. Strictly speaking (and more simply still) it is the result of allowing consistently for the time value of money in all comparisons across time. The results would be the same if the present values were taken as of period *N*, since this involves multiplying the numerator and denominator by the same (large) constant 1.018245^N. And this is precisely the point – the emphasis in the strategy literature on sustainable returns "sometime in the future" tends to produce a rather elementary distortion of economic perspective, in many cases implying a radical understatement of the importance of the adjustment path in the wealth picture. Needless to say, correcting that distortion does not involve abandoning the insight that a profitable position that is sustainable is more valuable than one that is unsustainable. Also, if one "tunes in" only toward the end of industry evolution, the cross-sectional differences in sustainable profits may look impressive -- much more impressive than they do in the broader perspective of wealth creation. This observation arguably goes to the heart of the question we raise about the appropriateness of the focus on flow profitability favored in the literature of the RBV.

We can complement this analysis of the relative importance of sustainable profits by giving such profits a somewhat artificial advantage in the comparison. In the third row of Table 4, we compute present values for sustainable profits on the assumption that the ultimately sustainable portion counts as “sustainable” from the start. If the point of the usual emphasis on “sustainability” is to direct attention to strategic actions that promote it, these early-stage profits should not count, for they are realized without any such action. We compare this expanded measure with the total NPV generated with the far-from-equilibrium start, counting the equilibrium “share” of the profits from the first period. Even with this criterion, only in the C=100 case do sustainable profits account for the majority of the total, and this is partly because (as noted previously) the industry starting position is closer to its ultimate equilibrium when customization cost is high. In that case, the equilibrium rent per unit is approximately of the same magnitude as the annual service price of the generic input, which seems like a generous allowance for the proportion of returns from idiosyncrasy.

While the first two columns of Table 4 show the base-case and the C=100 case, the next two columns provide the equivalent results on the “oligopolistic expansion” setup. Interestingly, even in the case of oligopolistic output restraint, the ratio of the sustainable portion (including market power effects) to total wealth creation is not substantially different from the competitive case. Again, sustainable profit is not the most important contributor to wealth creation, at least by most metrics.

A more direct comparison of how different factors affect the NPV of total profits at the level of the industry comes from comparing what happens when we change the factors driving sustainable profits (i.e., when we change the conditions surrounding customization cost) and compare them to other changes in terms of competitive dynamics. One particular robustness check is the change of the fraction of net profits available for capability investments, θ. By focusing on the top row of Table 4 we can see that NPV increases by 173%, as C changes from 10 to 100. Columns 5 and 6 offer yet another comparison, as they provide the NPV figures for the C=10 case and with net profit reinvestment θ to 0.2 (Column 5) or to 1 (Column 6) as opposed to 0.5 in the base-case (Column 1). Changing θ increases operating profit by 30% or reduces it by 24% respectively. These substantial changes can be compared to a mere 3% increase in profit NPV when we add the oligopolistic expansion constraints.

This suggests that a research focus on the factors that increase or decrease sustainability may be substantially less valuable than a focus on factors that affect the dynamic adjustment process. It also suggests, again, that a high-confidence assessment of the relative significance of transient and sustained profits will require better calibration to the broad historical record.

Insert Tables 4 and 5, and Figures 5 and 6about here

Since the comparisons in Table 4 are based on summary values at the industry level, they show aggregative effects and hide the drama of the evolutionary struggle between more and less effective firms. Some of that drama is suggested in Figures 5 and 6, which present illustrative firm-level detail for the maximum heterogeneity, C = 10 case. Panels 5 a-b display how the NPV at time zero is accumulated over time. The graph asks, “what would the NPV of a given firm (Y axis) look like if we stopped the world (or the discounting of profits) at period X”? Thus, a positive NPV contribution at a particular period displays as a rise in the curve. In these calculations, we charge off the firms’ initial endowments of 3 resource units, at a customization cost of 10 per unit, as investments occurring at time 1. (Other antecedent investments in learning, which yield the particular capability through which a firm competes, are implicit in our story, but not accounted for here.) Out of the total of 11 firms in our scenario, we display the top 6 firms in terms of efficiency in 5a and the bottom 6 in 5b, with firm “A6” appearing in both – and with a vertical scale difference enhancing visibility for the less efficient firms. One immediate conclusion is that our “fade away” rule for the downward adjustment of R produces very little sacrifice of NPV; the appearance is almost as if firms were promptly making discrete choices to discontinue operations when profitability turned negative. That this is not quite the case is seen in 5b, where small declines in NPV are visible, as firms briefly operate at a loss before they shrink and shut operations. Also in 5b, we see that the least efficient firm never makes it into positive territory for its shareholders, while the next-to-worst makes it there but quickly gives up the gains in the fade-away process. In Figure 6, we see that the impact of falling output price on the flow of operating profit is a dominant influence from the start; only a hint of the impact of firm growth is seen. The operating profit streams for the top two firms level out a positive value. This reflects their long term coexistence, discussed previously. Finally, the dip below zero reflects the brief period of unprofitable operation for firms just before they exit, as discussed.

Effects of Changing the Number of Firms

Consistent with our expectation that eleven firms should suffice to illustrate the competitive case, this change had very minor effects on the results, apart from very straightforward consequences such as the decrease in the Herfindahl index and an increase in the number of surviving firms (see Table 5). In both the competitive and "oligopolistic" settings, the principal qualitative results are substantially unchanged.

It should be noted, however, that the combination of our oligopoly extension and a firm count manipulation almost certainly *would* change the results significantly, especially if we reduced the firm count substantially instead of increasing it. In that case, the industry structure would bring the oligopoly markup into play in a significant way from the very start, and its impact would further increase over time. This observation has important implications for the connection between sustainable profit and wealth creation in an industry evolution context. In many historical examples of industry evolution, significant concentration took substantial time (decades) to develop – as in our simulations – after perhaps a brief early stage with just a few entrants on the scene. In these cases, the impact of concentration may be quite modest in present-value terms. In other cases, however, high concentration is present in the initial conditions -- often reflecting the fact that strong basic patents are held by the industry founder, who promptly becomes its dominant firm. The early profits of such a founder are likely to be important in the overall wealth picture.

DISCUSSION

Our model formulation centered on customization cost seems, on the one hand, to be highly consistent with the theory of sustainable advantage presented in the RBV literature, and on the other, to resolve some outstanding puzzles about the relationship of that theory to standard microeconomics. We show how profitability that is *sustainable in long-run equilibrium, even under competitive conditions* may derive from the possession of idiosyncratic rent-earning resources.

In our formulation, "idiosyncrasy" matters not just because it might have something to do with the feasibility of imitation by rivals, but because it has everything to do with the input market conditions that allow a firm to trap the rent stream from a resource that it does not necessarily own – a resource that, in its generic form, is priced in the market every period. In the interest of framing this conceptual issue in the sharpest possible terms, we have posited that the specialized version of the resource that is effective in an individual firm is literally unique in the sense that the customizing investments that create it produce no value in any other firm. This extreme formulation is only a guidepost, but a valuable one, for understanding the complex reality in which the degree of idiosyncrasy varies with the routines of individual firms.

While the heritage of Ricardian rent theory remains highly visible, our picture of the origins of rents departs from that tradition in ways that seem plainly congenial to RBV thinking. Far from being "the original and indestructible powers of the soil" (with access thereto being acquired simply by a transaction), the ultimate source of the value that resources contribute is something that the firm creates through a time-consuming process – and creates to an extent that is guided and constrained by market competition (see, again, Dierickx & Cool, 1989). Our formulation is, similarly, consistent with Barney's (1986) critique of the idea that superior returns can be captured simply by buying firms that possess superior resources; our image of how firms capture value from "attached" resources certainly does not imply that outright purchase of access to large packages of semi-attached resources is advantageous. Yet we have, here again, suppressed important aspects of RBV thinking in the interests of simplicity. The only resource "leveraging" that is reflected in the model is the simple kind featured in basic evolutionary models – the profit-motivated effort to do more of the (profitable) same, which is accomplished by replicating routines. We have consigned to the prehistory of our process the creation of the idiosyncratic routines themselves, which are the typically the fruit of a complex interweaving of creativity, learning and investment over extended periods (Dierickx & Cool, 1989; Montgomery, 1995; Winter 1995; Levinthal 1997).

As our introductory discussion emphasizes, the interpretation of any "rent-as-profit" story must address the key question, "can this return legitimately be called 'profit'?" While there can be no definitive resolution of this question – because the legitimacy criterion is ultimately a matter of terminological taste – we argue that the persistent returns in our model can reasonably be termed profits. As a first observation, of considerable practical significance, we note that such returns reflect continuing benefits to which no continuing outlay corresponds, and, for this and other reasons, it is unquestionably the case that standard accounting will show such a return as a part of net income. Similarly, there is little question but that the types of costs involved in customizing a generic input are typically expensed rather than capitalized; the books of real firms show no asset corresponding to most capability investments. Thus, to the extent that propositions about sustained profitability have a connection to empirical evidence that is based on accounting measures, our candidate for a profit concept serves the purpose. Second, we note that such profits are indeed the sort of net return that firms *pursue*, i.e., strive to increase through effort, investment and the exercise of strategic judgment (though much of that process is suppressed in the model, as we have conceded). Third, we can respond to the possible objection that our profits "merely" reflect the normal return on investment. Actually, "normal return" plays a narrow but critical role in our model; it controls the *margin* of the process of capability creation for the most capable firm. On infra-marginal units in the resource creation process, the most capable firm typically makes above-normal returns. Other firms may benefit from their capability investments as well, even though they fail in the long run to yield rents corresponding to normal returns. Here and generally, the industry evolution framework serves to underscore the point that economic competition is typically analogous to athletic competition in a track event, not a field event. In the broad jump or the pole vault, your medal chances have little to do with *when* you start your run at the pit or the bar – but that is not true in the 100 meters, where timing is of the essence. "Time compression diseconomies" strictly limit the ability to recover from a tardy start!

In the existing strategy literature, the "uncertain imitability" model of Lippman & Rumelt (1982) offers an analysis of sustained profitability with key similarities to our own. Their model, like ours, examines competition among heterogeneous firms. The *ex ante* perfect symmetry among potential firms is broken in their case by a chance mechanism that assigns different, and permanent, efficiency levels to different firms; in our case this mechanism is implicit in the learning contingencies that give rise to the differing routines and efficiency levels at the start of our dynamic process. There are costs that must be incurred in order to establish a potential firm as an actual firm; in their case this cost is the one-time purchase of the information underlying efficiency: information that can be assessed by the purchasing entrepreneur only in probabilistic terms. Since these are not continuing production costs, yet are voluntarily incurred only in anticipation of returns, the equilibrium picture is one of firms earning differential rents. Further entry is blocked by the combination of the investment requirement and the uncertain outcome of an effort to imitate existing success. In our model, "customization cost" plays a similar role, but it is an obstacle to growth rather than to entry.

Our analysis also differs from Lippman & Rumelt in a number of respects, which mostly have to do with its grounding in evolutionary economics. First, like most evolutionary models, ours respects the additivity axiom of production theory. By contrast, the production functions assumed by Lippman &Rumelt are not merely inimitable, but also non-replicable by the firm itself. Returns to scale are diminishing, and this is the basic explanation for the fact that firms of different efficiencies can coexist in the equilibrium they describe. The question of why the scale of application of information should be limited *even for its initial possessor* is left dangling.[20] Second, our model shares the common evolutionary premise that mistaken decisions not only take place, but also have lasting consequences for the evolution of the system – in fact, these negative consequences can be viewed as the obverse side of the benefits that behavioral variety brings to the system. Lippman and Rumelt, by contrast, adhere closely to the full-rationality paradigm of mainstream economics, and manage with great analytical skill to describe an equilibrating process in which all of the worst mistakes (i.e., the investments with no returns) perish promptly and without consequence for other actors. Third, and closely related to the foregoing, we view the competitive process as occurring in real time and compute the implications of behavioral rules embedded in time. Behavior is represented as driven by information that actors might plausibly know. (Of course, we do not assert that our assumed rules are unique or even prominent in the class of rules with the requisite plausibility.) Finally, we claim that the notion of one-time costs of capacity expansion, as in the creation of a new establishment, has the virtue of greater realism than the idea of an entrepreneur's one-time purchase of a "recipe," a purely informational resource.

Our perspective also has some commonalities with Sutton (1991), who shows that "endogenous sunk costs" (i.e. the need to commit funds irreversibly to activities such as advertising, or R&D that can attract consumers) can account for patterns of industry structure, and for sustained profitability. We too consider the strategic ramifications of costs committed by the firm, though our interests lie with the supply side and with a behaviorally plausible conceptualization of production, rather than issues of demand. However, our concept of "customization cost" relates to sunk costs that are *not independent of scale*, which is the crucial consideration that underlies our highlighted possibility of persistent heterogeneity in a model involving price-taking firms. Also, unlike Sutton, we do not assume that firms can have access to any appropriate technologies, nor do we focus on equilibrium conditions. We emphasize the adjustment process, arguing that wealth creation in the economy is largely the result of

[20] We do not mean to imply that this question has no plausible answer, only that Lippman and Rumelt do not offer one. Indeed, our own formulation of "customization cost" illustrates one path to an answer: it is possible to treat information as non-rivalrous without necessarily assuming that replication is costless. See the discussion of "non-standard examples of information economics" in Winter and Szulanski (2002).

temporary profit streams, reflecting successive waves of industry creation, creative destruction, and re-generation.

More generally, our paper shows the importance of complementing the analysis of flow profitability at one point in time (and of long-run equilibrium) with the analysis of the net present value of cash flow (and of the adjustment path), since what we should be concerned with is the total wealth that is created in a sector. That is what mainly matters to the appraisal of outcomes, from the viewpoint either of the focal private actors or of society at large.[21] Short-run and equilibrium perspectives, by contrast, serve primarily the interests of diverse analysts and academic theoreticians, whose life is made simpler thereby.

Limitations

Our exploration of the dynamic consequences of efficiency differences has not, of course, produced a "rule for riches" – or even a promising set of hints. We have not accounted for the mechanisms that separated the best firms from the worse in our arrays. The ultimate sources of superior efficiency – whether in creativity, persistent effort, or luck – remain beyond the scope of our analysis. The very diversity of these potential sources suggests that the prospects for future insights through the application of the sorts of techniques used in this paper should be assessed with restrained optimism. We have employed strong simplifications to clarify the economic logic of "semi-permanent attachment" and "sustained profitability." These simplifications unfortunately tend to obscure other highly relevant aspects – the implications of entry and exit, the time-consuming nature of capability building, the financial implications of owned assets, the consequences of secondary markets for partially customized assets (which may exist, even though thin). Our experiments involve a stylized start to the industry evolution process, and invoke a rough judgment about how far from equilibrium such a start might realistically be. We did not attempt to model a "natural start" in which the contending firms arrive on the scene from diverse origins, inventing their routines on the fly; nor did we explore how the "distance from equilibrium" issue is entangled with the complexities of ongoing technological change and the demand changes associated with aggregate economic growth. Most importantly, our effort to clarify the subtle issues involving pricing of resources that are "semi-permanently attached" to the firm led us to exclude the more substantial and straightforward "attachment" of asset ownership. In so doing, we also excluded from the present-value analysis the wealth changes

[21] Some of the private wealth doesn't correspond to social wealth – particularly, the fruits of market power illustrated in our "oligopolistic expansion" calculations.

attributable to changes in the market prices of owned assets – which in many cases form an important part of the picture. Whereas the positioning school has tended to discuss timely purchase of key resources in terms of the creation of entry barriers, we would argue that this view has, at a minimum, to be complemented by an acknowledgment of the straightforward wealth consequences of changing market valuations – changes that the combination of input scarcity and industry growth will very commonly produce. (Jacobides, Knudsen & Augier, 2006; Jacobides & Winter, 2007; Winter, 1995)

Finally, while we have discussed "wealth creation", we have not examined that concept itself in depth. We did not address its relationship to the "destruction" aspect of "creative destruction" – i.e., the destruction of private wealth that often accompanies wealth creation of the kind we discuss. This is important not only because the dynamics of new and old are often closely linked, but also because the broader perspective is essential to understanding the implications for "the wealth of nations" as opposed to private wealth. Also, we did not make the link to shareholder returns and wealth at the investor level. Since the stock market registers opinions about the future as well as current results, it is reasonable to assume that wealth creation at the shareholder level is, in the aggregate, shifted toward the early stages by an even greater amount than our calculations suggest. However, the uncertainty that attends these situations implies that the net wealth creation at this level is the net result of a lot of ups and downs in the prospects of individual firms. Individual stock prices are moved, in both directions, as the news of the evolutionary struggle flows in. If, as we recommend, the assessment of wealth creation moves up on the agenda of the strategy field, there will be a need for strong empirical methods appropriate for such long-term appraisals and for interpreting shorter-term movements in strategic terms. For such work, we urge attention to the very compelling model provided some time ago by George Baker's research on Beatrice (Baker, 1992).

Conclusions and Future Agenda

Summing up, this paper uses formal modeling (both the analytical, closed-form variety, and computations) to advance our theoretical understanding of some key strategy issues in five ways. First, we elucidate the nature of the "semi-permanent attachment" of resources, providing a simple model consistent with the overview of appropriation issues provided by Coff (1999). Second, we show that heterogeneous firms can co-exist in a competitive equilibrium, and establish that the degree of heterogeneity is directly linked to the firms' growth calculus. Third, we show that sustainable profits in an industry *can well emerge without any "monopolistic"* *imperfections or information asymmetries on either the product or the resource market.* This strengthens the case for the "economizing" viewpoint by directing attention to the dynamic logic of competition and industry evolution. Fourth, we provide an indicative list of dynamic factors that shape industry evolution and profit distribution (capability heterogeneity; resource customization cost; ease of expansion), and show how these elements interact to drive both short-run and sustainable profits. And fifth, we show that sustainable profits often represent only a small part of the total process of wealth creation, and that changes in the factors that affect the adjustment processes of profits over time may be much more important than changes in factors that determine the ultimate sustainability of profits.

In future work we propose to look further into a number of strategic issues that cannot be traced through a conventional model. Several of these will be pursued with more attention to analysis at the individual firm level, where the consequences of heterogeneity are played out. High on that list is the question of how imitation affects the picture. Our preliminary explorations tend to confirm the basic insights of the RBV regarding the implications for long-run flow profitability, but also suggest the importance of confronting other aspects, such as the impact of different imitation rates on NPV. Our computational model already includes the capability to address imitation, likewise the impact of changing elasticities of demand and supply. We are also equipped to consider the implications of correlation between the (firm-specific) customization cost C_j and the firm efficiency parameter a_j. We can also explore the role of "learning curve" effects. We will probe more deeply into the problem of the division of the rent between resource owners and firms.

Significant as this future agenda is, some important issues have already been clarified. This paper provides a bridge between the tenets of the RBV and microeconomic theory, as well as evolutionary theorizing. It puts forward a novel explanation for why differentially capable firms might co-exist in competitive equilibrium, focusing on the role and nature of capability development. It considers the nature of profits, as distinct from returns to resources. Finally, we propose a shift of emphasis from sustained flow profitability to the net present value of cash flows, and from comparative static analysis to the dynamic adjustment process and its determinants. The need for such a shift in strategic analysis is the paper's most important message.

APPENDIX:

ANTICIPATED CHANGE IN VARIABLE PROFIT PER UNIT OF THE RESOURCE

An important fact about the trajectories displayed in early stages of industry evolution is that initially high-profit prospects gradually give way to the more "normal" prospects of approximate equilibrium. Both *a priori* logic and experimental calculations suggest that plausible investment calculations for profit-seeking, boundedly rational and imperfectly informed actors cannot be derived by imputing to such actors the assumption that the profit circumstances of the current period will persist indefinitely. In search of a simple way to represent the way that prevailing trends affect actor calculations, we settled on a simple "adaptive expectations" scheme focused on variable profit per unit of the resource. In effect, this scheme extrapolates recent experience of percentages changes in the quantity defined in equation (9) above, namely:

$$V(t)=\left[\Pi^{*}(t)+1\cdot K(t)\right]/R^{*}(t)=P(t)\cdot a\cdot\left[R^{*}(t)/K(t)\right]^{b-1}-W(t)$$

It is more convenient to do the analysis in the changes of the natural log of *V(t)*, rather than the period-to-period percentage changes. (The convenience has to do with the fact that representing the relevant compounding across time would require use of geometric means, if we were to deal in the percentage changes directly.)

Thus, in the spirit of adaptive expectations, we define $Lvdel(t) = \log\left(V(t)/V(t-1)\right)$, the most recent change in the log of variable profit per *R*, and

$Lvdav(t) = (1-d)\cdot Lvdav(t-1) + d\cdot Lvdel(t)$ is the moving average process for that change, yielding geometrically declining weights for the variable profit series with common ratio *(1-d)*, $0 \le d \le 1$. When *d = 1*, this formulation extrapolates the most recent change – *not* the most recent value of $\log\left(V(t)\right)$. Then the period *t* estimate of the per-period relative rate of variable profit change is $z = \exp\left(Lpdav(t)\right) - 1$. In our calculations, we assume that firms allow for a single period of change at that rate, then extrapolate constancy from there on. In the industry evolution context, this one-period change is expected to be negative, but becoming less so as time goes on. Then $(1+z)\cdot V(t)$ would be the projected variable profit per *R* for the following period, and $(1+z)\cdot V(t)/i$ is the present value if that variable profit stream is assumed to continue indefinitely.

REFERENCES

Baker, G. 1992. Beatrice: A study in the creation and destruction of value. *Journal of Finance*, 47(3): 1081-1119.

Barnett, W. 1995. Telephone companies. In G. Carroll, & M. Hannan (Eds.), *Organizations in Industry: Strategy, Structure, Selection*: 278-289. New York: Oxford University Press.

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

Barney, J. 1991. Firm resources and sustained competitive advantage. *Journal of Management*, 17: 199-220.

Becker, G.S. 1993 [1964]. *Human Capital: A Theoretical and Empirical Analysis, With Special Reference to Education,* 3rd ed. Chicago: University of Chicago Press.

Bloom, N., & Reenen, J. V. 2007. Measuring and explaining management practices across firms and countries. *Quarterly Journal of Economics*, 122: 1351-1408.

Brandenburger, A., & Stuart, H.W. 1996. Value-based business strategy. *Journal of Economics and Management Strategy*, 5: 5-24.

Chew, W.B., Bresnahan, T., & Clark, K.B. 1990. *Measurement, Coordination, and Learning in a Multiplant Network.* Harvard Business School Press, Boston, MA.

Christensen, C. M., Suarez, F. F., & Utterback, J. M. 1998. Strategies for survival in fast-changing industries. *Management Science*, 44(12 Part 2): S207-S220.

Coff, R.W. 1997. Human assets and management dilemmas: Coping with hazards on the road to resource-based theory. *Academy of Management Review*, 22: 374-402.

Coff, R.W. 1999. When competitive advantage doesn't lead to performance: The resource-based view and stakeholder bargaining power. *Organization Science*, 10: 119-133.

Coff, R.W. 2010. The coevolution of rent appropriation and capability development. *Strategic Management Journal*, 31: 711-733.

Demsetz, H. 1973. Industry structure, market rivalry, and public policy. *Journal of Law and Economics*, 16(1): 1-9.

Dierickx, I., & Cool, K. 1989. Asset stock accumulation and the sustainability of competitive advantage. *Management Science*, 35(12): 1N4-1511.

Dixit, A. K., & Nalebuff, B. J. 1991. *Thinking Strategically: The Competitive Edge in Business, Politics, and Everyday Life*. New York: Norton.

Dosi, G., Marsili, O., Orsenigo, L. & Salvatore, R.. 1995. Learning, market selection and the evolution of industrial structures. *Small Business Economics* 7(6): 411-436.

Foster, R. N., & Kaplan, S. 2001. *Creative Destruction: Why Companies That Are Built to Last Underperform the Market -- and How to Successfully Transform Them*. New York: Currency/Doubleday.

Gort, M., & Klepper, S. 1982. Time paths in the diffusion of product innovations. *Economic Journal*, 92: 630-653.

Helfat, C. E., & Lieberman, M. B. 2002. The birth of capabilities: market entry and the importance of pre-history. *Industrial and Corporate Change*, 11: 725-760.

Jacobides, M. G. 2000. Capabilities, Transaction Costs, Information Technology and Profitability in the Un-Bundling Mortgage Banking Value Chain: A Study of the Drivers and Implications of Changing Vertical Scope. Unpublished Doctoral Dissertation, University of Pennsylvania, Philadelphia, PA.

Jacobides, M.G. 2008. How capability differences, transaction costs, and learning curves interact to shape vertical scope. *Organization Science* 19(2): 306–326.

Jacobides, M. G., Knudsen, T. & Augier, M. 2006. Benefiting from innovation: Value creation, value appropriation and the role of industry architectures. *Research Policy*, 35: 1200-21.

Jacobides, M. G., & Winter, S. G. 2007. Entrepreneurship and firm boundaries: The theory of *A* firm. *Journal of Management Studies*, Forthcoming.

Jovanovic, B. 1982. Selection and the evolution of industry. *Econometrica*, N: 649-670.

Kato, M. 2010. The role of investment efficiency in the industry life cycle. *Industrial and Corporate Change.* 19 (1): 273-294

Klepper, S. 1997. Industry life cycles. *Industrial and Corporate Change*, 6(1): 119-143.

Klepper, S., & Graddy, E. 1990. The evolution of new industries and the determinants of market structure. *RAND Journal of Economics*, 21(1): 27-44.

Klepper, S., & Simons, K. L. 2000. Dominance by birthright: entry of prior radio producers and competitive ramifications in the U.S. television receiver industry. *Strategic Management Journal*, 21: 997-1016.

Koopmans, T. C. 1957. *Three Essays on the State of Economic Science*. New York: McGraw.

Levinthal, D. 1997. Adaptation on rugged landscapes. *Management Science*, 43: 934-9N.

Lippman, S., & Rumelt, R. P. 1982. Uncertain imitability: An analysis of interfirm differences in efficiency under competition. *Bell Journal of Economics*, 13: 418-438.

Lippman, S. A., & Rumelt, R. P. 2003a. The payments perspective: Micro-foundations of resource analysis. *Strategic Management Journal*, 24(October special issue): 903-927

Lippman, S. A., & Rumelt, R. P. 2003b. A bargaining perspective on resource advantage. *Strategic Management Journal*, 24(11): 1069-1086.

Makadok, R. 2001. Towards a synthesis of resource-based and dynamic capability views of rent creation. *Strategic Management Journal*. 22(5): 387-402

Makadok, R. 2010. The interaction effect of rivalry restraint and competitive advantage on profits: Why the whole is less than the sum of the parts. Management Science, 56: 1090-1102.

Malerba, F., Nelson, R., Orsenigo, L., & Winter, S. 1999. History-friendly models of industry evolution: The computer industry. *Industrial and Corporate Change*, 8: 1-36.

Mass-Collel, A., Green, J. R., & Whinston, M. D. 1995. *Microeconomic Theory*. New York: Oxford University Press.

McDonald, G., and Ryall., M.D. 2004 How do value creation and competition determine whether a firm appropriates value? *Management Science.* N: 1319-1333.

Montgomery, C. 1995. *Resource-Based and Evolutionary Theories of a Firm: Toward a Synthesis*. Boston, MA: Kluwer.

Nelson, R. R. 1994. Why do firms differ, and how does it matter? In R. P. Rumelt, D. E. Schendel, & D. J. Teece (Eds.), *Fundamental Issues in Strategy*: 247-270. Boston: Harvard Business School Press.

Nelson, R. R., & Winter, S. G. 1982. *An Evolutionary Theory of Economic Change*, Cambridge, MA: The Bellknap Press

Pacheco de Almeida, G., & Zemsky, P. 2007. The timing of resource development and sustainable competitive advantage. *Management Science* 53: 651-666.

Peteraf, M., & Barney, J. 2003. Unraveling the resource-based tangle. *Managerial and Decision Economics*, 24(4): 309-323.

Porter, M. E. 1980. *Competitive Strategy: Ttechniques for Analyzing Industries and Competitors*. New York, NY: Free Press.

Rivkin, J. 2001. Reproducing knowledge: Replication without imitation at moderate complexity. *Organization Science* 12: 274-293.

Rubin, P. 1973. The expansion of firms. *Journal of Political Economy*, 81: 936-949.

Rumelt, R. P., Schendel, D.E. & Teece, D.J,. 1994. *Fundamental Issues in Strategy*. Boston: Harvard Business School Press.

Schmalensee, R. 1985. Do markets differ much? *American Economic Review,* 75: 341-351

Spence, A. M. & Porter, M. E. 1982. The capacity expansion process in a growing oligopoly: The case of corn wet milling. *The Economics of Information and Uncertainty*. J. J. McCall. Chicago, University of Chicago

Sutton, J. 1991. *Sunk Costs and Market Structure*. Boston: MIT Press.

Teece, D. J., Pisano, G., & Shuen, A. 1997. Dynamic capabilities and strategic management. *Strategic Management Journal*, 18(7): N9-533.

Utterback, J. M., & Abernathy, W. J. 1975. A dynamic model of process and product innovation. *Omega*, 3: 639-656.

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

Williamson, O. E. 1999. Strategy research: Governance and competence perspectives. *Strategic Management Journal*, 20(12): 1087-1108.

Winter, S. G. 1964. Economic 'natural selection' and the theory of the firm. *Yale Economic Essays*, 4: 225-272.

Winter, S. G. 1971. Satisficing, selection and the innovating remnant. *Quarterly Journal of Economics*, 85(May).

Winter, S. G. 1995. Four Rs of profitability: Rents, resources, routines and replication. In C. Montgomery (Ed.), *Resource-Based and Evolutionary Theories of the Firm*: 147-158. Boston, MA: Kluwer.

Winter, S. G., & Szulanski, G. 2002. Replication of organizational routines: conceptualizing the exploitation of knowledge assets. *The Strategic Management of Intellectual Capital and Organizational Knowledge*. C. W. Choo and N. Bontis. New York, Oxford University Press: 207-221.

TABLE 1 : Model Glossary

Parameters and Single-Period Variables

$Q_j(t)$	Output	Endogenously derived
a_j	Production efficiency coefficient (capability)	Firm-specific, varied within & between scenarios
R_j	Level of resource useable by firm *j* in period *t*	State variable (short-run); choice variable (dynamically)
R_j^*	Level of resource used by firm *j* in period *t*	Choice variable
B	Cobb-Douglas production elasticity for *R*	Parameter
K_j	Supporting resource	Choice variable
$W(t)$	Price of resource *R*	Endogenously derived (constant in calculations here)
V	Price of resource *K*	Parameter
$\Pi_j(t)$	Operating profit	Endogenously derived
$\Pi_j^*(t)$	Optimized level of operating profit	Endogenously derived
D	Demand constant	Parameter
$P(t)$	Price of final good	Endogenously derived
ε_D	Elasticity of demand	Parameter
W_0	Supply constant for resource	Parameter
ε_S	Elasticity of supply	Parameter

Dynamic Update Parameters and Variables

C	Customization cost	Parameter, varied between "setups" (sets of scenarios)
$V_j(t)$	Variable profit per unit of resource	Calculated (by firms) derived value
I	Interest / required rate of return	Parameter
$H(t)$	Herfindahl index at t	Parameter derived from period t output levels
$z_j(t)$	Expectations for changes in per unit profit	Calculated (by firms) derived value - see Appendix
D	Weight on recent observations to derive $z_j(t)$	Parameter
$T_j(t)$	Investment test, net present value	Calculated (by firms) derived value (+ means invest)
$\Pi_j^n(t)$	Net cash flow	Derived value- realized profit minus investment costs
θ	Fraction of net (free) cash flow investable in *C*	Parameter
Λ	Fraction of unused resource lost to the firm	Parameter

TABLE 2: Experimental Design

***Within* each setting:**

- 11 Scenarios , with varying degrees of heterogeneity in capabilities
  - Scenario 0: all firms equal with a=1
  - Scenario 10: Firms most dissimilar (interval 0.1, max 1.5, min 0.5)
  - In all scenarios the initial mean is the same (set to 1 on annual basis)
- Each graph (Figures 1a to 4b) represents the evolution of the industry for *all 11 scenarios* over 250 periods, with periods 101-245 elided.

***Between* different settings:**

- Settings differ in the customization cost parameter.
  - Base-case setting: C=10 (Figures 2a--2e)
  - "High customization cost" setting, C=100 (Figures 3a--3e)
  - "Low customization cost" setting, C=1 (not displayed)
- All scenarios in all settings were re-run for the "oligopolistic expansion" case. Some of the results presented in Table 4, which also shows different values of θ
- Total net present value (NPV) was compared between with "high customization cost" and "low customization cost" settings and also for the "high expansion" case. Results are shown in Table 4

TABLE 3: Parametric Choices and Robustness Checks

Pa ameter		***Parameter Value or Range***	***Robustness Checks***	***Observations***
a	Production coefficient (capability)	$a = (1- x/20$ to $1+x/20)/4$, or $a = (1- x/20$ to $1+x/20)$ on annual basis, $x = 0, 1, 2, \ldots, 10$		Breadth of a dispersion produces expected result
b	Cobb-Douglas prod. elasticity for R	0.8	0.5 to 0.95	No significant changes to qualitative results
D	Demand constant	262.5, or 1050 on annual basis		N/A
ε_D	Elasticity of demand	1	0.5 to 2.5	Interesting results, unrelated to reported ones
C	Customization cost	1, 10, 100	1 to 200	Dynamic instabilities for C=1; small values onwards work well
i	Interest / required ROR	0.018245, or 0.075 on annual basis	0.0375 to 0.25	No significant changes to qualitative results; impact as expected
i'	Interest / required ROR for oligopoly	$\left(1+0.075\cdot(1+2H)\right)^{0.25}-1$ or $0.075\cdot(1+2H)$ on annual basis	$0.075\cdot(1+H)$	No significant changes to qualitative results; impact as expected
d	Weight on observations to derive $z_j(t)$	0.5	0 to 1	No significant changes to qualitative results; impact as expected
θ	Fraction of net cash flow invested in C	0.5	0.2 to 1	No major impact; higher θ leads to rapid movement toward equilibrium with erratic overshooting.
λ	Fraction of nused resource lost	0.5	0.5 to 1	No significant changes to qualitative results
N	No of firms (firm types) per run	11	6 to 21	No significant changes to qualitative res lts - see Figure 7
	Price of resource K	0.25, o 1 on annual basis	N/A	K chosen to have unit price, no robustness checks necessary
$W(t)$	Price of resource R	1.75, or 7 on annual basis	1 to 3	No significant changes to qualitative results.

TABLE 4: Profit Level and Sustainability, Under Competition and Oligopoly

(Scenario 5)	C=10	C=100	C=10, Oligopoly	C=100, Oligopoly	C=10, θ = 0.2 (slow expansion)	C=10, θ = 1 (rapid expansion)
Total Profit NPV, seen from period 1	2,855	7,796	2,947	7,973	4,099	2,178
Sustainable Profit NPV in period 250 % of Total NPV	0.08%	0.17%	0.11%	0.20%	0.05%	0.11%
Sustainable Profits (from period 1) over total NPV	28.9%	65.9%	42.1%	74.9%	18.8%	41.2%

TABLE 5: Comparing 11 and 21 Firms

(C = 10, maximum heterogeneity)

5a. Aggregate Measures at Period 250

	P	***Q***	***R***	***Qmax***	***Op Profit***	***Net Profit***	***K***	***Herfindahl***
11 firms	5.615	46.749	113.683	46.749	13.818	13.818	198.947	0.577
21 firms:	5.624	46.674	113.618	46.674	13.961	13.961	198.832	0.422
ratio 11/21	0.998	1.002	1.001	1.002	0.990	0.990	1.001	1.368

5b. Output of More Efficient Firms at Period 250

Firm #	1	2	3	4	5	**Total**
11 firms	33.716	13.033	0	0	0	46.749
shares	72.1%	27.9%	0	0	0	100.0%
21 firms	26.579	12.360	7.734	0.001	0	46.6738
shares	56.9%	26.5%	16.6%	0.002%	0%	100.0%

FIGURE 1: The Model's Schematic Overview

Static Model: Within Period *t* Production and Output / Price Determination

Every firm *j* sets supporting input $K_j(t)$ appropriate to its existing resource level $R_j(t)$, then maximizes profit choosing by $R_j^*(t)$ as the X that maximizes

$$\Pi(t) = P(t) \cdot a \cdot X^b \cdot K_j(t)^{1-b} - W(t) \cdot X - 1 \cdot K_j(t)$$

$$s.t.\ 0 \le X \le R_j(t)$$

Market for good clears

$$\sum_j a_j \cdot R_j^*(t)^b \cdot K_j(t)^{1-b} = D \cdot P^{-\varepsilon_D}$$

Market for resources clears

Aggregate *R* is a function of *W* (in examples here, *W* = constant)

$$S \cdot (W - W_0)^{\varepsilon_S} = \sum_j R_j^*(t)$$

Dynamic Update: Between periods (*t* to *t+1*), all firms decide their expansion (in terms of *R*)

Each firm *j* builds expectations on what extra profit it will get in the next period for extra resource unit R_j

$$V_j(t) = \left[\Pi_j^*(t) + 1 \cdot K_j(t)\right] / R_j^*(t) = P(t) \cdot a_j \cdot \left[R_j^*(t) / K_j(t)\right]^{b-1} - W(t)$$

Firm *j* then makes investment test T_j, based on costs & benefits

$$T_j(t) = \frac{(1 + z_j)}{i} \cdot V_j(t) - \frac{(1 - b) \cdot W(t)}{i \cdot b} - C$$

In the "oligopolistic" model *only,* required rate of returns

$$i'(t+1) = i \cdot (1 + 2H(t))$$

Firm *j* then considers its available cash flow for investing in capacity expansion

$$\Pi_j^n(t) = \Pi_j^*(t) - C \cdot max(0, R_j(t) - R_j(t-1))$$

If the firm expands according to $T_j(t)$, then expands by

$$R_j(t+1) = R_j(t) + \begin{cases} max\left(0,\ \theta \cdot \dfrac{\Pi_j^n(t)}{C}\right), & T_j(t) > 0 \\ 0, & T_j(t) \le 0, R_j^*(t) = R_j \end{cases}$$

If the firm didn't use resource last period, then adjusts down

$$R_j(t+1) = R_j(t) - \lambda \cdot (R_j(t) - R_j^*(t))$$

Each firm j, given its $R_j(t+1)$, chooses the corresponding amount of *K*,

$K_j(t+1) = R_j(t+1) \cdot W(t) \cdot (1-b) / b$. Then, the static model runs anew.

FIGURE 2

Features of Industry Evolution, Base Case (*C* = 10)

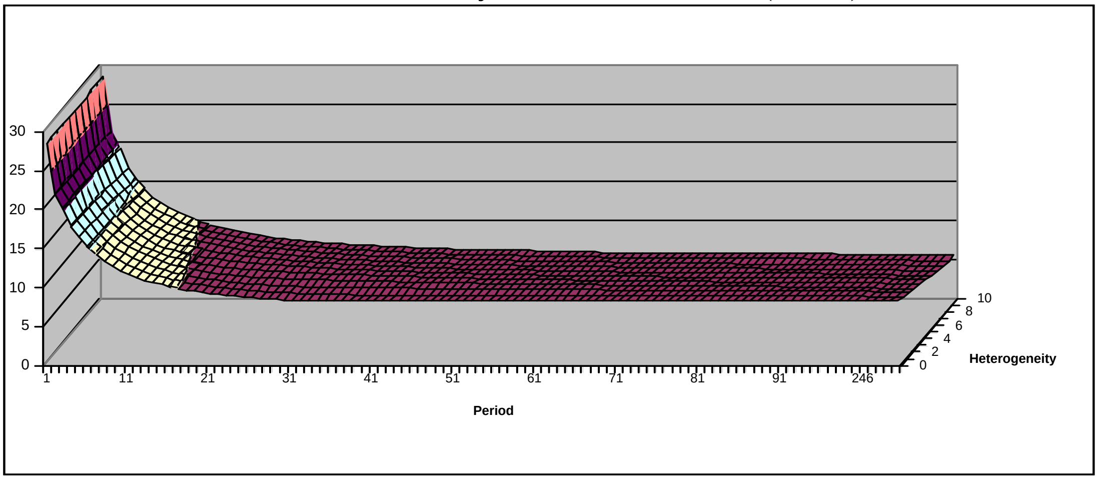


Figure 2a: Output price for C=10

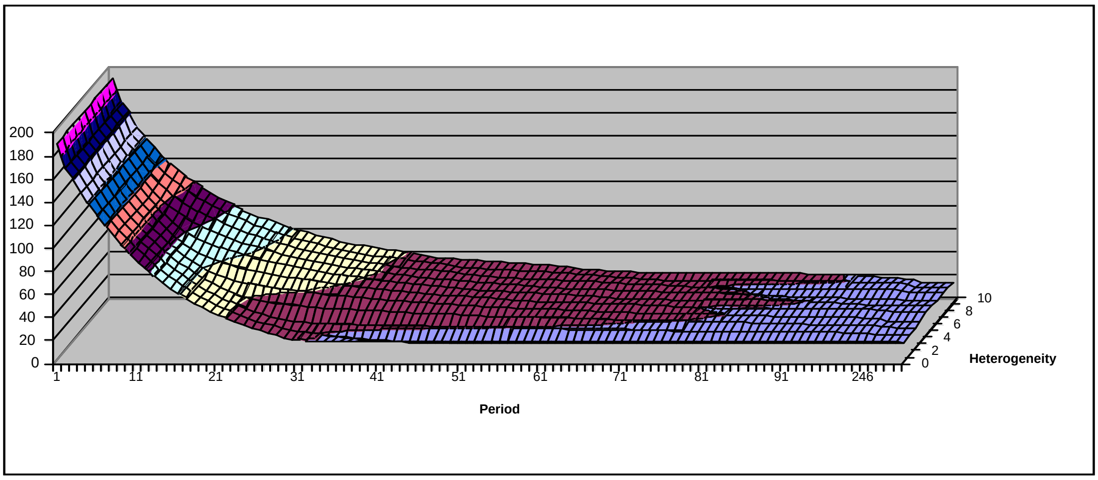


Figure 2b: Total operating profit for C=10

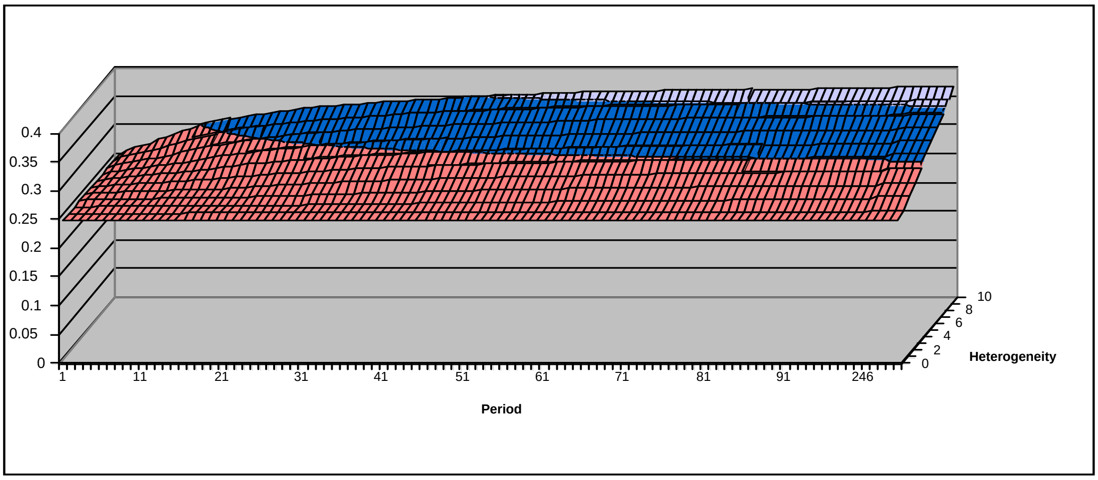


Figure 2c: Average capability in use for C=10

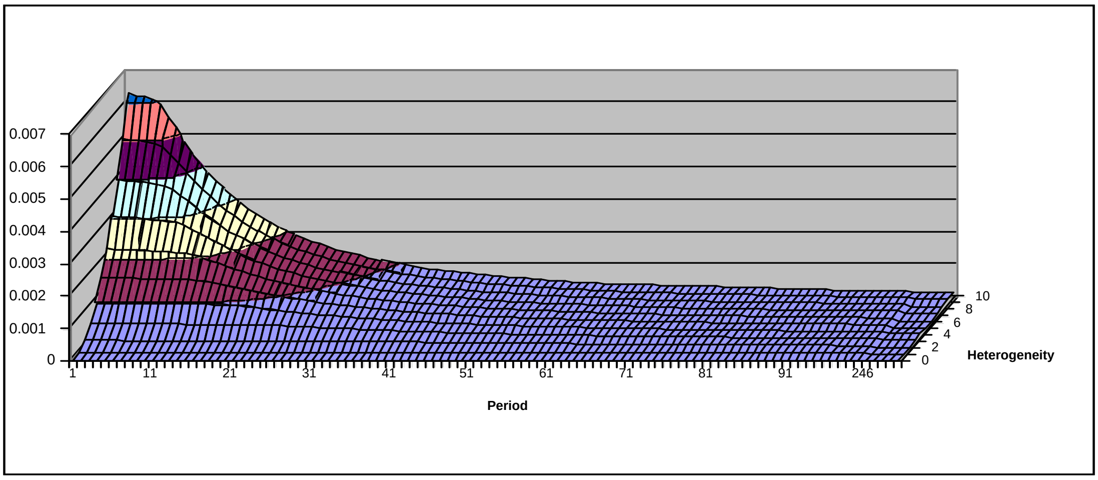


Figure 2d: Share-weighted variance of capability for C=10

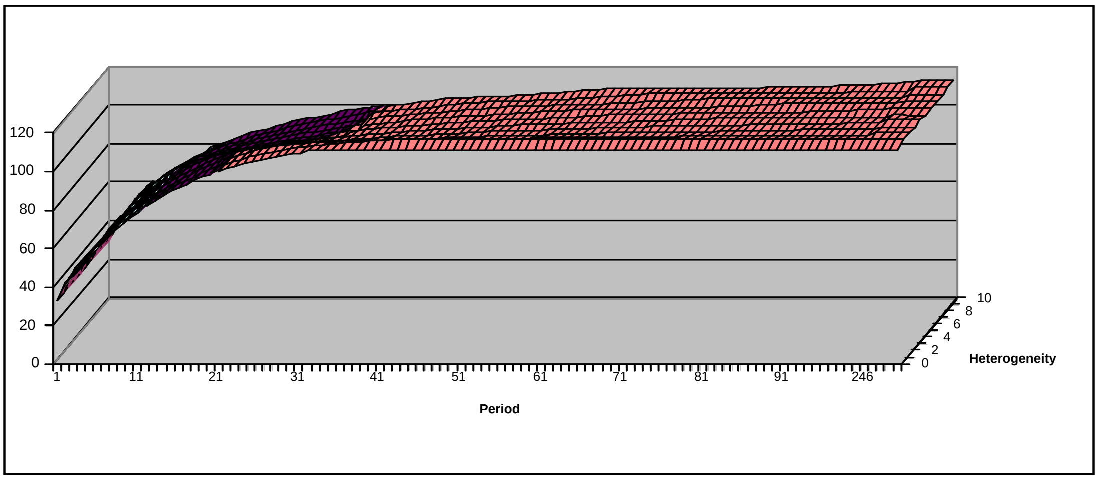


Figure 2e: Total resource use for C=10

FIGURE 3

Features of a High Customization Cost Industry (*C* = 100)

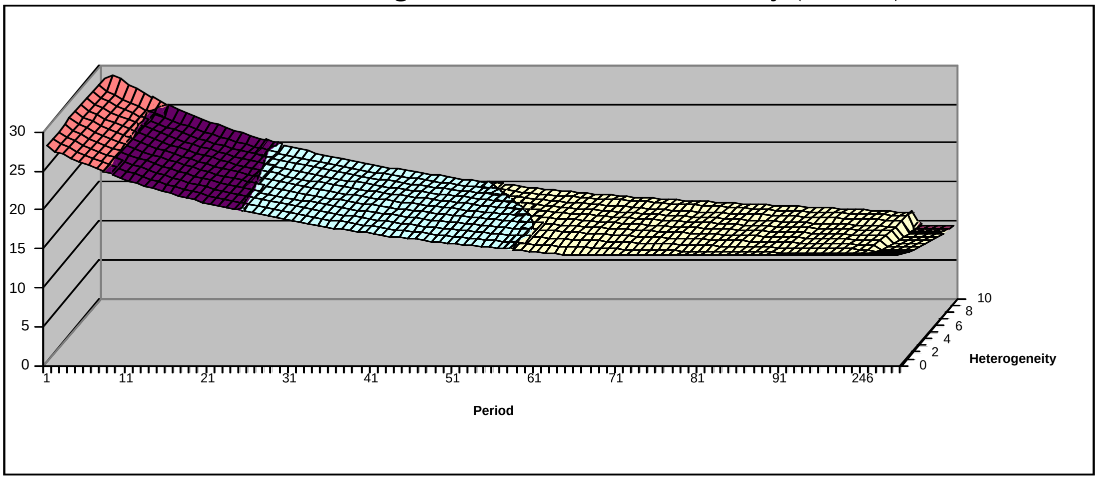


Figure 3a: Output price for C=100

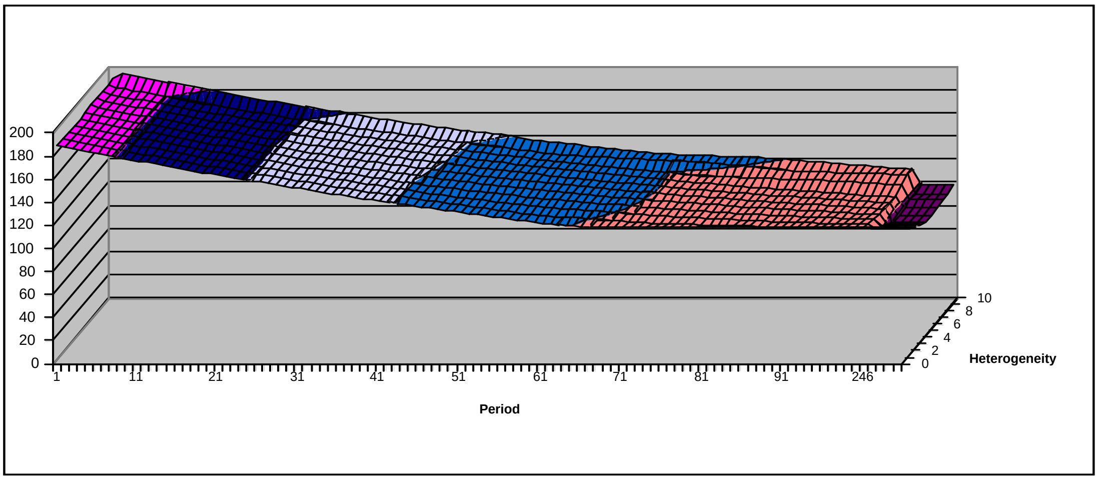


Figure 3b: Total operating profit for C=100

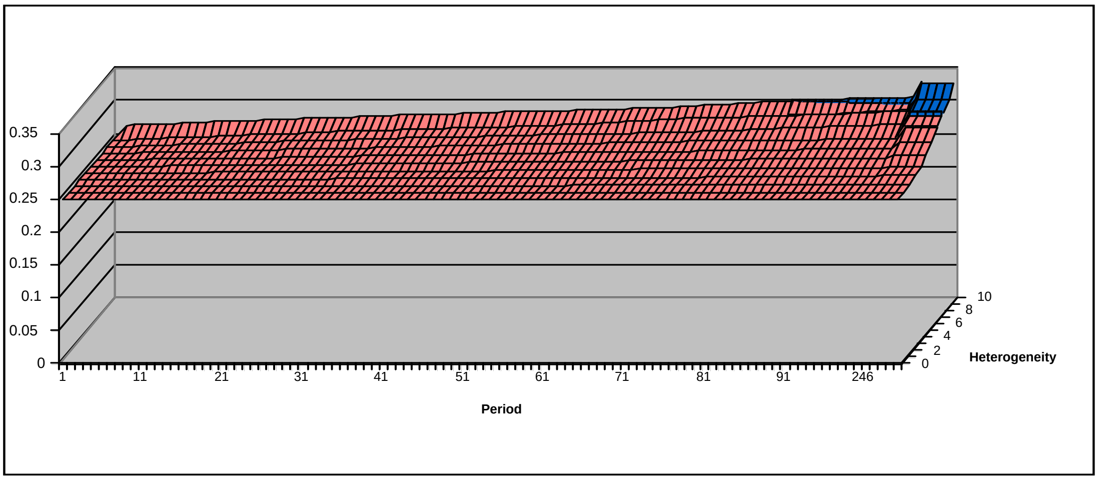


Figure3c: Average capability in use for C=100

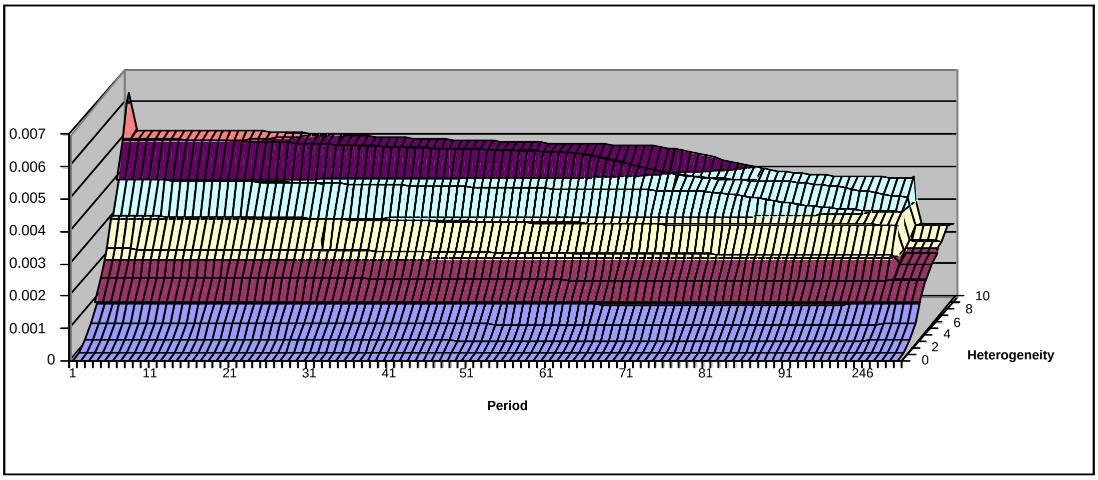


Figure 3d: Share-weighted variance of capability for C=100

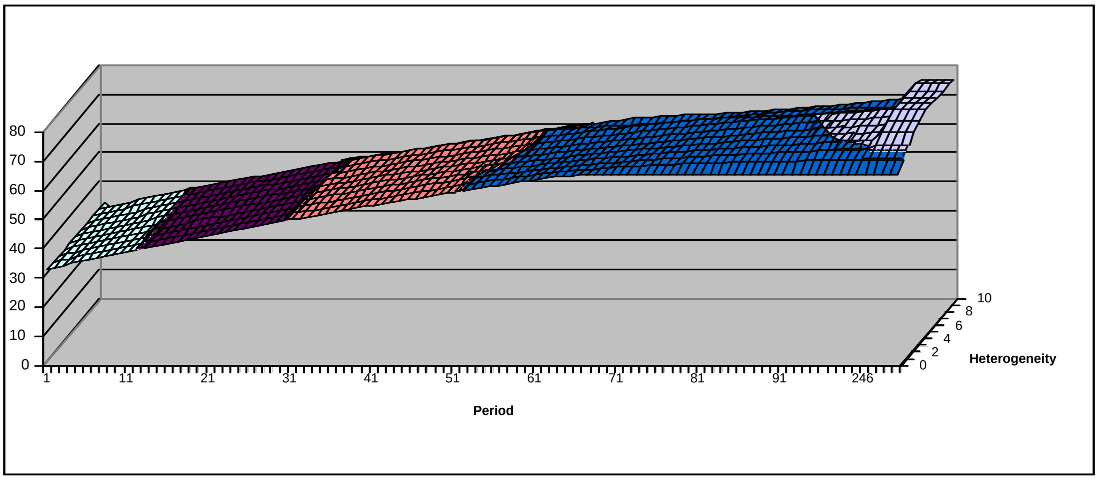


Figure 3e: Total resource use for C=100

FIGURE 4

Concentration in the Base Case and with High Customization Cost

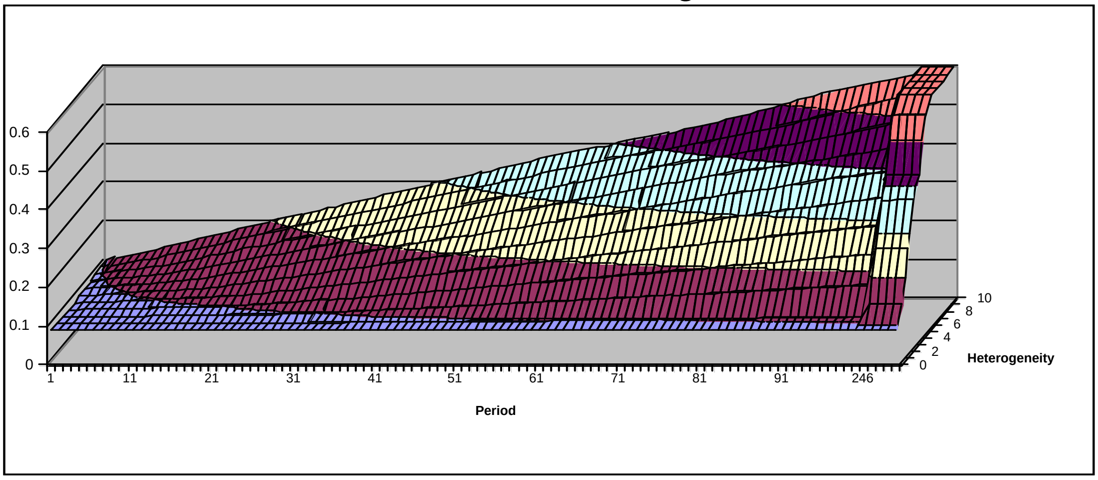


Figure 4a: Herfindahl Concentration Index for C=10

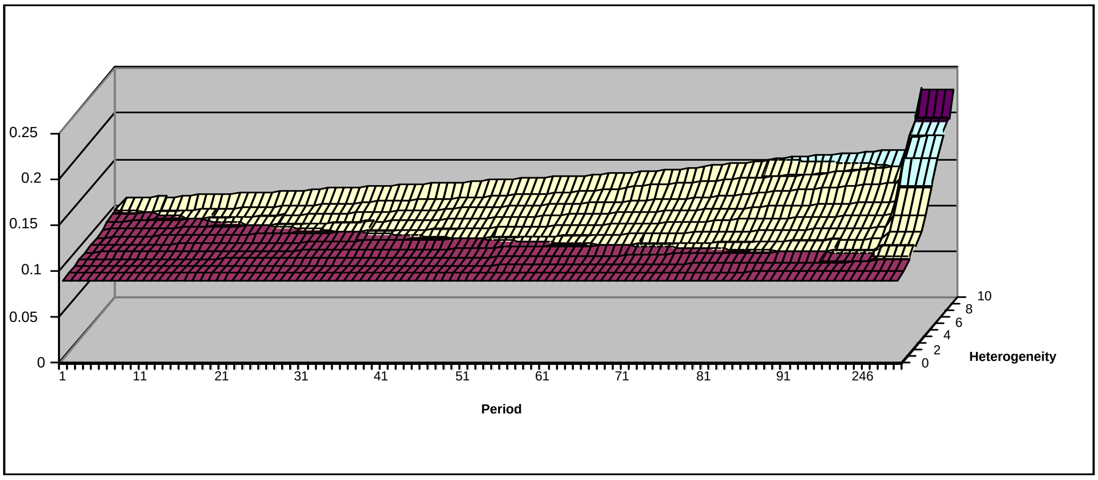
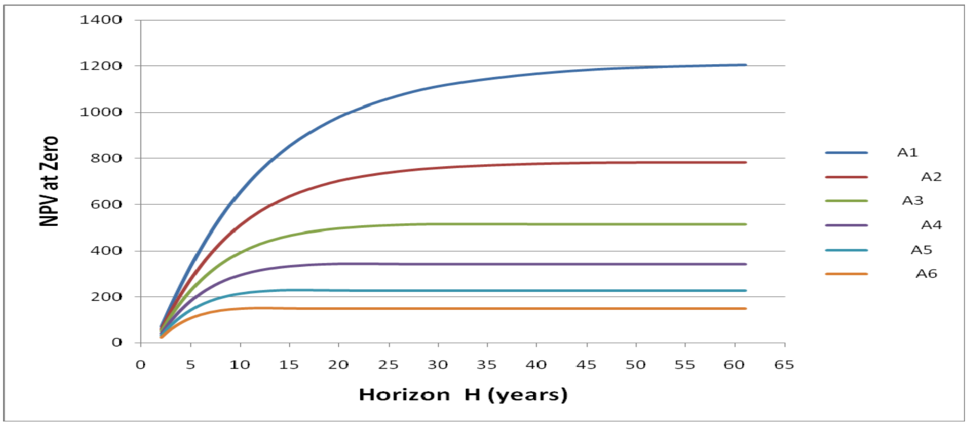
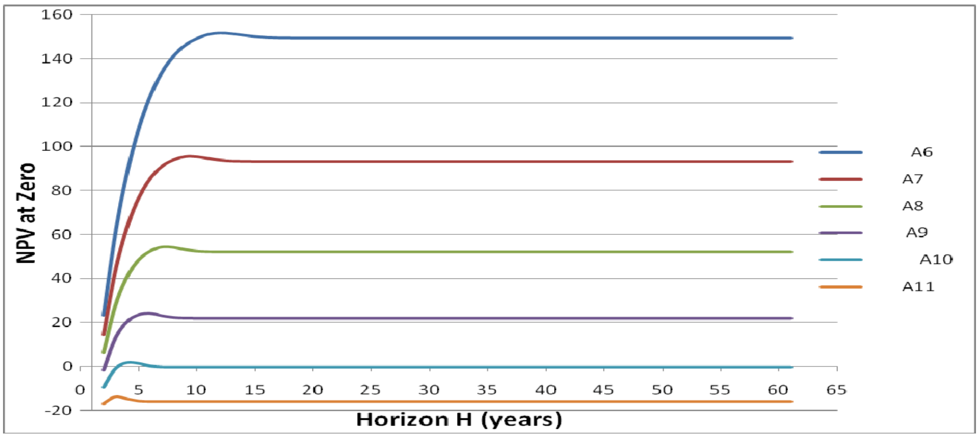


Figure 4b: Herfindahl Concentration Index for C=100

FIGURE 5
Creation of Time Zero NPV, by Firm, Maximum Heterogeneity



5a. Time Zero Present Value Realized by Horizon H (Six most efficient firms)



5b. Time Zero Present Value Realized by Horizon H (Six least efficient firms)

FIGURE 6: Operating Profit by Firm, Maximum Heterogeneity