A POSSIBLE MICRO-FOUNDATION FOR THE RBV AND ITS IMPLICATIONS

Birger Wernerfelt, MIT

July 17, 2018

1 bwerner@mit.edu. I am indebted to Jeffrey Reuer for encouraging me to write this. An anonymous referee and the Editors make very helpful comments on an earlier, incomplete, draft.
A Possible Micro-Foundation for the RBV and its Implications

Abstract

A central question in strategy asks in which markets and segments a given firm should compete. Any answer to this question is based on an implicit theory of the scope of the firm, and any theory of the firm implies an answer. The economics literature contains several slightly different theories of the firm, and we here focus on one which is based on advantages of specialization and what could be described as ‘increasing returns to scale’ in negotiations. We outline the theory and show how its recommendations are very similar to those of the RBV. There are, however, subtle differences and we highlight several of those.
I. INTRODUCTION

As will be well known to the readers of this journal, the RBV has, over the last thirty years, become very influential in strategy and also in management more generally. At the same time there has, maybe because of this influence, been a lot of debate about the foundations and theoretical status of the RBV (Barney, 2011; Foss, 2011; Mahoney and Pandian, 1992; Priem and Butler, 2001). The present essay seeks to contribute to this debate by proposing a micro-foundation for the RBV. Specifically, starting with a model based entirely on individual motivations and actions, we derive a set of normative principles that are very similar, though not identical, to those normally associated with the RBV. The ways in which the model recommendations differ from those of the RBV then allow us to suggest some minor refinements to the latter.

The argument deepens and sharpens earlier attempts by the author (Wernerfelt, 2013; 2016), and is based on the close relationship between theories of strategy and theories of the firm. In particular, theories of the firm tell us which activities a firm should undertake, and whether these should be inside or outside its boundaries. So by starting with a specific theory of the firm, we can derive its strategic implications and compare them to those of the RBV.

The case for an individual level foundation for the RBV has been made most forcefully by Felin and Foss (2005). They propose differences in individual histories (education or past

---

2 Barney (1991), Prahalad and Hamel (1990), and Wernerfelt (1984) have about 125,000 Google Scholar citations between them as of 2018.

3 Kogut and Zander (1992), Leiblein and Miller (2003), Nickerson and Bigelow (2008), Poppo and Zenger (1998), and Reuer, Arino, and Mellewigt (2006), are among many prior papers to make use of the connection between strategy and theories of the firm.

4 The general principle of “methodological individualism” has widespread support in the social sciences. There are, however, some dissenting voices. Hodgson (2012) and Winter (2013), for example, both question the feasibility of, and the need for, any micro-foundation for organizational phenomena.
employment) as one among several possible drivers of heterogeneity, and our theory falls in that category as well. However, it would make little difference if the heterogeneity was ascribed to other, less endogenous factors, such as genetics.\(^5\)

The lure of a micro-foundation is that it could help us to new and sharper advice and predictions about important aggregate (non-micro) issues: What is a firm, why does it exist, where should it compete, and when does it perform better? What is a resource (capability, competence), where does it come from, how would we recognize it, and what makes it valuable? Perhaps even more importantly, a micro-foundation would provide a scaffold with which we much more easily could develop new aggregate level constructs, knowing that they would be consistent with an existing body of work.

The key difficulty is that such a micro-foundation requires us to settle on a specific theory of the firm. The standard framework used in economics leaves no role for firms (a set of contracts can do the job), so to construct a theory of the firm, one needs to adopt at least one non-standard assumption. Several such assumptions, or sets of assumptions, have been put forward, but the economics profession has yet to agree on a candidate.\(^6\) We here choose one non-standard assumption, sketch the resulting Adaptation-Cost Theory of the Firm (ACT)\(^7\), and show how it portrays the firm in a way that is broadly consistent with the main tenets of the RBV.

The argument is presented as being about corporate strategy: Assuming that the firm already competes in a certain segment/market, when should it expand its scope and in which direction? However, this it only done to give the firm an initial identity and thus provide a starting point for

\(^{5}\) Coff and Kryscynski (2005) and Helfat and Peteraf (2015), among others, also root competitive advantages in individual differences, but focus on inherent heterogeneities.


\(^{7}\) Ch. 5 in Wernerfelt, 2016.
the analysis. Given that the firm has this resource, what should it do? The reader can think of the original business as small, dying, about-to-be-entered, or operating in a sub-optimal segment. If so, the argument is about business level strategy. Our intent is to provide a micro-foundation for both business and corporate strategy uses of the RBV.

To fix ideas and provide intuition, we start with an example from an almost universally well understood business.

II. A SIMPLE EXAMPLE

Before sketching the ACT, it is helpful to introduce the argument through an example.

Like all apartment buildings in the city, the brownstone owned by Joe needs a sequence of varied repairs performed over time. Some of these could be classified as plumbing, others as carpentry, electrical, heating, etc. Like all workers, Peter can in principle perform any of these services, but he is much more efficient if he can specialize. This specialization can take place along two dimensions: Peter can be a “superintendent” and do all of his work in the same building, or he can be a “tradesman” and perform a single category of repairs, say plumbing, in several buildings.

These jobs are governed differently, but follow a surprisingly regular pattern. Superintendents are invariably employees, while tradesmen sell their services in a more or less competitive market. It is not hard to understand why. Suppose first that Peter is a superintendent and performs a variety of odd services in Joe’s apartment building. Given the wide variety of possible repair services, it would be very inefficient if the two of them had to negotiate a new fee every time the need for a new service presented itself. Instead, they once and for all negotiate a
blanket contract according to which Peter will perform any simple repair service on demand in exchange for a fixed salary. By thus agreeing to “follow orders”, Peter becomes an employee in Joe’s firm. This is not necessary if the need to change orders is a very rare event, like in a renovation that is specced out in advance, but employment is used whenever adaptations are reasonably frequent.

While it is efficient for superintendents, the employment solution would not work very well for a tradesman. If Peter focuses on plumbing and works for many building owners, he does not want to take orders from all of them. It is much better, especially when there are several plumbers and many buildings, if Peter posts a price for each job. The competition between plumbers eliminates the need to negotiate over this price, and the dependence on the job allows the total fee to cover several different situations. Peter would not take orders in his role as a tradesman and would therefore be an independent contractor, rather than an employee in Joe’s firm.

In principle, the most efficient solution would allow Peter to simultaneously specialize on both a building and a single trade. This solution is, however, not feasible because Joe’s building does not need enough repairs of any one category. For example, a single apartment building can rarely use a full-time plumber. So the choice is between a superintendent who specializes in a single building and a tradesman who specializes in a single category of services.

Now suppose that Joe could buy a few other brownstones that are more or less identical to the one he has now. In this case, he might be able to employ Peter as a full-time plumber. As a plumber in Joe’s multi-building firm, Peter would not reap the entire benefit from specializing in a single building, but by only working in a small, relatively homogeneous set, he would be more efficient than if he responded to random incoming calls. So by diversifying into several “related” buildings, Joe can leverage the capacity (time) of his specialized personnel.
We will now present the theory more precisely. Readers who are unfamiliar with economics style models may skip directly to Section IV with little loss of continuity.

III. THE ADAPTATION-COST THEORY OF THE FIRM

a. Baseline Model

We will start out by looking at the case in which the scope of firms is fixed at a single business and labor is the only input. After presenting the ACT for that case, we generalize it by allowing firms to choose larger scopes and use other inputs.

Consider an economy in which each business needs one service in every period and each worker can perform any service for any business, but only one service for one firm per period. Workers can specialize or choose to be generalists in both the service and the business dimension, and this has natural implications for productivity. Specifically, a worker is most productive if he performs a service he is specialized in, medium productive if he is a service generalist, and least productive if he is specialized in one service but performs another. Since business specialization works in the same way, workers would ideally like to specialize in a single service and a single business. However, because each business needs a different service each period (and thus has to adapt), it is not possible to doubly specialize in that way. The best workers can do is to specialize in either a specific service or an individual business.

If workers are business generalists but specialize in different services and many businesses need each service, the service markets will have several participants on both sides and we assume that they are frictionless. On the other hand, if each worker specializes in a different business, the parties are effectively “locked” into each other. In such cases we assume that each business-
worker pair incurs bargaining costs and that these are increasing in the number of services covered by an agreement. That is, while the parties could negotiate a price per unit of quantity (or quality) for one service in each period, they may also agree on an average price that covers several, and even all, possible services. The key, and new, assumption is that these costs are sub-additive in the number of services covered. This means that there are economies of scale in bargaining, such that negotiating a unit price for any of \( N \) services is less than \( N \) times as costly as negotiating a unit price for one.

Depending on the relative costs of different kinds of production and bargaining, the most efficient way to organize the economy is one of three:

**Markets.** All workers are business generalists but each of them specializes in an individual service and goes from one business to another between periods.

**Sequential Contracting.** All workers are service generalists but each of them specializes in one business and in every period negotiates, with the entrepreneur, a contract covering the service needed by the business in that period.

**Employment/Firms.** All workers are service generalists but each of them specializes in one business and negotiates a blanket contract with the entrepreneur. The contract gives the entrepreneur the right to choose which service the worker is to perform in each period with no further negotiation. (So the worker “takes orders” and we use this to define him/her as an “employee” and the business as a “firm”.)

---

8 To keep things simple and rule out complete contingent claims contracting, we assume that only one price is agreed upon, even if it covers more than one service.
Note that the difference between the two latter possibilities depends on the frequency with which adaptation is needed. If the periods are very short, it would be inefficient to negotiate at the start of every one of them, but if periods are sufficiently long, there are benefits from postponing bargaining costs, rather than incurring them all upfront. Markets are more efficient that either of the last two when the benefits of service specialization outstrip those of business specialization.

b. Model with Endogenous Firm Scope

We now change the baseline model in too look at the scope of the firm. Instead of only differentiating between a business in which a worker specializes and all other businesses, we now assume that each business has a neighbor and that neighboring businesses are similar ("related") in two ways. First, they require similar human capital such that a worker who is specialized in one neighbor can work for the other with only a small decrease in productivity. Second, they tend to need the same services, though not necessarily concurrently. Aiming to keep things simple, we model the second property by assuming that each neighborhood has one service that is “common” in the sense that it is needed by one business in the neighborhood in every period.

In this case, again depending on relative productivities, a fourth class of organization can be efficient.

---

9 This is tested in Novak and Wernerfelt (2012).
Multi-business Employment. Entrepreneurs own two neighboring businesses and employ one worker who specializes in one of the businesses and in the common service. Another worker specializes in the other business, but is a service generalist. This worker is often an employee as well, but need not be.

This allows the second worker to be almost “doubly specialized”, performing the service in which he is specialized, either for the business in which he is specialized or for its neighbor. Multi-business employment can only be the most efficient mode of organization when both above-mentioned properties of neighborhoods hold. Markets will be more efficient if workers do not have an easier time changing between neighbors than between other businesses, and single business employment or sequential contracting will be better if neighboring businesses do not need some of the same services. It is also necessary that needs change somewhat frequently, since no employment solution can be dominant otherwise.

As is typical in economics, the model assumes that the players do not make any mistakes. In equilibrium, firms settle on their optimal scope before any trading takes place. Since we want to use the model to understand a situation in which a firm makes a strategic decision to change its scope, we have to specify an initial disequilibrium state in which the firm has a narrow scope but “should” have a wider scope. Suppose therefore that a single business firm has an employment contract with a worker who is specialized in a service that is common in its neighborhood (and therefore is needed in every period either by the business itself or by its neighbor). Since the business itself does not need this service in every period, it has excess capacity of the worker’s expertise (and because of the sub-additive bargaining costs, it is inefficient to employ him on a
part time basis). This excess capacity can be deployed in the neighboring business, and the parties can implement this solution by merging the neighbors and having the merged entity write an employment contract with the service specialized worker. So the increase in scope is motivated by a desire to utilize excess capacity.

c. **Inputs Other Than Labor**

There are two essential components in argument. First, workers are more efficient if they can specialize in individual services and businesses. Second, negotiations/contracting costs are subject to economies of scale (sub-additive) such that a single broad employment contract is cheaper to negotiate than a set of individual prices for each service in the job description (at least as long as the need to adapt does not arise very rarely).\(^{10}\) The latter premise implies that individual workers are indivisible in the sense that it is inefficient to employ one on a part time basis. Similarly, it is also inefficient for a specialist to work full time for an employer and but only use his expertise part time. In the latter situation, the employer controls excess capacity of the specialist’s time and related diversification is an efficient way to leverage this excess capacity.

We only very rarely, if ever, think of firms as leveraging the skills of individual workers. Both because those workers might demand higher wages and because most firms employ a large number of workers, thus making most individuals relatively unimportant. More often, we think of “skills” on a larger scale; high performing teams, large scale equipment, brand names, groups

---

\(^{10}\) This is experimentally tested in Maciejovsky and Wernerfelt (2011).
of loyal customers, etc. So it is important to know whether the argument extends to productive factors other than labor.

(a) Note first that the argument trivially extends to the case in which only teams of workers are effective, as is often the case in R&D, new product development, etc. Such a team could have excess capacity and it would be uneconomical if two different firms were to employ it on a part-time basis.

(b) Think next of physical assets with limited capacity, such as machines or entire plants. Such assets benefit from specialization and continuity in use; they are more productive if they can be dedicated to providing a small set of items for a narrow group of customers with relatively similar needs. Moreover, they can better serve these customers if they are owned by them, since this reduces the need to negotiate over adaptations. So ownership can play much the same role as employment but for hours on a machine rather than hours of labor. Similarly, short term rentals are analogous to markets for service specialists.\(^{11}\)

(c) Intangible assets with high but still bounded capacity, such as brand names, groups of loyal customers, corporate culture, or intellectual property, are by far the most important drivers of strategy. These assets are typically most productive in their original “core” use and then loose effectiveness as they are leveraged further afield. Furthermore, it is notoriously very hard to trade or rent fractions of them. So also this class of assets will typically be owned by the user. They often have some excess capacity and that could be leveraged in related diversification (and often are).\(^{12}\)

---

\(^{11}\) Large earth moving equipment is an example of a frequently rented physical factor.  
\(^{12}\) On the other hand, that are exceptions. It is, for example, not unheard of that luxury firms license (rent out) their brand names. This happens when the deal can be governed by a relatively robust contract, with few anticipated needs for adaptation.
IV. IMPLICATIONS FOR CORPORATE STRATEGY

The theory of the scope of the firm outlined above translates directly into a set of prescriptions about corporate strategy.

(1) If possible, firms should leverage specialized factors with excess capacity.

Compared to non-specialized factors, this would give the firm a productivity advantage in the markets where it is deployed.

(2) Firms should deploy the excess capacity in businesses (or segments) that are related to their current business. Relatedness is defined relative to the particular factor and consists of two requirements: That the services of the factor are valuable in the new segment, and that it can be productive in both old and new segments while still creating value in both.

For example, a brand name that is developed for a low end segment, may not transfer to a high end product. Conversely, if you deploy a high end brand in a low end segment, you may end up destroying value in your old segment. As a result, factors differ widely in the number of businesses that are related through them: Skilled management teams can be deployed very widely, while some pharmaceutical IP cannot.

(3) A factor only supports expansion as long as there exists businesses that, through the factor, are related to the firm’s existing scope, and do not require other factors that the firm cannot get at market prices.

While a skilled management team can be deployed widely, it cannot support entry into businesses that require a lot of proprietary technology. On the other hand, some pharmaceutical IP may allow the firm to enter some very sparsely populated industries.

(4) Excess capacity could be leveraged even if the factor is not unique, but unique factors can support more extensive leverage.
Even a non-unique factor could give you a cost advantage in the markets where you deploy it, as long as competing firms have to pay for it. However, fewer businesses will be related through a non-unique factor because it confers less value and thus can bear less customization.

(5) *Firms may have excess capacity of a factor if it is uneconomical to trade fractions of the factor or to rent the whole factor out on a short term basis.*

If you can trade any capacity you do not need yourself, you would never hold any excess, but simply trade it away. A factor that cannot be traded in fractions, but in its entirety, could still support expansion, though the firm has the additional option of selling it and leaving the leveraging to others.

(6) *Firms should expect decreasing rates of profit as they diversify more widely.*

The result holds by definition if we define “less related” or “wider” diversification as expansion into businesses where the transferred factor creates less value. This can either be because the factor creates less value in any business or because it loses more efficacy in the specific transfer. Management skills may be an example of the first case, and the trans-national transfer of some brand names may illustrate the second. Most standard definitions of relatedness, and the ways they are operationalized in empirical work, fit the proposed definition.

V. SIMILARITIES WITH, AND DIFFERENCES FROM, THE RBV

Readers of this review will recognize many aspects of (1) - (6) from the RBV. Strategy is based on “resources”, productive factors that are semi-permanently tied to the firm (Caves, 1980). The literature often has the firm starting from scratch in the sense that both the original segment and later diversification are chosen in light of the resources. However, the idea is always to focus on those resources that can support competitive advantage, characterized by Barney (1991) as valuable, rare, inimitable, and non-substitutable (VRIN), and leverage them in those industries or segments where they are most valuable. It may still be necessary to invest in complementary “mundane” resources (Foss, 2005, p. 104), but it is the VRIN resources that matter. Trade in these resources is thought to be hampered by “transactions-cost”, and they are often
characterized as indivisible (Rumelt, 1974; Teece, 1982; Wernerfelt and Montgomery, 1986). It is finally well documented that related diversifiers do better than unrelated ones (Lang and Stulz, 1994; Montgomery and Wernerfelt, 1988; Rumelt, 1974; Wernerfelt and Montgomery, 1988). So the theory of the firm sketched in Sections II – III portrays strategy in a way that closely resembles the RBV.

On closer inspection, however, the theory slightly and subtly differs from several standard RBV beliefs:

(a) The theory defines “relatedness” between two businesses relative to a specific productive factor. It consists of two conditions: that both businesses use the services of the factor, and that the factor can be deployed in both businesses (as opposed to one of them) with relatively little extra customization. Since this concept of relatedness is multidimensional, many older and larger firms, with several unique resources, will have different neighbors along different dimensions. For example, some may be related through brands and others through technology. This raises the possibility that the firm could expand along more than one dimension. The RBV literature often defines relatedness in a much less concrete way and it is often implied that the concept applies independently of any specific resource. Furthermore, the RBV has not yet paid attention to the types of management problems that can arise when firms diversify in different directions based on different resources, or when two businesses are related in more than one dimension.

(b) The theory argues that scope of the firm is limited by the capacity of its factors and the number of sufficiently closely related businesses. As we read the RBV literature, it has been mostly focused on the first of these forces, perhaps because it is thought to be easy to find related industries once you have excess capacity of a unique resource.

(c) The theory suggests that trade in some productive factors, in some situations, are subject to sub-additive negotiation/bargaining costs and that this makes it inefficient to trade
fractions of them. It says that such factors, in such situations, should be leveraged by the firm to the extent possible, or otherwise traded in their entirety. In situations where needs are simple and change infrequently, the firm may leverage them at arms-length, but normally leverage takes place inside the firm.

In contrast to much RBV literature, the argument makes no distinction between unique and non-unique factors; it is the sub-additive trading costs and the need for adaptation that matters. It also does not require the factors to be ontologically indivisible; what matters is that the trading costs render them economically indivisible.

(d) The theory also makes it clear that resources take primacy over market. Firms should choose their markets in light of their resources, not the other way around. This does not mean that “Five Force Analysis” (Porter, 1980) is irrelevant; the firm may have a choice between several industries and will want to select that in which its returns will be the largest. This speaks to two subgroups within the strategy community. First, our (anecdotal) impression is that many mainstream scholars organize their strategy courses to start with industry analysis and only then proceed to resources, instead of the other way around. Secondly, some rabid RBV scholars (a group which used to include the author), have held the view that industry analysis does not survive game theoretical reasoning because it tells everybody to do the same thing. But this is obviously incorrect in light of the theory presented here. It is just important that the industries be evaluated in light of the firm’s resources.

(e) Finally, the negative relationship between diversification and performance can have two causes within the theory: More widely diversified firms may have resources that create less net value or they may have spread similar resources more widely than other firms. Again here, the RBV literature has been mostly focused on the first force, largely ignoring the second.

VI. DISCUSSION
Based on a micro-foundation, we have proposed a number of refinements to the RBV, constructed a framework to flesh out more new results, and clarified the meaning of the aggregate-level terms “firm” and “resources”.

In terms of further research, we see three directions. One comes from the concept of the firm: The idea that adaptability is an essential property has not been used before and should have a number of interesting implications in both theoretical and empirical directions. Secondly, while the role of resource capacity has been studied before, it has not had the prominent position which the ACT affords it. By pushing this further, it might be possible to develop ways of measuring and enhancing the capacity of a particular resource. Third, the fact that “relatedness” should be understood in the context of an individual resource, has interesting implications at both the normative and empirical level, especially when combined with the fact that many firms have several resources.

In conclusion, we admit that this essay, not entirely shamelessly, promotes the author’s own work, both in strategic management and in economics. In truth, however, it is just an example of the general point, that there is a close relationship between theories of strategy and theories of the firm.\(^{13}\) Scholars in both fields could benefit from contact with scholars in the other, and in the long run, the two bodies of work will doubtlessly converge. In our view, this cannot happen fast enough.

---

\(^{13}\) The functions and role of management is another area of inquiry in which scholars from economics and strategy are working in parallel with little contact. Economists came to the area much later, but recent work pioneered by John Van Reenen and his co-authors (Bloom and van Reenen, 2007; Acemoglu, Aghion, Lelarge, Van Reenen, and Zilibotti, 2007; Bloom, Propper, Seiler, and Van Reenen, 2015; and Bloom, Sadun, and Van Reenen, 2016 are but a few examples.)
REFERENCES


