Beyond strategy: a critical review of Penrose’s ‘single argument’ and its implications for economic development

How to cite:

© 2013 Routledge
Version: Accepted Manuscript
Link(s) to article on publisher’s website:
http://dx.doi.org/doi:10.1080/09672567.2013.792364

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online’s data policy on reuse of materials please consult the policies page.
Beyond strategy: A critical review of Penrose’s ‘single argument’ and its implications for economic development

Dr Richard K. Blundel, The Open University, Walton Hall, Milton Keynes MK7 6AA, United Kingdom; e-mail: r.k.blundel@open.ac.uk

Abstract

This paper offers a critical review of the ‘single argument’ that underpins Penrose’s (1959) study, The Theory of the Growth of the Firm (TGF). It aims to complement and counter-point recent examinations of Penrose’s influence on strategic management, and on the resource-based view (RBV) in particular, including Jacobsen’s (2011) study. The paper examines six components of the argument, tracing their inter-connected journey towards TGF’s relatively neglected final chapters, which address the economic consequences of the growth of large firms. It also reflects on the implications for economic development research, with reference to Penrose’s later (1992) critique of contemporary liberalisation policies.

Keywords

Edith Penrose, firms, growth, economic development
1. Introduction

The entire study is a single argument no step of which can be omitted without the risk of misunderstanding later conclusions. Edith Penrose, Preface to ‘The Theory of the Growth of the Firm’ (1959: xxii)

Edith Penrose was, for many years, a prolific and widely respected figure in academic and public policy circles. The Theory of the Growth of the Firm (TGF) was well received on publication, yet the following two decades saw few empirical applications or conceptual developments of its core ideas. Wenerfeldt’s influential (1984, 1995) article, ‘A Resource-based View of the Firm’, in the Strategic Management Journal, prompted an increase in citations in the strategic management literature. Levels of awareness rose further with the publication of Hamel and Prahalad’s (1990) managerially-oriented, The Core Competence of the Corporation. The renewed interest in TGF prompted David Musson, a senior commissioning editor at Oxford University Press, to propose a third edition (Musson 2002). The new edition was duly published, with a new Foreword by the author (Penrose 1995). Edith Penrose died in 1996. In 1999, the 40th anniversary of the publication of TGF, was celebrated in a special issue of Contributions to Political Economy, and subsequently in an edited book (Pitelis 2002), with contributions from eminent colleagues from several fields including strategic management, industrial and regional economics, technological entrepreneurship and international management. Since that time, her influence on the resource-based view has been the subject of fierce debate in the Journal of Management Studies (Clark et al. 2004), and her wider contribution to strategy and economics has been reviewed in a special issue of Managerial and Decision Economics (Thompson and Wright 2005), Dunning (2003) has reflected on her influence on international business scholarship, and a further special issue (of Management International Review) has considered her impact on multinational enterprise research (Pitelis and Verbeke 2007). More recently, Jacobsen (2011) has examined the influence on Penrose of an overlooked precursor, Robinson (1931), tracing in some detail the stream of ideas that flowed through TGF and, in due course, exerted a profound impact on the resource-based view.

Penrose warned her readers that TGF must be seen as an integrated whole (Penrose 1959: xxii). Yet, despite the plaudits, TGF has shared the fate of many ‘classic’ studies. While selected facets of the study are routinely cited, this is often done in ways that do not reflect its wide-ranging yet tightly woven argument (Foss 1999a, Pitelis and Wahl 1998a, 1998b, Pitelis...)

---

1 The exceptions include a small group of economists, who attempted to formalise one component of Penrose’s argument, the so-called ‘Penrose effect’, or managerial limit to growth in the 1960s and 1970s (Marris 1964, Rubin 1973, Shen 1970, Slater 1980b, Uzawa 1969), and Richardson’s (1972) study, which provided the initial bridge between TGF and the subsequent literature on inter-organisational networks (cf. Penrose 1996).

2 Augier and Teece (2007: 189 n12) note that David Teece’s (1982) paper on the multi-product firm was the first to apply Penrose’s ideas to strategic management issues.

3 The fourth edition of TGF was published in 2009 with a new Introduction by Christos Pitelis, the editor of Penrose’s collected papers (Pitelis 2009).

4 The link with E.A.G. (Austin) Robinson is recalled in a striking image of Edith and her husband E.F. ‘Pen’ Penrose, driving from her home in Syria to attend a job interview in England: ‘In 1959 Edith drove across the Syrian desert (Pen did not drive), through Turkey and on to England in an old Hillman estate car so that Edith could attend an interview at Cambridge University. [Austin’s wife, the economist] Joan Robinson, had read page proofs of The Theory of The Growth of The Firm and instigated the invitation. As the story goes, Austin Robinson was less impressed and a job offer was not forthcoming.’ (Best and Gurnsey 1999: F199). We can only assume that other forces were at work since, as Jacobsen (2011: 13 n12) indicates, TGF contains several references to Robinson’s 1931 study, The Structure of Competitive Industry. It also signals many points of agreement with those of her prospective employer (ibid.: 13-18).
In addition, much of the recent attention on Penrose’s ideas has been located within the field of strategic management, and the resource-based view (RBV) in particular. This literature will remain central to any assessment of Penrose’s contribution. It has generated many insightful contributions (e.g. Kor and Mahoney 2000, 2004, Lockett 2005, Lockett et al. 2009, Turvani 2001, 2002), as well as some striking divergences of view (cf. Lockett and Thompson 2004, Rugman and Verbeke 2002, 2004). However, one unfortunate side-effect has been to draw attention to particular sections of TGF, to the extent that other important components of the argument have become marginalised, and arguably less well-understood.

The aim of the present paper is to help redress the balance by offering a complement and a counter-point to recent debates in the strategic management arena. We argue that the explanatory potential of TGF extends beyond business and corporate strategy. Though it has proved capable of informing research, policy and practice in many fields, including innovation studies, inter-organisational relationships and regional growth dynamics (e.g. Best 2001, Kay 1997, Turvani 2002), relatively little attention has been paid to the book’s overall conclusions, which are largely concerned with the economic consequences of contemporary patterns of firm-level growth and their implications for public policy. In a period when much of the world is enduring severe economic, social and environmental dislocation, it seems timely to reconsider this broader Penrosean legacy. The approach we have adopted is to trace the inter-connections between six major components of the single argument, with a particular focus on their application to economic development and international political economy. In doing so, we also seek to highlight the connections between TGF and concerns that are well-represented in Penrose’s other works (e.g. Penrose 1965, 1971, 1985, 1992). Retracing Penrose’s ambitious journey requires a series of leaps across disciplinary boundaries and the spanning of multiple levels of analysis. While it is impossible to do full justice to the original arguments, our aim is to broaden the discussion of Penrose’s contribution to economic thought by drawing attention to TGF’s concluding chapters, which extended its core argument to address the relationship between large and small firms in a growing economy (Chapter 10), and the process of industrial concentration (Chapter 11).

2. Intellectual antecedents of the ‘single argument’

Before beginning our review, it is worthwhile considering briefly the range of sources that Penrose drew upon when constructing her ‘single argument.’ One set of ideas, which centred on the role of human purpose in the ‘black box’ of the firm, is derived from Penrose’s 1952 paper, ‘Biological Analogies in the Theory of the Firm’ and the ensuing debate with Armen Alchian (Alchian 1950, Penrose 1952, 1953). Another set can be traced to Penrose’s (1951) doctoral thesis, ‘The Economics of the International Patent System’, which was concerned with the
creation and transmission of knowledge. TGF builds on the work of a number of other writers who are cited in the book. Many of these were fellow economists, and it is clear that – at least in retrospect – Penrose’s work can be located within a broader post-Marshallian tradition (Finch 2000, Jacobsen 2011, Loasby 1991, 1999). Yet, despite her thorough grounding in orthodox economics, she adopted a highly unusual and eclectic approach, drawing freely on related disciplines, including contemporary writers on industrial organisation, such as Chester Barnard, P.S. Florence and J.K. Galbraith. Direct industrial experience provided the other major source of ideas, and one that helps to explain her openness to insights from other disciplines. Penrose’s post-doctoral research involved her in a fieldwork study during the summer of 1954 at the Hercules Powder Company. Penrose had originally intended that the resulting detailed case analysis of Hercules would form a chapter within TGF and stated that it, ‘was designed to illustrate the arguments of that study’ (Penrose 1960: 1)\(^8\). However, as Kay (1999: 67) has pointed out, it is also clear that this experience was, in its turn, an important influence on the development of those arguments.

It is difficult to appreciate the sheer audacity of Penrose’s departure from the prevailing neo-classical orthodoxy. Reflecting on her experience ‘25 years on’, Penrose noted that in the early 1950s, when she joined Fritz Machlup’s research project examining the growth of firms, ‘I elected to work on the theory’. (Penrose 1985: 6 – emphasis added). Her colleagues, the confident exponents of a sophisticated theory of the firm, were already equipped with established analytical techniques and saw little merit in theories of organisation:

‘[S]tudents of industrial economics were regarded as in a border area of “applied” economics. Sociologists, institutionalists, behavioural psychologists, business analysts (and especially business school teachers), though undoubtedly commendable fellows, were clearly of lesser scientific standing. They had no “hard” integrated theoretical foundation for their alleged disciplines.’ (Penrose 1985: 6)

Penrose was particularly fortunate in her choice of mentor. Fritz Machlup’s training was also orthodox: he had studied economics at the University of Vienna under two of its leading lights, Friedrich von Wieser and Ludwig von Mises. However, like Penrose, he was not a prisoner of his discipline. On graduating from Vienna, Machlup had combined his intellectual interests in economics and the philosophy of science with a practical business career in the family’s cardboard-manufacturing partnership (Chipman 1987: 267-268). Having subsequently relocated to the United States, he pursued a mix of ‘pure’ and ‘applied’ forms of research (Machlup 1967). As a consequence, Machlup was perhaps one of the few prominent economists of his generation prepared to countenance an exploration ‘inside’ the firm. Penrose reflected, seemingly with some amusement, on the reluctance of others to follow suit:

‘Few economists thought it necessary to enquire what happened inside the firm – and indeed their “firm” had no “insides”, so to speak. I do not say they were wrong, only that being theoretical economists, they saw reality differently from other people.’ (Penrose 1985: 7)

\(^8\) The publishers determined otherwise, and the case was published separately. However, while there may not have been sufficient space for the Hercules study, Penrose did include the final two chapters. (Chapters 10 and 11), which extend her argument to address a wider set of economic issues.
In the period following the Hercules fieldwork, Penrose’s attention was drawn decisively away from the traditional ‘theory of the firm’, towards the work of the pioneering economists of industry structure such as E.A.G. (Austin) Robinson and, ‘those from other disciplines treating the firm as an organisation’ (Penrose 1985: 7). Penrose was able to maintain the support of her mentor as a result of careful efforts to distance her innovatory theory of growth from the concerns of neo-classical orthodoxy. She shared Machlup’s views on the purpose-specific role of theory, a position she maintained throughout her career. While others might agonise over the merits of ‘integrationism’ (Foss 1999b), Penrose argued against attempts to blend the neo-classical theory of the firm with emerging theories of industrial organisation.

3. Linking components of the Penrosean theory of growth

Having considered some of its more influential antecedents, we can now turn our attention to the links between six major components Penrose’s ‘single argument’ (Figure 1). This discussion is based on a detailed reading of TGF and of Penrose’s other published works. It also draws on the diverse literature that has addressed aspects of TGF, and which includes contributions from economists, management researchers, business historians and others. While a review of this literature suggested a degree of consensus over TGF’s central themes, there were also differences of emphasis, often reflecting disciplinary backgrounds. For example, Slater’s (1980a) Foreword to the Second Edition of TGF offers valuable insights, but is largely oriented towards the implications for theoretical economists (cf. Slater 1980b), while Moroni (2006) integrates specific Penrosean arguments into a re-examination of the neo-classical theory of the firm. For the purposes of this paper, six components of TGF have been re-constructed in order to clarify the book’s main conceptual developments, and to show how they build towards the themes of the final chapters (Figure 1). The framing of the components has been informed by Kor and Mahoney’s (2000) 10-point categorisation and by Penrose and Pitelis’s (2009: 23-4) 13 point summary of TGF’s main ideas. While the six components do not represent an exhaustive summary, they place particular emphasis on the dis-equilibrating aspects of Penrose’s single argument (Foss 1999a, Loasby 1999a) and its many inter-dependencies. In contrast to previous listings, they also make a more explicit link between Penrose’s concept of ‘the interstices’ (Penrose 1959: 223), and preceding phases of the argument. This is an important addition because it helps to open up the relationship between TGF and Penrose’s other research in the fields of economic development and international political economy.

*INSERT FIGURE 1 ABOUT HERE*

9 Jacobsen (2011) provides a detailed account of the E.A.G. Robinson’s influence on Penrose and TGF, with a particular focus on the implications for strategic management and the resource-based view.
10 Penrose was characteristically caustic in her assessment of those who attempted integration: ‘Williamson finds in the development of the M-form a means of joining more fully the neoclassical theory of the firm and “bureaucratic theory”. He may be right to the extent that in the narrow sense the “profit maximisation hypothesis” becomes more applicable in the “real world”, but not if one holds as I do, that the two types of theory are designed to answer different questions and are therefore not to be compared in any meaningful way.’ (Penrose 1985: 13)
11 Kor and Mahoney’s (2000) ten-point schema is set out in the order in which the concepts appeared in Penrose (1959), while Penrose and Pitelis’s 13-point list (2009: 23-4) summarises their own account of TGF’s main ideas.
Component (1) ‘Authoritative communication’: bounding the firm

In an early passage entitled, ‘The Firm as an Administrative Organisation’, Penrose began to define the firm in a way that facilitated the kind of theorising that she wished to undertake. The basis for abstraction was related to the main objective of the book, since it determined those aspects of the firm that needed to be selected, and those that could be ignored:

‘The object of the present study is to investigate the growth of the industrial (non-financial) firm as an economic entity in its broadest sense [...] Consequently the definition of what constitutes a “whole firm” for our purposes depends on its essential function as an economic entity in the economy.’ (Penrose 1959: 15 – emphasis in original)

For Penrose, this test of relevancy meant treating the firm as a strategic decision-making unit (Components 3 and 4). This contrasted with the Coasian transactional analysis, as a basis for distinguishing firm-level co-ordinating activity from that taking place in the market:

‘The essential difference between economic activity inside the firm and economic activity in the “market” is that the former is carried on within an administrative organization while the latter is not.’ (Penrose 1959: 15)

Penrose illustrated the distinction by contrasting the activities of an industrial firm and a ‘purely financial’ investment trust. This led to the important conclusion that the area she termed ‘authoritative communication’ would define the boundaries of the firm for the purposes of The Theory of the Growth of the Firm. The term, ‘authoritative communication’ was an acknowledged borrowing from Barnard’s pioneering (1938) study, The Functions of the Executive. Her analysis of human motivation is, accordingly, focused on the perceptions and behaviours of the managerial team. In doing so, Penrose reflected her concern for human purpose in organisations (Penrose 1995: 9, Best and Humphries 2002: 56). The implication was that different analytical techniques might be needed, where researchers sought to explain the growth of holding company structures rather than the industrial firm, her stated subject (Penrose 1959: 20). She retained this view in later years, but recognised that by the 1980s, the ‘nature of the industrial firm’ had itself been transformed, ‘not, as I had suggested, to take the form of a financial holding company; instead it had metamorphosed into a hybrid product of the two types. This was a kind of genetic mutation largely due to electronic forms of communication and control. Only now is it being realized that this development requires a revision of much microeconomic theory’ (Penrose 1992: 252 n1). This insight, which is revisited in her new Foreword to TGF made an explicit call to examine the role of ‘authoritative communication’ across these new organisational forms, including inter-firm alliances and networks (Penrose 1995).

---

12 In a footnote, Penrose explained that she was not attempting to extend their work, but merely drawing upon those elements that might enhance her own theorising: ‘I am concerned only with those aspects of these large and complex subjects which will be of use in the theory of the growth of the firm to be developed later.’ (Penrose 1959: 16, n2)
Component (2) ‘Resources’ and ‘services’: a vital clarification
A short (two page) section entitled, ‘The Firm as a Collection of Productive Resources’ might be regarded as the ‘birthplace’ of the Penrosean approach to growth. Penrose’s first move was to establish the firm as more than simply an administrative unit. The firm had a ‘cohesive character’ (Penrose 1959: 24), which derived from the ‘authoritative communication’ of its managers and which justified separating it, for analytical purposes, from other economic groupings (Component 1). The activities of an industrial firm could be further distinguished by the fact that managers were free to deploy its productive resources in different ways over time (Penrose 1959: 24). Penrose’s apparent motivation for clarifying the nature of the resource base was to establish a more appropriate measure of firm size, a pre-requisite for her theory of growth. Having placed resources at the centre of her analysis, she made a second move, introducing the vital distinction between ‘productive resources’ and the ‘productive services’ that they create:

‘Strictly speaking, it is never resources themselves that are the ‘inputs’ in the production process, but only the services that the resources can render. The services yielded by resources are a function of the way in which they are used – exactly the same resource when used for different purposes or in different ways and in combination with different types or amounts of other resources provides a different service or set of services. The important distinction between resources and services is not their relative durability; rather it lies in the fact that resources consist of a bundle of potential services and can, for the most part, be defined independently of their use, while services cannot be so defined, the very word service implying a function, an activity. As we shall see, it is largely in this distinction that we find the source of the uniqueness of each firm.’ (Penrose 1959: 25 – emphasis added)

In other words, resources represented an economic input with an option value, whereas services were the output, the activity itself. In a brief but telling footnote, presumably directed at fellow economists, Penrose explained that she had rejected the conventional term ‘factors of production’ to describe these inputs, ‘precisely because it makes no distinction between resources and services’ (Penrose 1959: 25, n1). The interactions between these concepts are central to TGF, and are inseparable from the notion of ‘productive opportunity’ (Component 3). The following extract from an interview that Penrose gave to a doctoral student, indicates that she shared the view that some later resource theorists were conflating ‘resources’ and ‘productive services’, thereby missing the essential issue of conversion13:

‘[I] argued that the resource-based literature did not fully explain value-creation (instead focusing on the value appropriation aspect). Professor Penrose expressed strong agreement with my observations […] The problem was, she said, that the resource-based literature had not fully pursued her position and had been too concerned with the analytical properties of resources. The literature had, hence, partly neglected her fundamental insight that resources were only a means to an end.’
(Haanes 1997: 17 cited in Foss 1998: 5)

Spender (1994) was a notable exception, which recognised the Penroseean emphasis on the firm as a, ‘dynamic body in action’ (ibid.: 355). It represented an explicit rejection of the search for advantage amongst resources that are logically prior to the firm’s activities:

‘When we overlook the resource application processes we miss what it means to bundle resources together so that they become a firm. […] Resource-based theory has paid little attention to the construction and management of the bundle.’ (Spender 1994: 354 – emphasis added)

TGF focused precisely on these processes of construction and management, by linking the concepts of resources, services and productive opportunity (Component 3). In its original (1959) guise, the ‘resources-services’ dynamic was contained within the relatively well-defined administrative boundaries of the industrial firm, but it later proved to be applicable at an inter-organisational level of analysis. Richardson (1972: 883) introduced the image of industry as a, ‘dense network of co-operation and affiliation by which firms are inter-related’, and Penrose later drew on these ideas when discussing the ‘metamorphosis’ of the firm in the late 20th century (Penrose 1995: xviii-xx).

**Component (3) Productive opportunity: expectations and conjecture**

The notion of ‘productive opportunity’ (Penrose 1959: 31) is built explicitly on the opening conceptualisation of the firm (Components 1 and 2). For Foss (1999a: 95) it is arguably the key concept of TGF. However, it also represented a much bolder step in her argument, and one that has remained problematic for economic theorists. This is because it is predicated on both Knightian uncertainty - the variable (and inherently unquantifiable) capacity to perceive opportunities – and on a Schumpeterian interpretation of their differential capacity to exploit them. Penrose was conscious of the lack of conceptual clarity with respect to entrepreneurship and attempted to clarify this ‘slippery concept’, by distinguishing between ‘entrepreneurial’ and ‘managerial’ services (Penrose 1959: 33-41). However, the assertion that firm-level growth is limited by such a subjective concept as ‘productive opportunity’ represented a direct challenge to neoclassical theory (Kor and Mahoney 2000: 115; Foss et al. 2008). By incorporating this insight into an integrated theoretical explanation, Penrose located her single argument in direct opposition to the mainstream:

‘A theory of the growth of the firm is essentially an examination of the changing productive opportunity of firms; in order to find a limit to growth, or a restriction on the rate of growth, the productive opportunity of a firm must be shown to be limited in any period.’ (Penrose 1959: 31-32)

The section, ‘The Role of Expectations in the Productive Opportunity of the Firm’ saw Penrose in her most subjective mode of enquiry.14 It was an explicit extension of her vitriolic critique of explanations that failed to take account of purposive behaviour (Alchian 1950, Penrose 1952, 1953). It appears that this component in her argument was already largely developed when she came across Kenneth Boulding’s (1956) anticipation of social constructionism, *The Image*. Penrose was evidently impressed with this, ‘imaginative little book’, noting that, ‘Image is so apt a word for my purposes that I promptly appropriated it’ (Penrose 1959: 5, n5). The influence of the firm’s ‘environment’ was not ignored; it re-

14 These arguments are revisited in Foss et al. (2008), where Penrose’s approach is synthesised with Austrian economics and resource-based theorising.
surfaced in her analysis of the ‘interstices’ (Component 6). Rather, it was, ‘put on one side’ in order that she could focus on the internal resources of the firm (Penrose 1995: xiii). Even so, Penrose was at pains to emphasise the limits of subjectivism, acknowledging both the ‘reality’ of the external selection environment and clarifying its temporal relationship to the firm-level decision-making:

‘In the last analysis the “environment” rejects or confirms the soundness of the judgements about it, but the relevant environment is not an objective fact discoverable before the event.’ (Penrose 1959: 41 – emphasis added)

This observation led her to the conclusion that, because environmental variables are always mediated by firm-level influences, analysis of the growth process must also begin at this level (Penrose 1959: 42)\(^1\). In the Foreword to the Third Edition, Penrose emphasised two important corollaries of this argument. First, that each firm’s ‘productive opportunity’ was unique, and relatedly, that neither the rate nor the extent of the growth of the firm was determined by exogenous factors; both were influenced by firm-level activity:

‘The “relevant” environment, that is the set of opportunities for investment and growth that its entrepreneurs and managers perceive, is different for every firm and depends on its specific collection of human and other resources. Moreover, the environment is not something “out there”, fixed and immutable, but can itself be manipulated by the firm to serve its own purposes.’ (Penrose 1995: xiii – emphasis added)

Penrose (1959) had asserted that demand conditions per se should not limit growth of an industrial firm willing and able to diversify into new products and markets. Three decades on, she was able to claim some empirical support for this statement (Penrose 1995: xiii). However, it seems likely that this evidence also reinforced her pre-existing concerns over the implications for economies in which large firms, or inter-firm networks, occupy dominant positions (Component 6). It was an extension of her previous insight that, while growth at a particular rate may prove effective for a firm, the same might not be said for its wider economic and societal consequences (Penrose 1985: 14, 1992), a theme that she returns to in her concluding chapter.

In a closely-related discussion, ‘The Quality of Entrepreneurial Services’, Penrose explored the pivotal role of the firm as a context in which entrepreneurial judgement (Casson 2005) and agency were facilitated. She argued that explanations of a failure to grow were often incorrectly attributed to demand conditions when the underlying cause was an absence, or limited supply, of these ‘specific types’ of productive services (ibid: 37)\(^2\). ‘There is a recursive dimension to this process, with the organisational framework of the firm providing the necessary environment for the ‘production’ of entrepreneurial services, and these services, in turn, becoming a ‘significant aspect’ of the firm’s changing productive opportunity (Penrose 1959: 35). As in the case of the resource-services dynamic (Component 2), the concept of productive opportunity later proved to be applicable beyond the boundaries of the firm. For example, inter-actions between highly networked technology firms can create new

---

\(^1\) Adam Smith’s [1776] exploration also starts with the firm (Best and Garnsey 1999: F193).

\(^2\) Organisational researchers have since extended this argument, challenging over-optimistic assumptions about the capacity of management teams to manipulate firm-level capabilities (e.g. Clark 2000, Scarbrough 1998).
sources of productive opportunity in a dynamic regional cluster (Best 2001). These extensions of the original TGF argument gain additional momentum when we make the next connection, between productive opportunity and the ‘receding managerial limit’ (Component 4).

**Component (4) ‘The receding managerial limit’: agency and constraint**

Penrose’s examination of the internal dynamics of the firm revealed that the same mechanisms operated both to stimulate growth, and to limit the rate at which it occurred\(^{17}\). She regarded the distinction between ‘economies of size’ and ‘economies of growth’ to be of fundamental importance to her argument (Penrose 1959: 262 n1). In an effort to isolate the relevant factors, Penrose identified three ‘classes of explanation’ for the observed limit on the rate of growth: managerial ability; product or factor markets; and uncertainty and risk. Of these, managerial ability was identified as an internal (or endogenous) factor, markets were external (or exogenous), while uncertainty and risk were the product of both internal attitudes and external conditions. Penrose’s achievement was to introduce a radically different understanding of growth, which was predicated on the idea that organisational factors could reduce the inherent, managerial limit on the growth rate of an individual firm. This limit was, ‘by its nature temporary’ (Penrose 1959: 5), since it receded as a direct consequence of the interaction of the managerial team (Components 1, 2 and 3). The receding managerial limit was first expressed in economic terms, highlighting its ‘dis-equilibrating’ character:

‘[A]fter the completion of an optimum plan for expansion a new “disequilibrium” has been created in which a firm has new inducements to expand further even if all external conditions (including the conditions of demand and supply) have remained unchanged.’ (Penrose 1959: 5)

This was followed by one of the most widely-quoted passages, which provides a succinct summary of this phase of the argument:

‘As management tries to make the best use of the resources available, a truly “dynamic” interacting process occurs which encourages continuous growth but limits the rate of growth.’ (Penrose 1959: 5)

In essence, the Penrosean firm required managerial resources with (necessarily pre-existing) experience within the firm in order to absorb new managers and other resources. Since this was not something that could be acquired from the market, it represented a necessary limit on the rate of growth (Penrose 1959: 45-46)\(^{18}\). However, the managerial constraint did not limit the extent of growth. On the contrary, the infusion of new blood initiated its own dynamic: once each increment of growth was completed, managerial resources became available for further expansion. This was because activities could be routinised, economising on cognitive

\(^{17}\) These were, in her terms: ‘The forces inherent in the nature of firms which at the same time create the possibilities for, provide the inducements to, and limit the amount of expansion they can undertake or even plan to undertake in a given period of time’. (Penrose 1959: 4-5).

\(^{18}\) Marris (1999: 51) noted how this concept echoed his father’s remarks on troop reinforcements (i.e. ‘they need time to bed down; they need time to learn each other’s ways’). The post-Marshallian economist, Philip Andrews also drew attention to internal reserves as a constraint on the rate of growth of firms, but in contrast to Penrose concentrated on physical capacity rather than managerial resources (Finch 2000: 390-1).
effort, and allowing managers to consider new possibilities. This concept became identified as the ‘Penrose curve’ or ‘Penrose effect’ and the main focus for formalisation\textsuperscript{19}.

Penrose’s attention appears to shift at this point in the argument. Having identified a series of mechanisms that help to explain the continuing growth of large firms, the final chapters of TGF turn towards the resulting patterns of growth over time (Component 5) and their wider economic, and to a lesser degree ‘social’, consequences (Component 6). With the concept of productive opportunity (Component 3), she had already made a connection between the individual-level, in the form of the subjective experiences and expectations of managers, and objective factors in the firm’s operating environment. However, Component 5 can be seen as providing the main bridge between the firm-level focus of the early parts of TGF, and the extension of the argument into the sphere of economic development, which is manifested in the final chapters (Figure 1).

Component (5) ‘History matters’: knowledge and evolution
Having established the primary mechanisms driving the growth of the firm, Penrose’s next step was to map out the patterns that were generated by the unfolding of these interconnected processes over time. Growth in the Penrosean firm was a product of its past activities, and the knowledge that these activities had generated:

‘One of the primary assumptions of the theory of the growth of firms is that “history matters”; growth is essentially an evolutionary process and based on the cumulative growth of collective knowledge, in the context of a purposive firm.’ (Penrose 1995: xiii)

In a Penrosean interpretation, the firm provides an institutional setting for conjecture and innovation in which productive services can be evaluated against, and directed towards, perceived productive opportunities, with Penrose’s ‘entrepreneurial’ and ‘managerial’ services distinction (Component 2) providing the basis for a variety-generating dynamic. As we have already noted, Penrose’s Foreword to the Third Edition of TGF acknowledged more recent work on knowledge and evolution, including studies that extended the Penrosean argument beyond the boundaries of the firm. She also complimented Best (1990) for demonstrating how Schumpeterian selection (Schumpeter [1943] 1954: 81-86) could be reconciled with the Penrosean mechanisms of knowledge-based growth on a regional and national scale\textsuperscript{20}. While gales of creative destruction might sweep away particular products or technologies, industrial firms were not simply passive containers; they could be rejuvenated through human agency (Kay 1997: 82). The Penrosean firm supported this evolutionary growth dynamic by making ‘connections’, between past activities and future options (Loasby 2001: 10, Penrose 1995: xvi). Penrose had re-conceptualised the firm as a, ‘dynamic body of knowledge in action’ (Arrow 1962, Spender 1994: 355). It provided a, ‘kind of temporary evolutionary equilibrium’ (Penrose 1995: xiv), where managerial teams could conjecture over the application of productive services to particular opportunities, thereby moderating the ‘perennial gale of creative destruction’ (Schumpeter [1943] 1954: 81-86, Kay 1997). Penrose was careful to qualify the scope for such conjecture and recognised that it was doubly

\textsuperscript{19} Penrose later noted that, ‘the so-called “Penrose curve” […] has been applied in a number of contexts, and even, to my surprise, to agricultural enterprises.’ (Penrose 1995: xii).
\textsuperscript{20} Best (1990: 138) had argued that the relative success of Japanese firms in the preceding decade could be explained by just such a conjunction.
situated. Firstly, the productive services that managers could prevail upon were the product of the firm’s ‘unique’ history, which in its later (Penrose 1995) variants is also entrenched in interlocked patterns of inter-organisational relationships. Secondly, the managerial team’s perception of productive opportunities was shaped by previous experience, and by the shared interpretive frameworks that it had engendered. These insights carried important implications for economic development: Penrose saw that the continued prosperity of any economy was dependent on a quasi-evolutionary process in which the outcomes were far from certain. She addressed this concern in the final chapters of TGF by considering whether the rise of ‘big business’ in the industrialised economies of the 1950s was compatible with the survival of smaller firms (Component 6).

Component (6) The ‘interstices’: securing economic development
Penrose’ final major move, in The Theory of the Growth of the Firm, was to introduce the concept of the ‘interstices’ of the economy. Interstices were defined as windows of opportunity for the growth of smaller firms, which arose due to the inability of larger firms to exploit all of the ‘productive opportunities’ generated by their own growth (Penrose 1959: 221-25). This component is somewhat underplayed in accounts of the Penrosean legacy21, yet it represents an important step in the ‘single argument’, extending its underlying logic beyond the intra-firm level, enabling it to incorporate inter-firm relationships (Richardson 1972, 1999), and opening it up to new fields of study, including industrial and regional dynamics (e.g. Best 1990, 2001, Blundel and Thatcher 2005). In the final chapter of TGF, Penrose emphasised that she was concerned with, ‘the appraisal of the place of large and growing firms in the modern world’. (Penrose 1959: 261 – emphasis added), and noted that the innovating competition of large firms could, ‘degenerate into almost senseless competition’ to satisfy consumer demand, (ibid.: 262 n2). She signalled her move from the analysis of growth within the firm to a consideration of the wider competitive environment in the closing sentences of Chapter 9, by removing one of the simplifying assumptions that had been used in earlier chapters of TGF. Her purpose was to ensure that the theoretical explanation remained sufficiently ‘general’ to encompass any mechanism that exerted a systematic influence on the growth of industrial firms. She pointed out that TGF had so far failed to account the mechanism that facilitated the growth of an entire sub-category, the small industrial firm in environments characterised by ‘big business’ competition:

‘Thus we evaded what is widely held to be the characteristic position of the small firm in a developed economy – an inability to compete with large firms, an inability which precludes its growth into those areas particularly suitable for the operations of larger firms.’ (Penrose 1959: 214)

If the ‘characteristic position’ of small firms was that pre-existing structural factors (i.e. in Penrose’s (1959: 214) terms, ‘the “environment” in the shape of competition from large firms’) determined their growth, then the other Penrosean mechanisms (i.e. Components 1 to 5), would be nullified and environmental conditions would set an absolute limit on the growth of small firms, ‘regardless of their resources or entrepreneurial ability’ (ibid.: 214). This was a crucial point for Penrose’s general theory of growth because TGF’s explanatory claims would be undermined if they failed to account for such a generalised phenomenon.

---

Chapter 10, ‘The Position of Large and Small Firms in a Growing Economy’ is an explicit attempt to resolve this problem, and to incorporate the growth of small firms into the Penroseean framework. In so doing, it refined the simplified conceptualisation of the environment that had been adopted in preceding chapters (Penrose 1959: 215). The main target for refinement was the highly subjectivist approach to the environment used to introduce the concept of the ‘productive opportunity’ of the firm (Component 3). Penrose acknowledged that this strand in her argument was based on the assumption that opportunities for expansion did, in some objective sense, ‘exist’\(^2\). This might not always be the case; larger firms in a sector could already enjoy size-related economies, while the expansion of their smaller rivals was compromised by structural barriers. In a rarely-cited section, Penrose examined, ‘The Continued Existence of Small Firms’ in the face of seemingly overwhelming competitive disadvantages. She developed a characteristically dynamic interpretation of the problem, which also served to highlight the limitations of prevailing static and cross-sectional interpretations:

\[\text{‘It would seem that at any given time a fair number of small firms would be in existence simply because they were young, and that at a later date the same firms would have developed into medium-size or large firms. This possibility, however, is rarely included among the explanations advanced for the existence of small firms, the analysis usually being presented in terms of the economies and diseconomies of size, using a kind of “static” or cross-section approach.’ (Penrose 1959: 220)}\]

Penrose listed four conventional explanations for the continuing existence a broadly-defined class of small firms, ‘that do not grow or, at least, that do not grow very much’ (Penrose 1959: 220): (1) Firms undertaking activities unsuited to large firms; (2) Firms protected by large firms for public relations reasons; (3) Firms in industries where entry is easy; (4) Firms in industries where ‘big firms have not got around to mopping them up’. Penrose’s fundamental point was that these four categories of explanation were insufficient to account for observed changes in the population of firms or for the continuing emergence of new entrepreneurial small firms in the face of such powerful incumbents (Penrose 1959: 221-222). Her solution to this anomaly began with the assumption that large firms did enjoy a generalised competitive advantage over the newer and smaller ones. Given the inherent limitation on the growth of larger firms, even under the ‘most favourable’ conditions (Component 4), their growth would necessarily result in the creation of unexploited productive opportunities. Small firms could grow into these temporary windows of opportunity, or ‘interstices’ (Penrose 1959: 222-223).

In Penrose’s original (1959) formulation, the interstices were a by-product of large firm activity. The therefore represented a kind of residual option set, which might be invaded by the large firm in a way that destroyed the small firm’s productive opportunity, ‘either driving it out of business or purchasing it outright’ (Penrose 1959: 223 n1). She used her own observations of corporate activity in the 1950s to demonstrate the dynamics of the interstices, indicating the ways in which the decisions of large US firms affected the small firm

\(^{22}\) She defers in characteristic style to the pragmatic and the empirical practitioner: ‘Now none but the most philosophically sophisticated businessman will accept the proposition that the opportunities for the expansion of his firm are simply his ideas about what his firm can do; he will insist that the opportunities he sees reflect the “facts” of the world, facts that may be known with indifferent accuracy to be sure, but facts none the less.’ (Penrose 1959: 216)
population. Rapid increases in demand might lead (profit-maximisation oriented) large firms towards increasing output of their existing products, since this was where ‘managerial effort per dollar of expansion’ was at its lowest. This contrasted sharply with the situation where new industries and technologies were being developed:

‘In the earlier stages of rapid industrial development the interstices may be very wide and numerous simply because the established firms are so few and because many new industries are coming into being. There seems considerable evidence, however, that very quickly each of the major industries tends to become dominated by a few large firms and a high degree of concentration develops early.’ (Penrose 1959: 224)

TGF’s concluding chapter is well summarised in its title, ‘Growing firms in a growing economy: the process of industrial concentration and the pattern of dominance’. It synthesises the earlier analysis of firm growth processes and the consequential effects on industrial structure. The chapter is distinctive for its extremely heavy footnoting, which seems to indicate a perceived need to re-integrate her complex web of argument into the mainstream economics literature (it also suggests that there would have been ample scope for a sequel to TGF that examine more its implications for economic policy in greater detail). Penrose considers several industrial policy issues in this chapter, including the impact of barriers to entry (Penrose 1959: 231-2) and corporate mergers (ibid.: 238-43). Her analysis confronts the prevailing assumptions of US corporatism head-on and includes a measured critique of Galbraith’s (1956) argument that market power would protect the incentive for technical development. She acknowledged the Galbraithian argument that ‘big’ business competition was different from that of previous times and could be, ‘even more beneficial to the economy and more effective in meeting consumers’ wants’ (Penrose 1959: 233), but while effective monopoly was self-evidently unacceptable, oligopolistic competition with entry barriers posed more ‘complicated’ dilemmas (ibid.: 234). Penrose chose to avoid the ‘essentially political’ approach to this question, preferring instead to focus on the effect on economic output (ibid.: 235). However, she also noted that there was ‘no easy way of resolving this issue on economic grounds alone, …’ (ibid.: 236). The final paragraph of TGF encapsulates the paradoxical nature of the growth process, and the challenges it poses for policy-makers:

‘If the case for big firms, and for big business competition, is a strong one, its strength rests on conditions that are not self-perpetuating, but may themselves be destroyed by collusion, by the extension of financial control, and by the struggle to resolve the contradictions in a system where competition is at once the god and the devil, where the growth of firms may be efficient but where their consequent size, though not in itself inefficient, may create an industrial structure that impedes its own continued growth’. (ibid.: 265)

Edith Penrose’s later writings indicate that she retained her parallel interest in economic development (Penrose 1985, 1992; Penrose et al. 1992, Pitelis 2011: 77). The international growth of firms was one of her other major research interests, particularly the activity and impact of oil companies in developing countries (e.g. Penrose 1965, 1971). Reviewing the ‘burgeoning’ literature in this field, she concluded that her original analysis, ‘seems by and large to apply equally well to expansion by foreign direct investment in its modern form – the processes of growth, the role of learning, the theory of expansion based on human and other resources, the role of administration, the diversification of production, the role of merger and acquisition are all relevant’ (Penrose 1995: xv). However, she also felt that new analyses
were required to cope with the more complex transnational corporation and its relationship with the market (ibid.:xvi). For example, in critiquing the ‘alleged superior “efficiency”’ of the Williamsonian M-form firm she revisited her earlier concerns regarding the restriction of the spread of entrepreneurial services in an economy that becomes increasingly dominated by large M-form conglomerates (Penrose 1985: 14). Recent efforts to set out new directions for Penrosean-inspired research on multinational enterprises (MNEs) (e.g. Buckley and Casson 2007, Pitelis and Verbeke 2007) have a rather different emphasis. The emerging research agenda is largely oriented around the interests of MNEs in securing firm-specific advantages (FSAs), exercising dynamic capabilities and ‘melding location-bound and non location-bound knowledge’ (Pitelis and Verbeke 2007: 142). While questions of this kind remain valid in their own terms, they do not engage directly with Penrose’s own later concerns, which focused around the impact of large firms on local and regional economies.

Rugman and Verbeke’s (2004) critique of Penrose’s contribution to the resource-based view provides an interesting perspective on this issue. One of their central criticisms revolves around what they regard as the ‘normative agenda’ that Penrose pursues in her work on international economics and international business (ibid. 214). This leads them to the proposition that, ‘If Penrose were alive today, it is rather safe to assume that she would harshly criticize the lack of focus on societal outcomes, in terms of both efficiency and equity, and the sole emphasis on rents, characterizing much of the present RBV work’ (ibid. 215). This assertion finds some support in one of Penrose’s final contributions to the field of economic development (Penrose 1992), a critique of recent economic liberalisation policies in developing countries. However, contra Rugman and Verbeke (2004), Penrose’s argument is firmly grounded in the language of economic analysis. In the paper, she questions two prevailing assumptions of economic liberalisation: firstly that case for ‘integration’ into the global economy is based on sound evidence; and secondly that global capitalism is itself ‘liberal’, given the rise of ‘multi-corporate conglomerate enterprises’ (Penrose 1992: 241). Penrose entitles her concluding section, ‘circumspect vigilance’ with the deliberate aim of highlighting the responsibility of government for the economic welfare of its people. Her analysis echoes the concluding sections of TGF, including the doubts she raised about Galbraith’s (1956) concept of countervailing power:

‘Those placing strong emphasis on the importance of economic liberalization tend to see ‘the market’ as the necessary and desirable countervailing power to that of the state. To a considerable extent most people would probably accept this. Equally dangerous to the liberal ideal, however, would be the hijacking of the state’s economic policy by corporate business; in the international field this would yield a kind of corporate world economic governance. But economic policy also encompasses various kinds of social policy and these raise issues affecting the allocation of resources on other than economic criteria, except to the extent that the overall ability and willingness of an economy to pay must be respected.’ (Penrose 1992: 251)

4. Concluding remarks

23 Pitelis (2011: 77-8), making reference to Penrose’s earlier work in the 1950s and 1960s, notes that her ‘uncompromising, objective account’ of the positive and negative effects of MNEs in developing countries attracted criticism from academics and company executives. Penrose’s later (1992) article reflects on a further two decades of MNE activity in developing countries and its critique of MNEs and host country governments is arguably more radical than that of her original (1965, 1971) contributions.
TGF is characterised by an unusual richness and diversity of concepts (Marris 1961: 144, Penrose 1959: xxii), which range from the psychological traits of entrepreneurs to the role of small firms in the industrial policies of nation-states. Penrose’s lasting achievement was to combine such disparate concepts into a coherent theoretical explanation of the growth of the firm. She was careful to isolate her project from that of neo-classical orthodoxy, but did not hesitate to make a radical break with academic convention. Her research methods were atypical, but perhaps the most significant signal of her ‘maverick’ status was a willingness to breach traditional disciplinary divides (Best and Humphries 2003, Clark and Blundel 2007). Penrose’s multi-disciplinary approach was an essential, pioneering act in pursuit of her research aims. She had, in effect, resumed Adam Smith’s search for ‘connecting principles’ in the field of production and wealth creation (Best and Humphries 2003: 63, Loasby 1999a, Skinner 1987). Her project took shape as an examination of the growth of knowledge within administrative boundaries of the firm, a theme that Marshall had developed out of Smithian classical economics, and which can now be seen as making a major contribution to an emerging neo-evolutionary research tradition (Finch 2000: 379, Foss 1998, Jacobsen 2011: 2-3, Lavezzi 2003, Loasby 1991). Penrose’s instinctive approach was to combine inductive research methods with disciplined analysis and, as Kor and Mahoney (2000: 129) have noted, while TGF is written in the style of deductive economic theorising, her research methods were informed by inductive reasoning. Penrose was also happy to draw on subjectivism (e.g. managerial perceptions of ‘productive opportunity’), and objectivism (e.g. the role played by ‘interstices’ and pre-existing industrial structures in constraining and enabling the growth of smaller entrepreneurial firms), in pursuit of a better explanatory theory. Furthermore, in extending the scope of her investigation to include the institutional context for growth, she was able to scrutinise prevailing models of economic development in industrialised and developing countries – and raised a number of questions that remain pertinent to the present day. TGF proved to be one of Penrose’s lasting achievements in economics, providing a coherent explanatory theory capable of accounting for the dynamic and idiosyncratic process of knowledge generation and application within the growing firm. Yet for two decades following its publication, TGF attracted limited interest from economists, with active engagement being limited to formalisation of the ‘Penrose curve’ (i.e. Component 4). Over the last three decades, strategic management scholars have re-engaged with Penrose’s ideas, encouraging a somewhat wider appreciation of TGF but still concentrating their attention on the first nine chapters (i.e. Components 1 to 4). In this paper, we have argued that there much to be gained by drawing on the full scope of TGF’s ‘single argument’ (i.e. Components 1 to 6), and in finding new ways of deploying this rich resource.

Marris’s (1961) review in the Economic Journal remarked that, ‘The book is indeed so packed with ideas that it would be impossible for all of them to be consistent’. (Marris 1961: 144). However, also he maintained the position that TGF, ‘lacked an economically interpretable account of the motives growth’ (Marris 1999: 48 – emphasis in original).
Acknowledgements

I would like to thank Michael Best, Peter Clark, José Luis Cardoso and the two anonymous referees for their helpful comments and suggestions. Though not consulted directly on this paper, I would also like to acknowledge the earlier generosity of Mark Casson, David Musson and Michael Rowlinson in making time to share their thoughts on Edith Penrose’s legacy. Lastly, belated thanks to Tony Lawson and Christos Pitelis for organising a research seminar on Edith Penrose and critical realism at Cambridge in 1999, which provided much of the initial inspiration for this study. Any errors and omissions remain my responsibility.

References


---

**Figure 1: Connecting the six components**

1: Authoritative communication: bounding the firm
2: ‘Resources’ and services: a vital clarification
3: ‘Productive opportunity’: expectations and conjecture
4: ‘The receding managerial limit’: agency and constraint
5: ‘History matters’: knowledge and evolution
6: The ‘interstices’: securing economic development

Strategic management

Economic development