

Northumbria Research Link

Citation: Burton, Nicholas (2016) Architectural co-evolution and correspondence in UK personal pensions. Doctoral thesis, Northumbria University.

This version was downloaded from Northumbria Research Link:
<http://nrl.northumbria.ac.uk/id/eprint/28428/>

Northumbria University has developed Northumbria Research Link (NRL) to enable users to access the University's research output. Copyright © and moral rights for items on NRL are retained by the individual author(s) and/or other copyright owners. Single copies of full items can be reproduced, displayed or performed, and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided the authors, title and full bibliographic details are given, as well as a hyperlink and/or URL to the original metadata page. The content must not be changed in any way. Full items must not be sold commercially in any format or medium without formal permission of the copyright holder. The full policy is available online: <http://nrl.northumbria.ac.uk/policies.html>



**Northumbria
University**
NEWCASTLE



UniversityLibrary

Architectural co-evolution and correspondence in UK personal pensions

Nicholas Burton

PhD

2016

A thesis submitted in partial fulfilment of the requirements of the University of Northumbria
at Newcastle for the degree of Doctor of Philosophy

Research undertaken in
Newcastle Business School
March 2016

Abstract

A firm's ability to survive and prosper is often a function of its ability to design and develop new products that meet the needs of heterogeneous markets. The way in which a product is designed can have profound implications for product market structure and who is able to profit from an innovation, but despite this few industry studies have examined how and why product and industry architectures co-evolve and correspond across time.

Notions of architectural co-evolution and correspondence are grounded in the modularity literature and assume a path towards increasing product modularity and industry specialisation. However, scholars have recently hinted that a reverse path towards increasing product and industry integration may be equally feasible. This research study contributes to the literature by proposing three stylised hybrid product and industry reintegration types that enhance our understanding of how and why reintegration may occur in product markets. Furthermore, the presence of a correspondence in the design characteristics between architectural layers (the so-called 'mirroring hypothesis') has also been suggested in the literature, such that product component design is often a blueprint for the way task, knowledge and firm boundaries are partitioned within a given product market. This research study finds that architectural correspondence is hard to sustain over time as firms often maintain a broader knowledge than task boundary for strategically important product components that offer differentiation opportunities or competitive advantage, contributing to the literature on contingencies that 'mist the mirror'.

Of particular interest to this research study is the UK personal pensions sector, a non-physical product, largely under-explored in the product modularity literature. By analysing the co-evolution and correspondence of a non-manufactured product over a 30-year period this research study breaks new ground. The research study makes use of a retrospective longitudinal research design, based upon semi-structured interviews with a purposive sample of 31 key personnel. The interview data was subject to a combination of matrix and template analysis.

Table of contents

<i>List of figures</i>	4
<i>List of tables</i>	6
<i>List of appendices</i>	8
<i>Acronyms</i>	9
<i>Acknowledgements</i>	11
<i>Declaration</i>	12
Main chapter headings	
1. Introduction	13
2. Literature Review	23
2.1 Architectural co-evolution	25
2.2 Architectural correspondence	67
3. Research Methodology, Design & Methods	82
4. Data analysis & Findings	139
5. Discussion	214
6. Conclusions and contribution	262
<i>References</i>	275
<i>Appendices</i>	299

List of figures

1. UK pensions landscape
2. The integrated-modular continuum
3. Component coupling and ownership properties of stylized product architecture types
4. Example of loose-coupling of two modular product components
5. Stylised product architecture typology
6. Within- and across-firm mirroring hypothesis
7. Outline of the research methodology
8. Example product time-line
9. Data analysis process
10. Insider – outsider: balancing the purposive sample
11. Product template: mid to late-1980s
12. Closed and integrated product architecture = with-profits personal pension
13. Product template: early to mid-1990s
14. Closed and modular product architecture = unit-linked personal pension
15. Product template: mid-1990s to mid-noughties
16. Hybrid product architecture = fund supermarket
17. Product template: mid-noughties to 2012
18. Open and modular product architecture = product platform
19. Product template: 2012-2014
20. Hybrid product architecture = product platform
21. Firm template: mid to late-1980s
22. Firm template: early to mid-1990s
23. Firm template: mid-1990s to mid-noughties
24. Firm template: mid-noughties to 2012
25. Firm template: 2012 to 2014
26. Industry template: mid to late-1980s
27. Industry template: early to mid-1990s
28. Industry template: mid-1990s to mid-noughties
29. Industry template: mid-noughties to 2012
30. Industry template: 2012 to 2014
31. Mirroring template: mid to late -1980s
32. Mirroring template: early to mid-1990s

33. Mirroring template: mid-1990s to mid-noughties
34. Mirroring template: mid-noughties to 2012
35. Mirroring template: 2012 to 2014
36. Product architecture evolution
37. Innovation modes
38. Dynamic effects of productive capabilities and efficiency considerations on modular product component innovation
39. Dynamic effects of productive capabilities and efficiency considerations on integrative innovation
40. Information-sharing: mid to late-90s
41. Varying information-sharing needs based on the value characteristics of the product component
42. Information-sharing: 2005 onwards

List of tables

1. Dimensions of stylised product architectures
2. The construction of the sample
3. The length of continuous professional experience of interview respondents
4. Pilot sample
5. Respondent recruitment process
6. Initial *a priori* matrix
7. Final matrix
8. Example of coding textual data to themes
9. Initial template: product, firm and industry themes mid to late-1980s
10. Initial template: mirroring mid to late-1980s
11. Initial template: product firm and industry themes early to mid-1990s
12. Initial template: mirroring early to mid-1990s
13. Initial template: product firm and industry themes mid-1990s to mid-noughties
14. Initial template: mirroring mid-1990s to mid-noughties
15. Initial template: product firm and industry themes mid-noughties to 2012
16. Initial template: mirroring mid-noughties to 2012
17. Initial template: product firm and industry themes 2012-2014
18. Initial template: mirroring 2012-2014
19. Final template: product firm and industry themes mid to late-1980s
20. Final template: mirroring mid to late-1980s
21. Final template: product firm and industry themes early to mid-1990s
22. Final template: mirroring early to mid-1990s
23. Final template: product firm and industry themes mid-1990s to mid-noughties
24. Final template: mirroring mid-1990s to mid-noughties
25. Final template: product firm and industry themes mid-noughties to 2012
26. Final template: mirroring mid-noughties to 2012
27. Final template: product firm and industry themes 2012-2014
28. Final template: mirroring 2012-2014
29. Phase one: the 90s – modular innovation
30. Phase two: 2000–2005 – architectural innovation
31. Phase three: 2005-2012 – architectural innovation
32. Phase four: 2012-2014 – integrative innovation

33. Integrative innovation modes
34. Evolution of systems integration capabilities
35. Summary of mirroring across time
36. Correspondence summary: closed and integrated product architecture
37. Correspondence summary: closed and modular product architecture
38. Correspondence summary: hybrid (unit-linked) product architecture
39. Value characteristics of product components on knowledge integration, information-sharing and location
40. Correspondence summary: hybrid (fund supermarket) product architecture
41. Correspondence summary: open and modular and hybrid product architecture

List of appendices

1. Interview date, firm background and job title of interview respondents
2. Interview schedule
3. Product architecture descriptions
4. Synthesised product timelines
5. Informed consent form for research participants
6. Summary of evolution towards a closed and modular product architecture
7. Summary of evolution towards a hybrid product architecture
8. Summary of evolution towards an open and modular product architecture
9. Summary of evolution towards a hybrid (platform) architecture

Acronyms

A-Day – Pensions simplification 2006

CIS – Collective Investment Scheme

CMS – Client Management Systems

DFM – Discretionary Fund Manager

D2C – Direct To Consumer

FCA – Financial Conduct Authority

FSA – Financial Services Authority

IFA – Independent Financial Adviser

IP – Intellectual Property

IPA – Interpretative Phenomenological Analysis

IPR – Intellectual Property Rights

ISA – Individual Savings Account

KBV – Knowledge-Based View

PEP – Personal Equity Plan

PIA – Personal Investment Authority

RBV – Resource-based view

RDR – Retail Distribution Review

SBU – Strategic Business Unit

SERP – State-Earnings-Related Pension

SIPP – Self-invested Personal Pension

SRA – Social Research Association

SSO – Standards-Setting Organisations

TCE – Transaction Cost Economics

TPA – Third-Party Administrator

Acknowledgements

To my wife Caroline for your understanding, support and unbelievable patience throughout the project and especially during the writing up stage

To my daughter Annabel, I hope my hard work will be an inspiration to you in your forthcoming A-level exams

To my principal supervisor Dr Pushkar Jha for your guidance and support, navigation through the red tape and your clarity and focus when the mists came down

.

To my secondary supervisors, Professor Peter Galvin and Professor Nigel King, for your advice and inspiration

To all interviewees for contributing to the research study

To Mark Polson and Richard Kelsall for assisting me with your expertise

To my NBS room-mates Rosaline and Richard for putting up with me during the process

To NBS for giving me a semester-long sabbatical free of teaching and marking

To my flock of Ryeland sheep for listening to me when I needed a break!

Declaration

I declare that the work contained in this thesis has not been submitted for any other award and that it is all my own work. I also confirm that this work fully acknowledges opinions, ideas and contributions from the work of others.

Any ethical clearance for the research presented in this thesis has been approved. Approval has been sought and granted by the School Ethics Committee on 29 September 2014.

I declare that the word count of this thesis is 83093 words, excluding references and appendices.

Name: Nicholas Burton

Signature:

Date:

1. Introduction

1.1 Background

The management of innovation across time is a key concern for managers. With external factors such as increasing deregulation, internationalisation of product markets, and fragmentation and specialisation occurring in many product markets, coupled with the economic and financial turbulence that began in 2007-8, a key challenge for firms is to be able to respond to the inherent complexity in managing new and existing product lines. In many instances, whether firms develop physical products or service-orientated propositions, many are responding to these challenges by shortening development cycles and increasing product variety in order to customise product or service designs for the needs of heterogeneous markets and segments.

The product design concept of modularity¹ has, therefore, quickly gained favour in many product markets, as diverse as bicycles (Galvin & Morkel, 2001); motor vehicles (Argyres & Bigelow, 2010), semiconductors (Funk, 2008), stereo systems (Langlois & Robertson, 1992), and air-conditioning systems (Furlan, Cabigiosu & Camuffo, 2014) to name but a few, and the research community has significantly increased in importance over the last 20 years, as indicated by the sheer volume of theoretical and empirical contributions. Although the concept of modularity spans across a number of diverse literatures and disciplines from mathematics and engineering to operations management, one strand of the modularity literature focuses on product design suggesting that product architectures can be partitioned into loosely-coupled² product components³ which interact together as a coherent whole through standardised interfaces, permitting greater degrees of substitutability within and across products and product families, and hence firms can benefit from economies of substitution (ie, Garud & Kumarswamy, 1995).

¹ The concept of 'modularity' is discussed in detail in Chapter 2. Here, modularity can be conceptualised as the degree to which any system can be separated/partitioned into distinct sub-systems/components

² Loose-coupling and tight-coupling were terms coined by Orton and Weick, (1990) to reflect the ease-of-separation of an architecture into distinct sub-systems or components. In other words, integrated product architecture are tightly-coupled and modular product architectures are loosely-coupled

³ A product component can be defined as a distinct portion of a product that embodies a core design concept (Clark, 1985)

Strong advocates of modularity as a design principle⁴ (ie, Fine, 1998; Sanchez & Mahoney, 1996; Schilling, 2000) have stressed its potential benefits ranging from faster and technologically-superior product component upgrades, self-reinforcing network externalities, through to more efficient supply chain partnerships, and there is growing empirical evidence which supports increasing product modularity (and to a lesser extent, increasing integration⁵) across a number of different product markets. On a similar note, some scholars (ie, Schilling, 1998; 1999) have stressed the negative aspects of sponsoring an integrated product architecture and its association with major reworks and redesign in the product development process, risk of architectural lock-in and high ex-ante, sunk costs of development.

In this research study, the degree of product modularity or product integration are conceptualised as ‘static’ design concepts, in other words it relates to the design characteristics of a given product architecture at a given cross-sectional point in time. However, as MacDuffie (2013) notes an analysis of product modularity as a design principle without a discussion of modularisation as a process⁶ is incomplete. Modularisation as a process refers to the various supply- and demand-side processes that usually precede a given state of modularity-in-design; in other words, modularisation-as-process is concerned with the *evolvability* of product designs.

Tracing to the work on technological trajectories (Nelson & Winter, 1977, 1982), technological paradigms (Dosi, 1982), and dominant designs (Abernathy & Utterback, 1978), product architectures can also be conceptualised as technological systems (ie, Schilling, 2000), rarely static and constantly evolving through periods of continuity and discontinuity. The product modularity literature has built upon these foundations to posit that product architectures tend to evolve towards more open and modular designs over time, with modularity becoming a common characteristic of many dominant designs (ie, Murmann & Frenken, 2006; Shibata, Kodama & Yano, 2005).

⁴ Modular product architectures have modularity-in-design such that sub-systems/components can be easily separated/partitioned

⁵ Integrated product architectures have integration-in-design such that sub-systems/components cannot be easily separated/partitioned

⁶ Modularisation-as-process was conceptualised by MacDuffie (2013) as the role of the designers’ knowledge in modularising product architectures. In this research study, I take a broader view than that proposed by MacDuffie. I conceptualise any supply-side or demand-side influence on the design of a product architecture as a ‘modularisation’ or ‘integration’ process.

Furthermore, prior literature that charts the evolution of productive capabilities (ie, Jacobides & Hitt, 2005; Jacobides, Knusden & Augier, 2006), knowledge (ie, Henderson & Clark, 1990), and transaction costs, (ie, Baldwin, 2008), all suggest that products and industries often co-evolve towards more modular or specialised designs over time. This research study complements the extant literature by showing how efficiency considerations and the trajectory of productive capabilities within the focal product market are instrumental in propelling architectures towards increasing product modularity and industry specialisation. More specifically, however, the findings of this research study highlight how the co-evolution of efficiency considerations and productive capabilities originate at the *product component level* across different time periods which results in systemic effects in future time periods.

Despite this large body of work on increasing product modularity and industry specialisation, a few empirical studies have challenged this assumption and have noted that products, firms and industries may co-evolve towards increasing (re)integration in building and construction (Cacciotori & Jacobides, 2005), disk drives (Christensen, Verlinden & Westerman, 2002); bicycles (Fixson & Park, 2008) and Swiss watch manufacturing (Jacobides & Winter, 2005). However, while these studies examine processes that help illuminate under what product market conditions reintegration occurs, there remains a gap in our understanding of how reintegration actually happens – for example, questions such as why and where in a product architecture does product component reintegration take root? Is product component reintegration associated with industry reintegration? These issues remain largely under-explored, and this research study aims to contribute to our understanding of how reintegration takes root in product and industry architectures. As a consequence, the research study aims to contribute to the literature by proposing three stylised hybrid product architecture types.

Many scholars have also hypothesised that product architectures are a kind of roadmap for partitioning task, knowledge and firm boundaries. In other words, the architecture of firms (and hence industries) corresponds to, and is a reflection of, product architecture (ie, Colfer & Baldwin, 2010; Sanchez & Mahoney, 1996), Advocates of ‘mirroring’⁷ may go as far as to suggest that the presence of an architectural correspondence has many important benefits for firms related to the speed, flexibility and efficiency of product development.

The hypothesised correspondence of product component, task, knowledge and firm boundaries has important implications for firms and the evolution of product markets. For example, partitioning product component, task, knowledge and firm boundaries draws an industry towards increasing specialisation and fragmentation which, in turn, has implications for knowledge and productive capability development processes and ultimately for the division of rents and surplus (ie, Jacobides, 2008; Jacobides, *et al.*, 2006). The concept of correspondence, however, is not without critics, with disadvantages cited relating to the decay of product component knowledge, and foregoing higher performance or quality outcomes that may be achieved through integration or via extensive collaboration and co-development with external firms (ie, Brusoni, Prencipe & Pavitt, 2001; Zirpoli & Becker, 2007).

Empirical support for correspondence across numerous studies is significant, but mixed – suggesting the need to focus less upon whether correspondence holds and more upon the contingent conditions under which it holds (Campagnolo & Camuffo, 2012; Furlan, Cabigiosu and Camuffo, 2014). By adopting a retrospective longitudinal research design, the research study aims to examine contingencies that may ‘mist the mirror’ across time. Moreover, the research study shows how architectural correspondence *and* non-correspondence can both be present at the same time. As a consequence, I propose further contingencies that may mist the mirror and bring much-needed clarity and a call for uniformity to the literature.

⁷ The ‘mirroring hypothesis’ is an extension of Conway’s Law (1968), and formalised by Colfer and Baldwin (2010) as the architectural correspondence or ‘mirroring’ between a product’s components, tasks, knowledge and firm boundaries. In the extant literature, various terms are used interchangeably such as mirroring/correspondence/matched versus non-mirroring/misting/non-correspondence/unmatched. In this research study, I will use the term correspondence and non-correspondence to signify ‘mirroring’ between product component, task, knowledge and firm boundaries.

1.2 Rationale

Given the broad and accepted understanding of increasing product modularity and industry specialisation, I was intrigued by the case of the UK personal pensions product market that appears not to fit this canonical pattern, seemingly undergoing numerous co-evolutions of product and industry architecture during the 30-year period of this research study.

The UK personal pensions product market is a unique product market with considerable size and value, with total invested assets of £2,405bn⁸ at the end of 2012. My focus in this research study is the *individual personal pensions* segment which is situated within the overall UK pensions product market as follows:

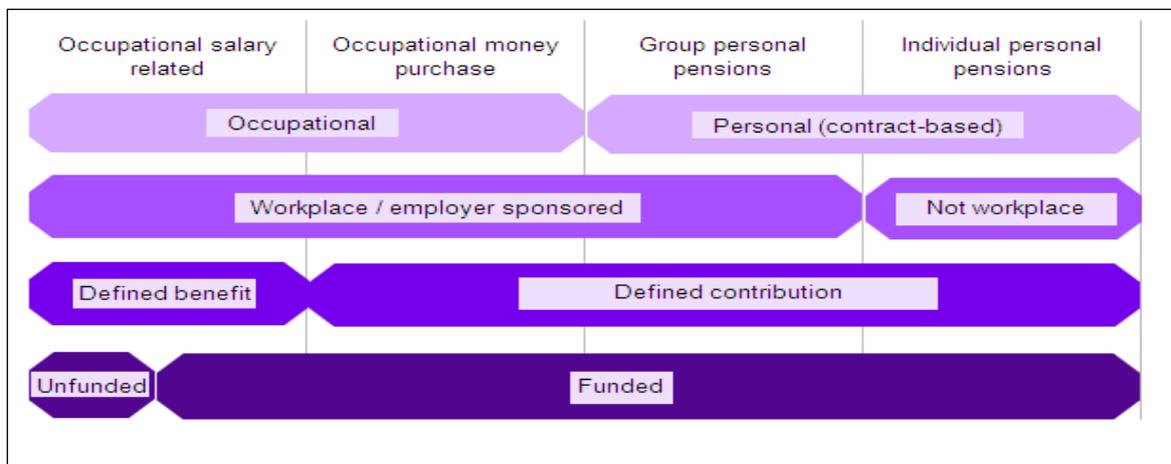


Figure 1: UK pensions landscape⁹

⁸ Source: Association of British Insurers (ABI) report 2012, downloaded <https://www.abi.org.uk/Insurance-and-savings/Industry-data/~media/DBDADF2BB9CD4C8B88419CD2B9E37D5E.ashx>

⁹ Source: ONS, downloaded http://www.ons.gov.uk/ons/dcp171766_269503.pdf

Individual personal pensions¹⁰ have been chosen on conceptual grounds, rather than representativeness (Miles & Huberman, 1994). The choice of product market context was also influenced by my prior employment in the sector as a product architect between 2000 and 2011. However, despite this personal familiarity, my over-arching objective was to select a setting that would illuminate the main concepts and, thus, I selected UK personal pensions precisely because it had undergone numerous periods of both continuity and discontinuity in the period between 1984 and 2014. Therefore, while my approach pays close attention to the specific dynamics of a particular product market context, it is acknowledged that it is likely to undermine its generalizability to other settings (Guba & Lincoln, 1982).

Referring back to the extant literature, few (if any) empirical studies have examined the co-evolution and correspondence of architectures in a setting characterised by non-physical, service-orientated products. The extant modularity literature has often examined manufactured products and this has carried through into examinations of the hypothesized correspondence between architectural layers. For example, as discussed in Fixson (2003, 2005) and Campagnolo & Camuffo (2010), among others, the aircraft engine, motor vehicle and computer hardware and software industries have dominated the architectural co-evolution and correspondence research landscape, and there is a compelling need to broaden out empirical findings to other settings and contexts.

As a non-physical product, an examination of the UK individual pensions product market contributes to scholarly work in at least two important ways: (1) with the exception of Jacobides (2005) examination of the US mortgage banking sector, how architectures co-evolve in a non-physical, knowledge-intensive sector is under-explored, (2) as a regulated product market, how industry regulation influences the direction and speed of architectural co-evolution across time is an important topic that deserves increasing attention. In other words, examining the processes underpinning co-evolution and correspondence in UK personal pensions may help illuminate to what extent such regulated product markets conform to our current understanding of how and why architectures co-evolve.

¹⁰ Hereafter referred to as 'personal pensions'

Furthermore, an understanding of architectural co-evolution and correspondence would be of significant interest to practitioners in the UK personal pensions product market. For example, how managers should decide between architectural choices and organize for innovative activities remains a strategically important concern. This research study aims to highlight that a dynamic view is often required such that decisions made in one time period may be less than optimal in subsequent time periods. By charting how and why the product market has co-evolved in a particular way across time may assist managers in future decision-making.

1.3. Research aims

The aim is to explore how supply-side processes¹¹ may be associated with the co-evolution and correspondence of architectures in the UK personal pensions product market. To achieve this, as I discuss in chapter 3, I constructed a retrospective longitudinal research design¹², as quoted in Fixson & Park (2008:1298-1299), Pettigrew (1990:269) highlights that *“theoretically sound and practically useful research on change should explore the contexts, content, and process of change, together with their interconnections through time”*

I hope, in turn, to make a contribution to the product modularity literature by answering the following questions:

- RQ1: What supply-side processes are associated with the co-evolution of architectures?
- RQ2: Are the design characteristics of product architecture associated with the design characteristics of firm and/or industry architecture?

¹¹ Supply-side processes may relate to processes such as knowledge and productive capability development, transaction costs, property rights, supply-side externalities, competitive dynamics, regulation and so on. This contrasts with demand-side processes such as user/consumer demands which are not the focus of this research study.

¹² I will use the term ‘retrospective longitudinal research’ to signify the methodological approach, as the term has been used in a number of similar studies by authors such as Snelgrove & Havitz (2010)

In this research study, as I explain in chapter 3, I started from a loose set of ideas and frameworks drawn from the extant literature, however the primary focus is to develop new insights and constructs that relate to the research questions via iteration between theory and data (Yin, 2014). I began the data collection phase with an *a priori* product architecture typology (as described in section 2.1.8) synthesized deductively from the extant literature in order to illuminate product architecture evolution in the UK personal pensions product market and to structure and tighten the data collection process (Miles & Huberman, 1994). However, the focus of the data collection and data analysis phases has been to inductively derive new insights into how supply-side processes are associated with evolutions of product architecture, explaining my choice of combining matrix and template analysis as a data analysis method. The matrix I developed emerges from the product architecture typology and this was followed by inductively-orientated template analysis to generate themes within the structuring device of the matrix.

Rather than using secondary sources to construct the product architecture typology, instead I synthesized extant literature from across a number of domains. I then asked respondents participating in the research study to match the stylized product architecture types to the *actual* product architectures that existed in the product market across time, in order to generate a retrospective product timeline and to establish construct validity – a similar idea to creating a personal lifeline in psychology research (Brannen & Nilsen, 2010).

Data was collected via semi-structured interviews with thirty-one key personnel across six different firms. As I discuss in chapter 3, I chose thirty-one respondents based upon practical considerations, but also because this generated over thirty hours of interview data, and allowed for many different explanations and perspectives to be reconciled (Miles & Huberman, 1994). Next, following the semi-structured interviews, I synthesized the thirty-one ‘product timelines’ in order to establish a ‘best fit’ which was subsequently member-checked (Horrocks & King, 2010) ex-ante with the pilot group, and ex-post with an expert panel. The synthesized product timeline then served as the structure for the data analysis phase.

In chapter 5, the findings are positioned within the context of prior work to identify a number of new insights and contributions. Moreover, given the practical and topical nature of the research question to practitioners in the product market, the findings of this research study also offer managerial insights for firms in the product market, as well as public bodies, such as industry lobbying groups and the regulator. The unique characteristics of the UK personal pensions product market may limit its generalisability, however this research study breaks new ground in analysing a regulated product market that has such importance to consumers and the wider UK economy.

1.4 Outline of the thesis

The thesis is structured around 6 chapters as follows:

Chapter 1 introduces the main theoretical basis of this research study and orientates it to the UK personal pensions product market; it then justifies the importance of the research study and its value as a contribution to knowledge and practice.

Chapter 2, Literature Review, provides a review of the extant literature relevant to this research study. It is divided into two sub-sections and begins with a discussion of the product modularity literature, in particular the notions of product architectures as integrated or modular, and as open or closed. Next, I introduce the idea of hybrid product architectures and technological platform architectures and propose a stylised product architecture typology which is subsequently used to structure and tighten the data collection and data analysis phases. I then turn to modularisation-as-process and discuss the principal theoretical perspectives that have dominated the strategic management research landscape. Finally, I turn to architectural correspondence, the so-called ‘mirroring hypothesis’, and review its theoretical underpinnings and mixed empirical evidence

Chapter 3, Research Methodology Design and Methods, places the research study within the ontological and epistemological research traditions, moving onto a discussion of my subtle realist commitments. I then turn to discuss the qualitative research design, based upon thirty-one semi-structured interviews with key personnel, across six firms. Interview data is analysed using a blend of matrix and template analysis to inductively derive themes and concepts, within the structuring device of an *a priori* matrix. The chapter continues with a discussion of the process of data collection and analysis, its ethical considerations and the justification of the data analysis methods employed. Finally, I then discuss quality measures and issues around researcher positionality.

Chapter 4, Data Analysis and Study Findings, presents the findings of the research study by illustrating the templates within the matrix structure. Within the retrospective longitudinal period between 1984 and 2014, five sub-periods are identified as of importance to respondents and the findings are structured as distinct time-periods in order to create a narrative of architectural co-evolution and correspondence across time.

Chapter 5, Discussion, makes use of the findings in order to reflect back on prior work in the field. This process of synthesis with the extant literature aims to develop an understanding of how processes influence architectural co-evolution and correspondence across time. I then draw out the study's main contributions.

Chapter 6, Conclusions and Contribution, summarises the research study's contribution to academic knowledge, management practice and public policy. Furthermore, future research directions are discussed to guide prospective studies. The chapter, and the thesis, concludes with a few personal reflections.

2. Literature Review

This chapter is divided into two sub-sections that correspond to the two over-arching research questions.

Section 2.1 reviews the extant literature in relation to RQ1 ‘what supply-side processes are associated with the co-evolution of architectures?’ Section 2.2 examines the literature in relation to RQ2 ‘are the design characteristics of product architecture associated with the design characteristics of firm and/or industry architecture?’

In section 2.1, I examine the extant literature in relation to product modularity to contrast the design properties of modular and integrated product architectures, and open and closed¹³ product architectures. Then, I formalise a stylised product architecture typology that draws upon existing frameworks (such as Sanchez, 2008; Shibata, *et al.*, 2005), but encompass the notion of a hybrid product architecture, and embeds the idea of evolution. Finally, I turn to ‘modularisation-as-process’, and explore how and why architectures may co-evolve. In terms of the range of ‘processes’, as I noted in chapter 1, my aim is to demark the boundary of this research study by focusing attention on *supply-side processes* that may be associated with architectural co-evolution.

In addition to proposing a stylised product architecture typology, within this sub-section of the chapter, I highlight two significant gaps in the extant literature where I seek to make a contribution. The first significant gap concerns connecting how and why product markets reintegrate with nuanced changes in stylised product architecture type. The extant literature on reintegration in product markets has tended to focus on either endogenous processes at firm level or exogenous processes at the industry level. In this research study, I aim to use empirical data to connect systemic processes that lead to reintegration, at least in this product market, with changes in product architecture in order to propose more nuanced stylised types of hybrid product architecture. Few (if any) prior studies have done so.

¹³ Whereas the integrated-modular continuum is primarily associated with the ease of partitioning and degree of coupling between components, open and closed primarily refers to the concept of ‘ownership’ of the architecture and its interfaces

The second gap concerns our understanding of modularity-as-process, how such processes combine to propel architectures either towards or away from product modularity and industry specialisation. In this research study, I aim to examine which processes combine and whether they combine at the product component level to evolve product components towards being more modular or more integrated.

Section 2.2 of this chapter then proceeds to examine the hypothesised correspondence between product component, task, knowledge and firm boundaries, known as the ‘mirroring hypothesis’ (ie, Colfer, 2007; Colfer & Baldwin, 2010). I review its theoretical foundations before analysing the growing body of mixed empirical evidence, and reviewing recent contingent perspectives that ‘mist the mirror’.

Within this sub-section of the chapter, I highlight two further significant gaps in the extant literature. First, I highlight that the lack of uniformity in how correspondence or non-correspondence is assessed in prior work, one possible explanation of the inconsistent and mixed results. As a consequence, I propose a ‘weak test’ and a ‘strong test’ of correspondence that may provide much-needed uniformity in future scholarly work. Second, I propose further contingencies that may mist the mirror based upon characteristics at both the product architecture and product component levels of analysis.

2.1 Architectural co-evolution

2.1.1 What is product architecture?

The concept of an architecture has been well-established across a number of research communities, such as strategic management, product development, organisation design and technology management (ie, Clark, 1985; Sanchez & Mahoney, 1996; Simon, 1962; Ulrich, 1995) and can be traced back to Herbert Simon’s seminal paper “The Architecture of Complexity” (1962). A product architecture can be conceptualised as a schematic, “*the scheme by which the function of a product is allocated to physical components*” (Ulrich, 1995:419), and encompasses the arrangement of product components and the way they fit or connect together, “*it is an arrangement of functional elements; the mapping from functional elements to physical components; and, the specification of the interfaces among interacting physical components*” (Ulrich, 1995:420). Product architectures, therefore, can be conceptualised to encompass three distinct features: (1) an architecture that is the blueprint for the design and the way in which components¹⁴ are arranged, (2) product components which contribute to the products’ function, and (3) interfaces which document how components fit and connect together (Baldwin & Clark, 2000).

Product architectures have been further stylised as two ideal types, integral¹⁵ or modular (Schilling, 2000; Ulrich, 1995) as shown in Figure 2:



Figure 2: The integrated-modular continuum

¹⁴ In this research study, I shall use the term ‘product component’ to distinguish it from ‘firm component’ or ‘industry component’. Other studies have used the simpler term ‘component’ or ‘module’.

¹⁵ Although Ulrich (1995) uses the term ‘integral’, I will use the term ‘integrated’ throughout this research study

An integrated product architecture is one where product components, interfaces and the nature of their relationship is complex, interdependent and non-standardised, whereas, a modular product architecture has relationships between product components and interfaces that are simple, independent and standardised.

The concept of modularity has been applied to many kinds of complex systems (Ulrich, 1995; Fixson, 2003, comprehensively review modularity across various domains). In the management literature, however, modularity has gained increasing attention as a concept that has the potential to be applied to many different phenomena, such as products, processes, knowledge, firms and industries (ie, Colfer & Baldwin, 2010; Sanchez & Mahoney, 1996). Modularity is a systems concept based upon the notion of partitioning a system into simpler sub-systems or components (Simon, 1962, von Hippel, 1990), invoking the initial ideas of information-hiding and parallelism (Parnas, 1972) and decomposition (Alexander, 1964; Simon, 1962), focusing on reducing or removing interdependence. Product architectures are often hierarchically-nested¹⁶ (Alexander, 1964; Clark, 1985; Schilling, 2000), and differ in the degree to which they have been partitioned into ‘loosely-coupled’ or ‘tightly-coupled’ product components (Orton & Weick, 1990). In other words, tightly-coupled product components exhibit a high degree of interdependence¹⁷ with each other and design changes to one product component may have knock-on design consequences for other product components (for example design changes to product component A have significant knock on design consequences for product component B), whereas loosely-coupled product components exhibit a high degree of independence from each other, and design changes to one product component often have no (or few) design consequences for other product components (for example, design changes to product component A have few or no design consequences for product component B).

¹⁶ In this context, hierarchically-nested refers to sub-components being part of larger components which, in turn, form part of an architecture. A product can, therefore, have many different levels within its systems hierarchy

¹⁷ For the purposes of simplicity, in this research study interdependence and independence have the same meaning as tight-coupling and loose-coupling, and refer to the components’ ease of separation/partitioning

In the extant literature, terms such as tight-coupling, integration and interdependence are often used inter-changeably, as are loose-coupling, modular and independence. I will also assume the terms have similar meaning and use the terms interchangeably, except where I note otherwise.

The primary focus of the integrated-modular continuum is the relative interdependence or independence (or tight-coupling or loose-coupling) of product components within a given product architecture. However, product architectures also incorporate the concepts of ownership and ‘interfaces’ (how product components connect together) and, as a result, a number of scholars have extended Ulrich’s continuum to encompass the degree to which a product architecture is either proprietary or non-proprietary, and the extent to which its interface specifications are either non-existent, specialised (firm-specific)¹⁸ or industry-wide¹⁹ (ie, Fine, Golany & Naseraldin, 2005; Sanchez, 2008; Schilling, 2000). Taken together the integrated-modular and open-closed continuums can be represented in a 2x2 matrix as follows:

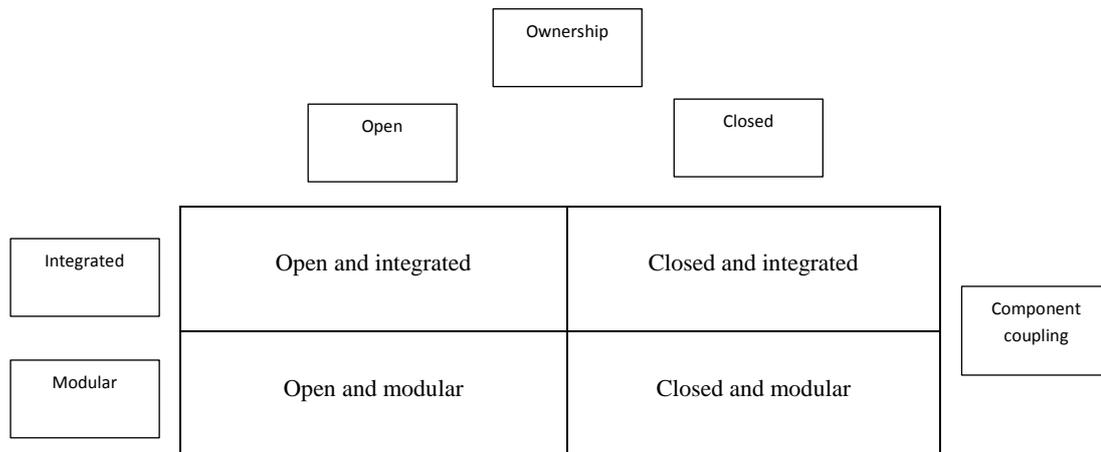


Figure 3: Component coupling and ownership properties of stylised product architecture types²⁰

¹⁸ Firm-specific interfaces are interfaces between product components that are defined, and used only within firm boundaries (ie, Fine, Golany, & Naseraldin, 2005; Schilling, 2000)

¹⁹ Industry standards are standards which have permeated across firm boundaries and have been used by many firms as an industry-wide basis to connect product components. Therefore, they may emerge over time or be imposed by a third-party such as a standards-setting organisation (such as the DVD forum) or by Government

²⁰ Based upon typologies in Sanchez (2008) and Shibata, Kodama and Yano (2005)

As I explain in section 2.1.7, a 2x2 representation as illustrated in figure 3 has a number of limitations, and throughout this section of the chapter, I will build upon this prior work to propose a revised product architecture typology. The idea that product architectures can be characterised as either modular or integrated, or open or closed, is now discussed in sections 2.1.2-2.1.5. Following this, in section 2.1.6, I introduce the idea of non-perfect or hybrid product architectures.

2.1.2 Integrated product architectures

Integrated product architectures usually incorporate product components that are tightly-coupled and interdependent with each other, and connect together via interfaces that are closed, often idiosyncratic, (Sanchez, 2008) or even non-existent (Mikkola, 2003, 2006). In such integrated designs, interdependent product components work tightly together to achieve a products' function, in other words, there is a 'many-to-one' relationship between product components and product functions (Ulrich, 1995).

Integrated product architectures are often created to serve a single use or market purpose (Sanchez, 2008), statically-optimised along some dimension (Sanchez & Mahoney, 2013) such as maximum performance or lowest cost (Sanchez, 2008). Moreover, Schilling (2000:316) described such integrated designs as exhibiting high levels of 'synergistic specificity' - the constituent component parts of the product architecture are highly 'specific' to each other. For example, desktop, inkjet printers are often optimised for the lowest cost and at the broadest level can be conceptualised as consisting of two primary product components, (1) the ink cartridge and (2) the printer outer shell. The ink cartridge is a replacement item and so is linked to a stream of rents. Hewlett Packard (HP) and Canon, for example, integrate the print-head with the ink cartridge into a single component, protected by patents, in order to minimise the risks of external firms reverse-engineering the ink cartridge as a stand-alone product component. As such, tight integration allows HP and Canon to optimise the product architecture and, at the same time, to appropriate rents from the replacement ink cartridge.

Integrated product architectures are usually very difficult to re-architect to new uses without re-engineering the entire product architecture or at least a significant proportion of it (Sanchez & Mahoney, 1996, 2013), because design changes to one product component often result in knock-on, and possibly unforeseen, design changes to other product components (Sanchez & Mahoney, 1996) which means that isolating product component level innovation is hindered, often requiring more difficult, expensive and time-consuming architectural innovation (Henderson & Clark, 1990) to untangle the inter-connectedness and web of complexity. As such, the design characteristics of integrated product architectures can act “...as a strong force against the system shifting to a more modular design” (Schilling, 2000:316) due to ‘technical imbalances’ in integrated product designs (Murmman & Frenken, 2006).

As discussed, integrated product architectures are often used by sponsoring firms to maximise the appropriation of rents (Schilling, 1999, 2000). For example, the usual logic is that in many product markets with ‘new to the world’ products, firms often develop “*product architectures that are idiosyncratic to the firm and that feature customised and highly interdependent components*” (Argyres & Bigelow, 2010:853). Because of the ex-ante sunk cost investments required to design and develop integrated product architectures, appropriation from innovation (Teece, 1986) is a key concern for firms, in order to recoup the initial investment in the products’ design and development (Sanchez, 2008). However, if a firm can establish and protect its own, idiosyncratic and integrated product design as a ‘dominant design’, a firm “...may be able to earn near-monopoly rents” and be in “*a good position to shape the evolution of the industry*” (Schilling, 1999:266).

The sponsorship of an integrated product architecture is not without risks, however, as such products are often non-compatible with complementary product components and are often more expensive and time-consuming to develop and improve than more open or modular product designs (Schilling, 1999). Moreover, Schilling (1999:269) contends that integrated product architectures often have a “*high risk of rejection under conditions of strong network externalities*” as a firm would need to generate its own firm-specific externalities through branding and marketing in order to generate a sufficient installed user base around its product architecture.

For example, early versions of hand-held devices were largely integrated. The Palm hand-held device was designed and manufactured by Palm who sold a bundled package of hardware, operating system and applications, but it was later rejected by the market under conditions of strong network externality effects. In contrast, however, Fixson and Park (2008) argue that Shimano, in the bicycle industry, with weaker network externalities, were able to significantly invest in marketing and branding to over-turn modular bicycle designs. In sum, Sanchez (2008:337) contends that the salient characteristics of an integrated product architecture is: *“it’s (intended) ability to deliver optimised performance or cost according to a fixed set of performance and cost criteria; its use of integrated component designs and idiosyncratic interfaces that are intended to achieve optimised performance or cost; and, the second characteristic thereby makes it time-consuming and costly to accommodate changes in the performance or cost parameters to be delivered by the architecture”*.

The product development process that underpins an integrated product architecture often follows a traditional path with sequential staging. This is often the case because technological changes to integrated product designs are often localised in the sense that *“changes are tried in one part, this defines further problems in other parts, which in turn may define problems in other parts”* (Murmman & Frenken, 2006:937). In a sequential development approach, *“after designing the product concept, design and development tasks are sequenced so that technology and component development tasks with the greatest need for new knowledge and with the greatest impact on other component design and development tasks are undertaken first”* (Sanchez & Mahoney, 1996:68). It then specifies the next level of product components and repeats until all product components are specified.

Sequential development processes are often associated with breakdowns, delays and recursive information flows (Sanchez & Mahoney, 1996) and are likely to ‘lose information’ as development proceeds from one stage to the next, such that *“technical interdependencies can be ‘designed into’ the architecture”* (Sanchez & Mahoney, 1996:69). On the other hand, however, integrated product architectures often enhance knowledge-sharing and learning in a product development team as team members rely on each other’s expertise (Mikkola, 2003), as their knowledge becomes more and more specific to the product being designed (Schilling, 2000).

2.1.3 Modular product architectures

Modular products architectures are a feature common to many product markets, such as bicycles (Galvin & Morkel, 2001); semi-conductors (Funk, 2008), air-conditioning systems (Furlan, Cabigiosu & Camuffo, 2014) and stereo systems (Langlois & Robertson, 1992). The design characteristic that lies at the heart of modular product architectures is greater interdependence within product components than across different product components (Ulrich, 1995), known as ‘loose-coupling’ (Orton & Weick, 1990), as shown in figure 4.

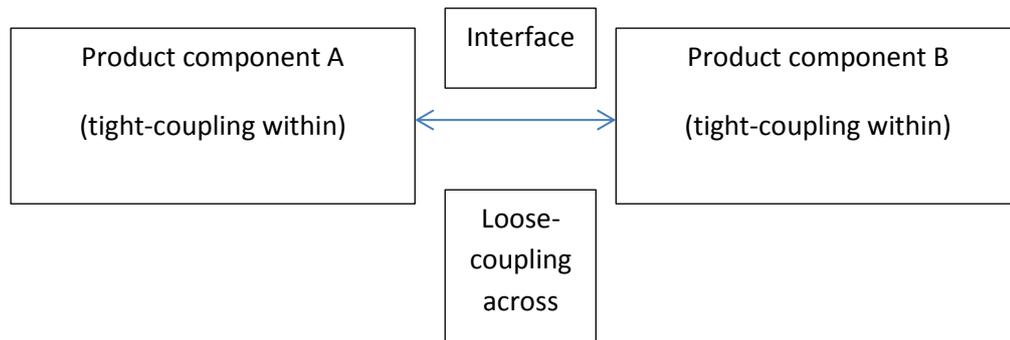


Figure 4: Example of loose-coupling of two modular product components

In perfect form, loose-coupling facilitates a one-to-one mapping between product functions and product components (Ulrich, 1995), and design changes to one product component require little or no modification to other product components, so long as there is adherence to a specified and stable interface. This means that modular designs also permit product components to be designed and produced independently by separate individuals, teams, divisions or firms (Sanchez, 2008), and in parallel (Parnas, 1972), which results in a potential reduction in development cycle time and an increase in speed to market (Sanchez & Collins, 2001). For example, assume Figure 4 refers to a reusable razor system, comprising two loosely-coupled product components: component A is the handle and Component B is the detachable and replaceable blade. Component A and B can each be designed and developed in parallel, potentially by separate teams or firms, so long as each team or firm proceeds with reference to the interface specification. Furthermore, Component B can be subject to modular or incremental innovation without reference to Component A.

In a modular product architecture, there exists the potential to (1) standardise product components and/or (2) the interface specifications. Product component standardisation often results in re-use across multiple products so as to benefit from economies of substitution (Garud & Kumaraswamy, 1995; Ulrich, 1995). For example, many of the Apple software components are re-used across multiple devices. However, modular product components do not need to be standardised. Rather, the architectural blueprint or ‘design rules’ (Clark, 1985) dictates which parameters are ‘hidden’ inside the inner workings of a product component and which parameters are ‘visible’ to other components (Parnas, 1972). According to Baldwin & Clark (2000), what is often hidden and encapsulated within a product component is the knowledge and intellectual property (IP) relating to the design parameters of the product component, and what is often ‘visible’ are the design rules of the architecture and the interface specifications that “*connect, interact or exchange resources in some way, by adhering to a standard interface*” (Schilling, 2000:318). Therefore, according to Langlois (2012:5), product modularity is advocated as a useful design principle to encapsulate high levels of complexity or uniqueness inside a product component, especially when a system becomes unwieldy and interdependencies so far-reaching that integrated design efforts become almost impossible (Parnas, 1972).

Interface standardisation is perhaps the key design characteristic of modular product architectures (Sanchez, 2008; Sanchez & Mahoney, 1996), as it keeps the interfaces between product components constant or frozen (for a time). It is a process that makes assets non-specific (Schilling, 2000), and can take two generic forms: (1) interface specification may take place within the boundaries of a single firm who create firm-specific or ‘specialised interfaces’ (Fine, *et al.*, 2005; Schilling, 2000), or (2) a specialised interface specification may permeate across firm boundaries, perhaps eventually to become an industry-standard. According to Tee (2011) and Christensen, *et al.*, (2002), interfaces are often codified initially within firm boundaries, but they then often permeate across firm boundaries and become adopted by external suppliers.

An example that typifies this progression from specialised to industry standards is the classic story of IBM's modularisation of its computers in the 1970s (it also illustrates the risks of losing control of proprietary interfaces). IBM began to partition its product architecture into modular product components and created firm-specific interfaces to connect components together. Many of these were shared with external suppliers who then reverse-engineered and sold stand-alone product components to fit with the IBM architecture. Eventually, IBM lost control of the interfaces, and industry standards developed across much of the computer industry which opened the product market up to new entrants, imitation and intense competition.

Despite the risks, standardised interfaces may often increase product component variety because component upgrades can be leveraged by easier substitution (Sanchez, 1995), so that firms benefit from economies of substitution (Garud & Kumaraswamy, 1995, Ulrich, 1995). In other words, modular design characteristics permit a far easier mixing and matching of product components to give a potentially large number of product variations (Sanchez & Mahoney, 1996; 2013; Schilling, 2000; Ulrich, 1995), or product families (Argyres & Bigelow, 2010) which may be a source of strategic advantage (Sanchez, 1995), or option value (Baldwin & Clark, 2000), by enabling a firm to respond more quickly to evolving and heterogeneous markets. Furthermore, product modularity may offer a route for firms to reduce the costs of differentiation based on product component design (Sanchez & Mahoney, 1996), or a firm may choose to insource product components that offer a source of differentiation, and outsource non-core product components (Pil & Cohen, 2006). In other words, modular product architectures are dynamically-optimised, rather than statically-optimised (Sanchez & Mahoney, 2013).

The presence of industry standards also makes it easier for firms to access a wide range of complementary product components²¹ to plug and play (Sanchez, 2008) into the product architecture. For example, laptops or PCs from different manufacturers may operate in slightly different ways, however, they are all designed to permit standardised USBs to plug into the product and be decoded and read.

²¹ Complementary product components are generally taken to be non-core product components that can connect to a product in order to enhance its benefit to consumers. For example, Apple invites App developers to contribute Apps to its range of devices.

Complementary product components may feasibly be developed within firm boundaries, however, interface standardisation and a plug and play design often extends product development across two or more firm boundaries and is particularly important in product markets where competitive advantage may depend upon the number or range of complementary products or services (Galvin & Rice, 2008; Sanchez, 2008; Schilling, 1999, 2000).

Product development teams, therefore, may find competitive benefits in sharing interface specifications with specialised providers of complementary product components (Sanchez & Collins, 2001). Interface standards, however, are not fixed and stable in perpetuity. They can be a contested domain, and often are supplanted by rival standards (ie, Garud, Jain & Kumaraswamy, 2002), or they may occasionally revert back to a private or proprietary form if strategic or performance considerations emerge (Fixson & Park, 2008). Ultimately, scholars such as West (2003, 2007) and Tee (2011) argue that strategic factors will play a moderating role in exactly how and if interface specifications are shared with external firms.

Modular product architectures not only provide firms with significant network externalities – or ‘modular externalities’ (Sanchez, 2008) – but network externalities may be present on both the supply- and demand- side (Sanchez, 2008; Schilling, 2000). On the supply side, it may help mitigate both ex-ante and ex-post transaction hazards (Schilling, 1999, 2000), as the convergence of an intermediate market around stable, standardised interfaces can reduce asset specificity (Argyres & Bigelow, 2010), small numbers bargaining problems and the risk of ex-post transaction costs (Williamson, 1985, 1991). On the demand side, network externalities may also promote the convergence of a significant installed user base around a particular product design as a result of the wide availability of complementary product components (Galvin & Rice, 2008; Schilling, 1999) or a reduction in the risks of technological obsolescence that may be perceived by consumers (Schilling, 1999).

Moreover, the presence of potentially self-reinforcing network externalities may offer strategic benefits to firms who decide to ‘open-up’ and create incentives for cooperative product development activities based upon a common approach to interface specification. For example, Gawer (2014) cites the example of the Nintendo Wii to highlight how self-reinforcing network effects may operate: the Nintendo Wii console has a large installed user base which increases the incentives for external firms to participate in designing Wii-compatible complementary products which, in turn, attracts more end-users to the Wii console.

The product development process that underpins a modular product design requires fully specifying product components and interfaces ex-ante before development begins (Sanchez & Mahoney, 1996). This means that a firm must possess knowledge about the architecture, its product components and how they connect together. However, as product development proceeds, and the blue-print of the architecture stabilises, innovation tends to turn to the product component level (Galvin & Morkel, 2001; Henderson & Clark, 1990) and subsequent incremental and modular innovation efforts may be focused on new market segments or heterogeneous consumer needs.

Modular product architectures are often associated with partitioning product component, task, knowledge and firm boundaries (Colfer, 2007; Colfer & Baldwin, 2010). For example, Sanchez (2003:382) writes that *“the standardising of component interfaces based upon the firms current architectural knowledge largely decouples architectural knowledge-based processes from the component-level knowledge used to develop specific component design during product development”* such that manufacturers/assemblers of modular products can specialise in architectural knowledge and outsource product component design and development across firm boundaries (Sako, 2003). In other words, in product markets where systems integration may be important, according to Hobday, Davies & Prencipe (2005:1128), *“the lead firm moves away from an in-depth control over component design and manufacture to the systems integration knowledge and skills needed to integrate the modules produced by others in the supply chain”*.

According to Sanchez and Mahoney (1996:72), the partitioning of architectural and product component knowledge can enhance a firm's ability to focus on architectural learning and capability development, rather than on lower-level product component knowledge, and that this may ultimately help a firm overcome the architectural challenges for incumbent firms as posed by Henderson and Clark (1990). Once architectural and product component level knowledge is partitioned, it can facilitate the division of labour, such that product components can be *“designed and developed concurrently and autonomously by geographically-dispersed, loosely-coupled development groups* (Sanchez & Mahoney, 1996:70). In other words, modular product development activities can proceed independently and concurrently across a range of individuals, teams, groups, divisions, or even firms.

2.1.4 Closed architectures

Product architectures may be conceptualised as either integrated or modular, or along the dimension of being either open or closed (Sanchez, 2008; Sanchez, Galvin & Bach, 2013; Shibata, *et al.*, 2005). A closed product architecture is one that is not able to be used by other firms in the industry, it is proprietary and a firm may choose to hide its intellectual property through formal mechanisms such as patents, trademarks, and copyright, or informal mechanisms such as secrecy, encryption, or complexity, or a combination of the two (Sanchez, *et al.*, 2013). Boudreau (2010:1851) makes a similar observation when highlighting that *“closed technologies are wholly-owned proprietary, vertically integrated and controlled by a single party”*. Sanchez (2008:341) also suggests, however, that *“a firm may have such advanced design or production knowledge about a component or its interfaces that other firms are simply unable to create equivalent components or use the firm's interface specification in architecting their own components or products”*. All of the reasons may make it near-impossible for other firms in a product market to gain knowledge about or access to a closed product architecture.

Sanchez, *et al.*, (2013:2) coined the term ‘architectural specificity’ to describe closed systems, suggesting that it is a feature of product architectures that have *“been strategically partitioned into kinds of components that are functionally different from those used in other firms’ architectures and/or that have adopted interface specifications for their components that are not compatible with the interface specifications used in other firms’ architectures”*. As a consequence, firms who choose to sponsor closed product architectures may be unable to plug and play complementary product components from other external firms, owing to the specificity and idiosyncrasy of their designs. Sanchez (2008:342) highlights the main defining characteristics of closed product architectures as *“proprietary, closed systems; they have significant design integration of optimised components; and, they are only understood by the specific firm and requires significant technical knowledge exchange and coordination”*.

Not all closed product architectures are integrated, however. For example, Takeishi (2002) and Takeishi and Fujimoto (2003) refer to the motor vehicle industry as a form of ‘closed modularisation’, Langlois (2002) distinguishes between forms of internal modularity and external modularity, and Chesbrough (2003) between internal modularity and market modularity. In other words, (1) internal modularity refers to product modularity within firm boundaries and (2) external or market modularity refers to product modularity across two or more firm boundaries. In his paper, Langlois (2002) also recounts the example of IBMs computer systems in the 1970s as a form of internal modularity, as IBM sought to keep its modular interfaces proprietary and to prevent external firms from supplying modular product components (it subsequently failed as has been well-documented).

The idea of a closed and modular product architecture has been formalised in the stylised frameworks offered by Sanchez (2008) and Shibata, *et al.*, (2005). A closed and modular product architecture should theoretically offer firms the potential to obtain rents from *“configuring greater product variety, more rapid upgrading of components, or lower costs for common components used across product variations”* (Sanchez, *et al.*, 2013:30).

Firms with closed and modular designs, therefore, may be able to develop modular product components and firm-specific interface specifications via their own firm-specific knowledge and capabilities, but it is also a decision to forego network externalities and the potential for technical advance that may emerge with a diversity of contributions from external firms (Woodward & West, 2009). Baldwin and Clark (2000) also highlight that a closed and modular product architecture is also a decision to forego the option value of sourcing the best quality or lowest cost product components.

2.1.5. Open architectures

An open architecture is a type of product architecture with high levels of commonality (Sanchez, *et al.*, 2013) and interface specifications that are standardised and permeate across the product market (Shibata, *et al.*, 2005), unencumbered by intellectual property rights or other means of secrecy or protection. Boudreau (2010:1851) highlights that open technologies are “...neither wholly owned, nor controlled by a single party, thus accessible to all” and firms interested in developing complementary product components can often ‘plug and play’ into the product architecture (Sanchez, 2008; Sanchez, *et al.*, 2013). Open product architectures are often used by more than one, and often many, firms. Industry standards are an important design feature of open product architectures and may be associated with core technologies, specified by government or industry groups such as the DVD standard developed by the DVD Forum, or with operational processes such as ISO standards (Schilling, 1999, 2000). Open innovation projects are often viewed as a form of ‘extreme’ open architecture (ie, Chesbrough, 2003; West, 2003).

Sanchez (2008:342) argues that “*because open-system architectures allow firms to access an architecture, they may create gains from trade, as well as significant positive network externalities, both for firms that participate in a product market based on an open architecture and for users or consumers of products and services leveraged from open-system architectures*”. Open architectures, therefore, are often associated with the presence of significant demand- and supply-side network externalities and gains from trade and/or specialisation and can be found in many product markets where interoperability is essential, including personal computers and consumer electronics (Langlois & Robertson, 1992), and telecommunications (Galvin & Rice, 2008).

Perhaps the classic example of an open architecture is the Linux operating system. On the demand-side, any user can use Linux, and on the supply-side, any firm can assemble or bundle Linux with other hardware or software, and any firm can design and develop a Linux-compatible complementary product, or contribute to the improvement of the operating system, subject to the open source standards. Sanchez (2008:342) summarises open architectures as: *“open systems, with strategic partitioning and interfaces known and understood by many firms and available for use by any firm; they have one to one mapping into technically separate individual components, and allow component substitution; the interfaces between product and process have been technically frozen to create a technically stable environment; and, they are decoupled and can be developed by technically separate process components and firms”*

As shown in Figure 3, the integrated to modular and open to closed characteristics of product architectures would seem to indicate a stylised 2x2 product architecture typology comprised of closed and integrated, closed and modular, open and integrated and open and modular product architecture types. However, while such a stylised product architecture framework as formalised by Sanchez (2008) and Shibata, *et al.*, (2005) may extend the original integrated to modular continuum proposed by Ulrich (1995), these contributions suffer from at least three limitations, (1) it is highly unlikely that an open and integrated product architecture exists in practice (ie, Shibata, *et al.*, 2005:15) because *“there are virtually no products for which the mapping relationship is complex and for which standard interfaces have been established...accordingly we may assume that as a rule, open architectures are [always] modular architectures”*, (2) few product architectures are likely to exhibit the perfect characteristics of an ideal or stylised type (Schilling, 2000; Ulrich, 1995), and in practice many product architectures are often comprised of a blend of integrated and modular product components, and firm-specific or industry-wide interfaces (a few scholars have previously hinted at the notion of a semi-open product architecture, ie, Boudreau, 2010; 2012; West, 2003), (3) the product architecture framework is ‘static’ and fails to capture the evolutionary nature of product architecture design across time.

I now turn in section 2.1.6 to the notion of non-perfect product architectures – what I shall call hybrid product architectures in line with West (2003).

2.1.6 Hybrid architectures

Despite the intuitive appeal of the open and closed, and modular and integral dimensions, few product architecture types are perfectly open or closed, or perfectly modular or integral (Schilling, 2000; Ulrich, 1995). In fact, many product architectures (at the product component level) are often a blend of both modular and integrated product components, as well as a blend of both specialised and industry standards. The notion of a hybrid product architecture seems entirely logical because the partitioning of a product architecture often has limits so that *“some interdependencies remain across modules, yet many fewer than the interdependencies contained within module boundaries”* (MacDuffie, 2013:11) and *“few products, services or systems would be composed exclusively of modular or interdependent interfaces – suggesting that architectures that are entirely modular or entirely interdependent would be the rare extremes at opposite ends of a spectrum”* (Christensen, et al., 2002:959).

‘Non-perfect’ product architectures have been described as ‘semi-open’ (Boudreau, 2010) or ‘hybrid’ (West, 2003) because products, as complex technological systems, are made up of multiple product components, and can often be ‘opened up’ one product component at a time (Boudreau, 2010). West (2003:20) identifies two types of hybrid strategy: (1) opening parts of the architecture by waiving control of the commodity layer, while retaining full control of other layers that may offer a source of potential differentiation along some lines, or (2) disclosing technology under restrictions so it cannot be easily copied by competitors. West (2003:20) hypothesises that the former strategy is likely to be advantageous to drive adoption and facilitate interoperability. However, as West suggests, without some approach to innovation, differentiation or some form of lock-in, incumbent firms will find it very difficult to appropriate rents unless there are advantages to be gained through marketing, customer service, or leveraging brand names. Boudreau (2010:1852) also suggests that product architectures can be often characterised as exhibiting a ‘degree of openness’, which may vary along a number of dimensions; for example, its *“...treatment of property rights, contracts, and rules, as well as their procedural characteristics”* and Gawer and Cusumano (2014:420) argue it may vary according to its *“level of access to information on interfaces to link to the platform or utilise its capabilities, the types of rules governing use of the platform, or cost of access (as in patent or licensing fees)”*.

Opening up an architecture to allow external firms to participate in its development is often characterised as a crucial decision because “...*opening has the potential to build momentum behind a technology, but could leave its creator with little control or ability to appropriate value*” (Boudreau, 2010:1849), the so-called ‘paradox of openness’ (Yacoub, 2015). Boudreau (2010:1852) posits that the paradox of openness can be conceptualised as a ‘balance of power’ between the product architecture owner and external contracting parties.

Boudreau (2010:1853) suggests that giving up control means sharing a firms’ architectural IP so that external firms can ensure the interoperability of their activities. By doing so, firms benefit from an increase in the diversity of complementary product components, providing external firms are motivated or willing to contribute. Again, a trade-off identified; simply granting access to a significantly large number of external firms, and fostering intense competition among firms in the intermediate market can reduce incentives for those firms, whereas opening up a product architecture to differentiated contributions may increase innovative efforts. On the reverse side of the coin, concentrating control in the hands of one firm may lead it to appropriate rents by squeezing suppliers’ ex-post. As a result, external firms may lack the incentive to invest in innovative activities. The key success factor, according to Gawer and Cusumano, (2014:421) is allowing “...*interfaces to be sufficiently open to allow outside firms to plug in components as well as innovate on these compliments and make money from their investments*”.

The paradox of openness has been conceptualised in a number of similar ways in the extant literature. For example, Schilling (1999, 2000) conceptualises the paradox as ‘diffusion versus protection’ and points out “*this creates a dilemma for the firm about whether to protect or diffuse its technology: although a firm might wish to protect its proprietary technologies in order to appropriate rents, product systems based on open standards might more rapidly accumulate an installed base and be compatible with a wider range of complementary goods*” (Schilling, 2000:329).

Schilling (1999:269) suggests that in many cases diffusion may be a better strategy than protection, however warns that open systems are often “*quickly commoditised, and may provide little appropriability of rents to their developers*”. Protection, on the other hand, is considered a suitable strategy “*only if there are few competitors, an existing range of complementary goods, and/or the technology has a great margin of improvement over other options*”. Schilling (1999) considers the adoption of a modular product architecture as a way to trade-off ‘diffusion versus protection’. In her view, “*modular product systems can utilise proprietary technology within components of the system, but use a standards-based interface to interact with other components or systems*”. In other words, by hiding IP within a product component boundary, subject to a standards-based interface, firms can benefit from both interoperability and rent appropriation. West (2003) conceptualises the trade-off in the same way as Schilling, but uses different terms - adoption versus appropriation. West reminds us that when a firm pursues a fully open product strategy it is likely to reduce rent appropriation as it opens up competition and lowers entry barriers. On the other hand, opening up a product architecture may encourage user adoption as it reduces consumers’ fears of being locked in to a single technology manufactured by a single firm (Katz & Shapiro, 1994). Several studies in fact have concluded that product designers should strive for the development of product architectures with near-decomposability but with an intermediate degree of product component interdependence in order to balance these strategic trade-offs (Boudreau, 2010; Ethiraj & Levinthal, 2004; Fleming & Sorenson, 2001; Rivkin & Siggelkow, 2006).

Hybrid product architectures may, therefore, be conceptualised as a product architecture that fosters intermediate integrated-modular and open-closed characteristics as a mechanism for firms to open up access to complementary product components, but without giving up full control (Boudreau, 2010). Boudreau (2010:1850) cites the example of Apple in its development of the iPhone, where Apple tightly-controlled the operating system, but allows thousands of external firms to develop software applications. This is, of course, a fine balancing act, the trick being to determine which product components to open up and which to retain proprietary control. These strategic decisions can lead to product architectures being further opened or closed as the trade-offs are evaluated and re-evaluated by firms across time. For example, Apple initially bundled Google Maps as part of its iPhone, but then replaced it with its own Apple Maps application as part of iPhone 5. Similarly, Netscape began life as a provider of a complementary product to Microsoft, who later enveloped the browser component within Internet Explorer.

I now turn to propose a stylised product architecture typology that aims to synthesise the various strands of the extant literature discussed, and builds upon prior conceptual and theoretical contributions. I propose a stylised product architecture typology consisting of closed and integrated, closed and modular, hybrid²² and open and modular product architectures in order to encompass the emerging literature on non-perfect product architectures. The product architecture typology I propose responds to the advice of Campagnolo and Camuffo (2010:260) in (1) specifying the unit of analysis, (2) measured along a continuum from integrated to modular, and (3) recognises that degrees of product modularity may change over time.

2.1.7. Stylised product architecture design types

Diagrammatically, I propose the following product architecture typology:

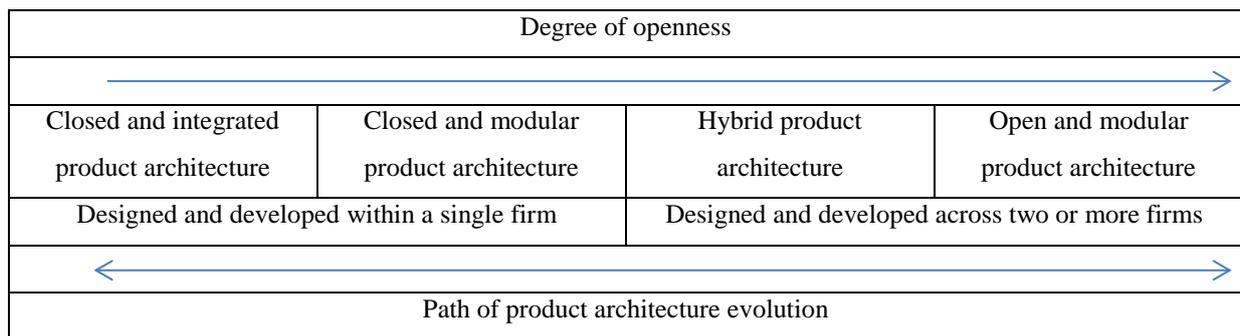


Figure 5: Stylised product architecture typology

²² It is worth noting here that in the discussion chapter 5, I develop 3 types of stylised hybrid product architectures, under conditions of reintegration, that further contribute to the extant literature

In figure 5, I illustrate that the degree to which a product architecture is open increases from left to right. In other words, product architectures are likely to evolve from being closed and integrated to closed and modular within the boundary of a single firm. Next, closed and modular product architectures may evolve into hybrid product architectures as product component design extends across firm boundaries, and, finally, towards an open and modular product architecture as further product components are outsourced. In addition, the evolution of product architectures is assumed to be bi-directional, such that the reverse path towards (re)integration is feasible as product components are re-absorbed within firm boundaries.

I define a *closed and integrated product architecture* as a proprietary architecture. It is designed and used by one controlling firm and a high number of product components are tightly-coupled and integrated with each other. A low number of product components may be loosely-coupled and modular. The product architecture may be protected by either formal or informal mechanisms, or both, and interface standards are either idiosyncratic to the controlling firm or non-existent. Examples from the UK personal pensions sector include many final salary occupational pension schemes. In such schemes, product components are tightly-coupled and optimised to achieve a specific retirement outcome, usually based on a percentage of final or average salary. To achieve the outcome, member contributions are tightly-coupled to the outcome which, in turn, is tightly-coupled to the underlying investment components. Interface standards between product components are idiosyncratic or non-existent. Occupational pension architectures can therefore be described as closed and integrated.

A *closed and modular product architecture* is also a proprietary architecture, protected by either formal or informal mechanisms, or both. A high number of product components are loosely-coupled and modular. A low number of product components are tightly-coupled and integrated. A high number of interfaces have been specified by the controlling firm. A low number of interfaces have been adopted as the basis of market exchange with a low number of external suppliers.

Many UK personal pensions are architected as closed and modular. Stakeholder personal pensions, for example, are designed on proprietary mainframe IT systems, however a number of product components are a form of closed modularity. For example, many stakeholder pensions have a range of investment funds that can plug and play into the architecture, so long as they are proprietary and meet the specialised interface standards. In the wider UK financial services sector, closed and modular product designs can be found across many sectors such as retail banking and mortgages.

A *hybrid product architecture* has elements that are proprietary and controlled by a single firm, and other elements that are controlled by two or more firms. It has an intermediate *mix* of tightly-coupled and integrated product components and loosely-coupled and modular product components, and an intermediate mix of firm-specific and industry-wide standards. In the UK personal pensions sector, such designs are common in unit-linked personal pensions. These hybrid designs are proprietary and incorporate a wide range of investment funds that can plug and play from a wide range of external firms. Interfaces, however, are specialised but each firm shares their own specifications with external firms as a basis for market exchange and co-development practices. Some industry standards are imposed by the regulator or negotiated via industry standard-setting organisations (SSOs).

An *open and modular product architecture* does not have one controlling firm, and is typically a non-proprietary architecture, used by more than one, or many, firms, unencumbered by formal or informal protective mechanisms. A high number of product components are loosely-coupled and modular, interacting and connected through a high number of industry standards. A low number of tightly-coupled and integrated product components or firm-specific interfaces may be present. Significantly open and modular product architectures can be found in the investments and retirement sector, encompassing a range of plug and play product components. For example, self-invested personal pensions (SIPPs) are often licensed from a small number of core systems architects, designed and administered on licensed platforms, with an almost limitless range of investment components from thousands of external firms. Industry standards exist for many product or process component interfaces.

I now turn to discuss prior empirical studies that have examined architectural co-evolution, followed by the supply-side processes that may be associated with such architectural co-evolution.

2.1.8 Paths of architectural co-evolution

The evolution of product architectures can be traced back to Simon's (1962:468) seminal work on hierarchy and systems. As technological systems, product architectures are rarely static in nature and may often evolve over time (Christensen, *et al.*, 2002; Fine, 1998; Fixson & Park, 2008; Schilling, 2000; Shibata, *et al.*, 2005) "*in response to changes in their context, or to changes in their underlying components in the pursuit of better fitness*" (Schilling, 2000:314-5).

Product architectures are often assumed to follow an evolutionary path towards increasing modularity. In other words, product architectures may often be designed initially as closed and integrated, and progress through various iterations of design changes until they exhibit more or less open and modular characteristics (ie, Sanchez, 2008; Shibata, *et al.*, 2005). Empirically, this phenomenon has been examined in a number of product market settings, such as stereos (Langlois & Robertson, 1992), numeric controllers (Shibata, *et al.*, 2005) motor vehicles (Argyres & Bigelow, 2010) computers (Baldwin & Clark, 2000), mortgage banking (Jacobides, 2005) software (MacCormack, *et al.*, 2012) semi-conductors (Funk, 2008) and bicycles (Galvin & Morkel, 2001). Shibata *et al.*, (2005) suggest that in most cases the evolution of product architecture shifts from one of "*complex, non-standard interfaces, through simple company-wide standard interfaces and ultimately to industry-wide standards*" (2005:15).

In a similar vein, the dominant logic in explaining shifts in firm and industry architecture is conceptualised as a set of ‘centrifugal forces’ (Jacobides, *et al.*, 2006) that push towards disintegration and ‘modular organisations’ especially in more mature industries. Schilling and Steensma (2001:1149) also highlight the dominant logic of disintegration, whereby “*the role of a tightly-integrated hierarchy is supplanted by loosely-coupled networks of organisational actors*”. In other words, analogous to the dominant logic of shifts in product architecture towards modularity, many scholars have also argued that firm and industry architecture tends to follow a path towards a loosely-coupled organisational form and disintegrated product markets populated by highly-specialised firms.

However, a number of authors have hinted at a reverse path towards integrated product designs. For example, Shibata, *et al.*, (2005:27) points out that “*returning to a product system that has a complex mapping of function to structure and a complex interface....is conceivable*” and Schilling (2000:312) also suggests that software packages and bicycles as examples where “*sets of components that were once easily mixed and matched may sometimes be easily bundled into a single integrated package that does not allow (or that discourages) substitution of other components*”. Moreover, MacDuffie (2013:12) contends that reverse swings are possible when “*technological change or customer demands for new functionality can redefine module boundaries [which] may increase interdependencies across modules, and can reverse maturation processes that lead to dominant designs*”. Empirically, Christensen, *et al.*, (2002:972) highlight the role of technology and argue that increasing integration is associated with a ‘performance gap’ in product markets. In other words, firms continually improve products to a point where they overshoot the demands of users such that “*disruptive technologies – simpler, more convenient products that initially do not perform well enough to be used in mainstream markets – can take root in undemanding tiers of the market and then improve at such a rapid rate that they squarely address mainstream market needs in the future*” (Christensen, *et al.*, 2002:961).

An alternative explanation is offered by Cacciatori and Jacobides (2005:1854) who examine the building and construction industry and observe that specialisation of productive capabilities eventually leads to cracks in the overall system at the aggregate level which causes pressure for reintegration, compounded by the presence of exogenous shocks. As it becomes evident that the existing specialised productive capabilities in a value chain are insufficient to meet the demands of users, firms respond by offering new products or services. However, Cacciatori and Jacobides (2005:1867) argue that firms often pursue a form of reintegration that preserves its original advantage. In other words, “*when threatened with the spectre of commoditisation, or intensified competition, firms will pursue aggressive strategies of reintegration to protect their profitability*” such that “*this provides them with new sources of revenue and allows them to enter higher margin parts of the industry, using their existing capabilities as the ‘thin end of the wedge’; they use their scope and capabilities as their entry point to integrated service provision which is expected to be more profitable than their existing business...*”.

In the Swiss watch industry, Jacobides and Winter (2005:405) argue that “*...the cycle pushing towards specialisation get reversed when new and superior capabilities arise from knowledge bases that are misaligned with the existing vertical structure of the industry*”. New productive capabilities emerge in a product market as a consequence of a technological innovation, in this context one based on miniaturisation and micro-electronics. As a consequence, the existing vertical structure of the product market was unable to effectively respond which favoured vertically integrated firms Japanese firms. The ensuing selection process then forced out the co-specialised structures of the incumbent Swiss manufacturers.

Finally, Schilling and Steensma (2001) argue that firm and industry architectures may resist or reverse the shift towards a more modular form. This may occur where a firm architecture has performance advantages or cost savings from integrating particular activities. Schilling and Steensma (2001) also draw upon the arguments of TCE to highlight that in the presence of asset specificity, transaction costs may overload the productive system which leads to a disincentive to shift towards a more loosely-coupled form. As such, the shift towards a loosely-coupled firm architecture is not a given; rather, it is a trade-off between “*the gains achievable through greater flexibility and the performance advantages of integration*” (Schilling and Steensma, 2001:1163).

Despite these contributions, Fixson & Park (2008:1312) highlight that few empirical studies have examined architectural co-evolution in complex supply networks, and therefore do not distinguish between horizontal and/or vertical structures, or moreover, how industry structures evolve across time. In this research study, I aim to contribute to the gap identified by Fixson & Park and explain how and why product markets reintegrate horizontally and vertically across time, but to go further and connect such processes with changes in product architecture. As highlighted, the extant literature on reintegration has tended to focus on either endogenous processes at firm level or exogenous processes at the industry level. I aim to use empirical data to propose more nuanced stylised types of hybrid product architecture, under conditions of reintegration. Few (if any) prior studies have done so.

2.1.9 Processes influencing architectural co-evolution

To consider how product architectures, and associated firm and industry architectures may co-evolve, I now turn to ‘modularisation processes’. As MacDuffie (2013:9-10) asserts, “*a more process-based view of modularity is needed to understand the now-commonplace observation that coordination and communication intensive activities can precede emergence of modular properties and standardised interfaces*” and observes that “*modularity-as-property and modularisation-as-process are so intertwined, so embedded in dynamics of reciprocal feedback that examination of one without the other is inevitably incomplete....*”.

The most striking observation from my review of the extant literature is the sheer breadth of theoretical perspectives that have been used to explain processes involved in architectural co-evolution. Theoretical explanations have been categorised by some scholars as either internal or environmental (ie, Suarez, 2004) or as supply-side and demand-side (ie Adner & Levinthal, 2001). However, the extant literature has primarily focused on supply-side explanations for examining co-evolutions in product and/or industry architecture, and the explanations are diverse and varied, including transaction costs (ie, Klein, Crawford & Alchian, 1978; Williamson, 1975), the exercising of real options (ie, Leiblein & Miller, 2003; Schilling & Steensma, 2002), the role of economic property rights (ie, Langlois, 2002; 2006), the role of knowledge (ie, Kogut & Zander, 1992; Monteverde, 1995, Tee, 2011), the role of resources (ie, Barney, 1991; Wernerfelt, 1984), the notion of comparative advantage (Jacobides, 2008; Jacobides & Hitt, 2005; Jacobides & Winter, 2005), competitive intensity (ie, Schilling, 2000); and technological evolution patterns (ie Dosi, 1982; Utterback & Abernathy, 1978; Schilling, 2000).

In this section of the literature review, my objective is to examine the *dominant* theoretical perspectives that permeate much of the work in the field. Upon reviewing the theoretical basis of prior contributions to the literature, the fields of transaction cost economics, productive capabilities and knowledge-based perspectives permeate much of the extant literature in high-ranking, peer-reviewed journals, such as Strategic Management Journal, Research Policy, and Academy of Management Journal. The scope of my discussion will, therefore, focus on these dominant theoretical perspectives.

2.1.9.1 Transaction cost economics

According to Sanchez, *et al.*, (2013) transaction cost economics (TCE) frames the problem of economic organising as a problem of contracting and examines why firms choose to make rather than buy (Williamson, 1985). While TCE is often associated with the choice of an efficient governance mode, it is an important perspective in the context of product architecture, as well as industry evolution as its main premise is that “*efficient organisation necessitates matching transactions which require higher levels of coordination with organisational governance forms which provide the necessary levels of coordination in a cost-efficient manner*” (Leiblein, 2003:939). In effect, its prediction is that of ‘discriminating alignment’ (Leiblein, 2003:940). TCE generally assumes that markets are more efficient than hierarchies, so long as the transaction costs of contracting do not exceed the costs of vertically integrating the activity with firm boundaries. In other words, firms look to minimise production costs, but also associated transaction costs (Williamson, 1985).

TCE distinguishes between ex-ante and ex-post transactions costs. Ex-ante transaction costs are associated with the “*costs of drafting, negotiating and safeguarding an agreement*” (Williamson, 1985:20) and with ‘small numbers bargaining’ problems (Williamson, 1975, 1979), whereas ex-post transaction costs include the costs associated with monitoring performance and enforcing the behaviour of parties (Williamson, 1985) as economic actors are subject to the “*frailties of motive*” (Simon, 1982:303), the potential for hold-up, and with challenges in measurement and coordination problems (Leiblein, 2003). Ex-post small numbers bargaining problems may also arise in the presence of specialised investments. Small numbers bargaining can increase transaction costs as “*when there are few suppliers available for a resource, the supplier, aware that the firm has few alternatives, can exploit the buyer’s dependence*” (Schilling & Steensma, 2002:389).

Both ex-ante and ex-post transaction costs are associated with three dimensions of economic transactions: (1) uncertainty, (2) frequency and, (3) asset specificity (Carter & Hodgson, 2006; Williamson, 1979:239).

Uncertainty increases the number of potential contingencies in a contract, and increases the ex-ante costs of writing the contract and ex-post costs of enforcing any claims upon it (Williamson, 1985). Therefore, additional ex-ante costs of drafting an accurate contract are incurred, and despite the best efforts of both parties the contract will inevitably contain a degree of ‘incompleteness’ (Baldwin, 2008; Langlois, 2003, Leiblein & Miller, 2003), which may lead to ex-post costs arising from adapting contractual arrangements that may require further negotiations or ‘hold ups’ if transaction partners engage in opportunistic behaviour. Incompleteness often arises as firms “*neither have perfect information nor possibly foresee all future contingencies*” (Wolter & Veloso, 2008:588). According to Sanchez, *et al.*, (2013), the frequency of transactions acts as a kind of multiplier on ex- post transactions costs associated with each separate contract the firm writes. In other words, a low frequency of transactions with an external firm can increase ex-post transaction costs, and a high frequency of transactions may help mitigate ex-post transaction costs. Finally, asset specificity is associated with the creation of ‘quasi-rents’ which may be subject to hold-up or ex-post opportunistic behaviour (Klein, Crawford & Alchian, 1978). When asset specificity and uncertainty combine, transaction hazards are most likely to occur (Leiblein, 2003) and therefore vertical integration may prevent unnecessary haggling, and it helps harmonise risk perceptions and expectations (Williamson, 1971).

Transaction costs are not limited to transactions in the intermediate market. For example, organising production activities within firm boundaries can create costs associated with depreciation, agency, co-ordination and shirking (Jacobides, 2005). As a result, firm boundary decisions are often associated with the relative trade-off between bureaucratic, production and transaction costs compared within- and across firm boundaries (Williamson, 1985). Jacobides (2008:310-11) further elaborates by arguing that transaction costs should be conceptualised as the ‘net’ transaction costs of using the market; in other words, transaction costs (TC) less bureaucratic costs (B).

How firms' architectural choices affect ex ante and ex post transaction costs is a recent concern in the transaction costs economics literature (Bach & Galvin, 2011; Baldwin, 2008; Langlois, 2006; Sanchez, *et al.*, 2013). In discussing the effects of increasing product modularity on transaction costs, Baldwin (2008:156) distinguishes between modular product component boundaries which have 'thin-crossing points' with low transaction costs, and the interior of product components which have 'thick-crossing points', and hence high transaction costs. Baldwin further distinguishes between so-called types of 'mundane transaction costs' that can be traded-off and incurred ex-ante in order to reduce ex-post transaction costs, and characterises the costs of creating modular product architectures as an ex-ante mundane transaction cost. Furthermore, Langlois (2006) characterises the creation of industry standards as a type of mundane transaction cost.

Sanchez, *et al.*, (2013) further apply Baldwin's (2008) notion of mundane transaction costs to a firm's choice of architecture; in their view, it depends on the relative attractiveness of incurring ex-ante versus ex-post transaction costs related to the creation and use of a given architecture. Choosing a closed architecture may give a firm strategic benefits resulting from exclusive control and use of its architecture, but it is also a decision to incur full development costs resulting in potentially high ex-ante mundane transaction costs and a 'normal' level of ex-post opportunism transaction costs (Baldwin, 2008). Ex-ante mundane transaction costs may include the costs involved in design, development, and establishing IPR, and ex-post transaction costs would be 'normal' by reference to the costs of protecting and defending its intellectual property rights (Sanchez, *et al.*, 2013). On the other hand, where a firm incurs ex-ante transaction costs in creating and/or using an open architecture, it may benefit from lower levels of ex-post transaction costs, as well as reduced ex-ante costs in relation to *future* product variants and upgrades (Sanchez, *et al.*, 2013). However, where firms engage in an *existing* open and modular product architecture, it is therefore likely that both ex-ante and ex-post transaction costs may be reduced, such as via sharing ex-ante development costs among product market participants.

A further important contribution to the TCE literature is that transaction costs are dynamic and transient (Jacobides, 2005; Jacobides & Winter, 2005; Langlois, 1992; 2003; 2006). Jacobides and Winter (2005:402) observe that firms try to “*manipulate and shape the transactional environment to their advantage*”. For example, new production processes or the specification of interface standards may help to reduce TC that “*enables specialisation, thus reshaping its institutional context for the next period*”. Falling transaction costs, therefore, tends to produce a shift in the direction of more modular product architectures and more specialised product markets. However, on the reverse side of the coin, there is often a contrary push back towards integration, increasing transaction costs, as tasks become “*increasingly routine, and transfers increasingly standardised, it sometimes begins to pay not to subdivide tasks further but rather to integrate tasks...*” (Langlois, 2006:1403).

Ultimately, Sanchez, *et al.*, (2013) suggest that a firm’s choice of architecture is likely to be associated with its ability to appropriate rents from its product development activities. Gains from specialisation are often available to firm’s who specialise in a particular activity, and gains from trade are realised by firms who trade with specialist firms by obtaining lower-cost and/or higher-quality product components (Jacobides, 2005). The pressure for or against decomposition of a product architecture into modular product components is associated with the respective gains achievable such that “*the balance between the gains achievable through recombination and the gains achievable through specificity determines the pressure for or against the decomposition of the system*” (Schilling, 2000:317-8). Schilling continues “*as a product is a nested hierarchy, each product component is a system of other components...and should gains from recombining win out, the system will be decomposed into further fine-grained components, until the system reaches a level composed of relatively homogenous inputs or demands*”. On the reverse side of the coin, if gains from integration win out, the system may migrate towards integration “*if the system is part of a larger system, and if there are strong gains to be achieved through making this component specific to a particular combination, it might become more tightly integrated with another even larger system*” and thus “*the trajectory of systems is bi-directional....towards or away from increasing modularity*”.

When a product architecture is closed, firm's often believe that it can succeed in the product market by using its own specialised productive capabilities to design and develop products (Sanchez, *et al.*, 2013). As a consequence, in such circumstances a firm is likely to use significant vertical integration as a means to capture gains from integration, although it may use outsourcing where it believes it does not have a comparable advantage.

Open architectures offer opportunities for firms to specialise in specific types of product components who hold a comparative advantage in designing and/or producing specific kinds of product components (Jacobides, 2005). Sanchez, *et al.*, (2013) suggest that both product component suppliers and assembler firms can benefit from both gains from specialisation and gains from trade when this occurs. For example, manufacturer/assembler firms may realise gains from trade in a number of ways. For example, a firm may source lower-cost product components from specialist firms owing to the economies of scale advantages the specialist firm has, or a firm may source higher-quality product components that may earn it a price-premium for the final, assembled product (Sanchez, *et al.*, 2013).

In sum, the modular structure of a product architecture and its associated task network is not fixed; as Baldwin (2008:156) argues, new modularisations can occur and new product component boundaries can have significant implications for firm and industry structure. In other words, dynamism in the transaction cost environment and its implications for respective gains from trade and/or specialisation versus gains from integration can often push a product architecture, and associated firm boundaries, both towards and away from modularity and specialisation as new innovations and modularisations occur.

2.1.9.2 Productive capabilities perspective

The resource-based view is an important theoretical perspective in the strategic management literature and can be traced back to the work of Penrose (1959) and seminal contributions by Wernerfelt (1984) and Barney (1991). The resource-based view is concerned with resources; their tacit-ness and how hard they are to imitate and this may impact upon a firm's competitive advantage and its sustainability. The 'view' has also been utilised as an approach to analyse firm boundaries (ie, Connor & Pralahad, 1996; Poppo & Zenger, 1998) predicting that the acquisition of hard-to-imitate resources will lead to vertical integration (Leiblein & Miller, 2003; Schilling & Steensma, 2002).

The capabilities perspective also has a long tradition in the firm boundaries literature, captured in the notion of 'comparative advantage' (Jacobides, 2008; Jacobides & Hitt, 2005). Whereas the resource-based view focused on resources, and a firm's attempts to gather and exploit unique assets in the pursuit of competitive advantage (Affuah, 2001; Leiblein & Miller, 2003), the notion of comparative advantage highlights the role of heterogeneous capability distribution along a value chain. According to Jacobides and Hitt (2005), heterogeneous productive capabilities arise "*as a result of a path-dependent learning process in which there is abundant opportunity for various contingencies to shape the way of doing things that ultimately emerges, even if all firms have access to homogenous resources*" (Jacobides, 2005). In other words, a firm's productive capabilities are path-dependent (Bach & Galvin, 2011) and arise from prior investments in knowledge, and the strategic choices previously taken that impact future opportunities (Jacobides & Hitt, 2005). Subsequent variations in productive capabilities have the effect of creating specialised knowledge bases that may influence product market specialisation and the gains from trade that are appropriable (Jacobides, 2005). Hence, firms do not need a competitive advantage, but rather a comparative advantage, relative to other firms, in order to appropriate gains from trade and/or specialisation (Jacobides, 2005; 2008).

The basic logic, therefore, of comparative advantage is that productive capability heterogeneity along a value chain provides the rationale for specialisation and ‘make-or-buy’ decisions. It would appear logical to assume that firms will choose to specialise in activities where they hold a comparable advantage, and use external firms where they do not, transaction costs permitting (Jacobides, 2008; Jacobides & Hitt, 2005). In other words, for specialisation to occur, “...*the gains from trade must simply outweigh the cost of using the market....*” (Jacobides & Hitt, 2005:1213).

As hinted at in the previous section, whereas the transaction costs tradition has illuminated the role of the aggregate sum of production costs and transaction costs in determining architectural choices, the productive capabilities tradition has foregrounded the role of heterogeneous capability distribution in a value chain. According to Jacobides and Hitt (2005:1212), “...*once we allow for differences in the productive abilities of different firms, the relevant choice is between the firm, with its internal cost levels, vs the price and transaction costs of interacting with another firm*”. Transaction costs do, therefore play a role. However, according to Jacobides and Winter (2005:398) “...*their role is akin to that of a sales tax or a tariff levied on international trade*” and they hypothesise that where “*productive capability asymmetries are high, such ‘taxes’ might not be sufficient to curb vertical specialisation*” and conversely “*even if the level of the TC tax is low, if the gains from trade are even smaller...then there will be no reason for vertical specialisation to emerge*”.

Jacobides and Winter (2005:400), moreover, suggest that heterogeneity in productive capabilities can be conceptualised as an evolutionary process; based upon product market selection forces, such as competition, the dynamic level of transaction costs; and, the emergence of new productive capabilities, as new forms of knowledge become relevant. Firstly, (1) the forces of product market selection tend to reinforce the existing institutional structure, such that when specialised firms have superior productive capabilities, selection tends to push for greater specialisation. On the reverse side of the coin, where integrated firms have superior capabilities, selection forces further integration. Second, (2) transaction costs are dynamic and transient and subject to managerial manipulation, either through specialisation or technological imitation.

Thirdly, (3) changes in vertical scope, such as increased integration or specialisation, necessarily affect the knowledge accumulation and productive capability development processes, especially as knowledge accumulation and capability development tends to occur more quickly in specialised firms. Ultimately, this may lead to increases in gains from trade and/or specialisation, providing increasingly specialised knowledge bases do not become a core rigidity. According to Jacobides and Winter (2005:405), these evolutionary forces suggest a drift towards ever-increasing specialisation in a product market.

However, the authors point to an evolutionary drift in the reverse direction. Specifically, they argue, *“the cycle pushing towards specialisation gets reversed when new and superior capabilities arise from knowledge bases that are misaligned with the existing vertical structure of the industry. This sets in motion a process that may eventually make vertical integration typical, endogenously increasing TC along the way”* (Jacobides & Winter, 2005:402). Jacobides and Winter (2005) also cite Williamson (1983) in arguing that integration is more likely whenever asset-specific investments are called for, which is often the case as new technologies, associated with new knowledge and productive capabilities, favouring existing integrated firms.

Cacciatori and Jacobides (2005:1864) examine a case of reintegration in the building and construction industry and assert that *“knowledge bases shape and, crucially, constrain the trajectories of capability development”*. In other words *“the vertical division of labour creates ‘trajectories’ that define and prescribe the way in which knowledge and competency will evolve over time”*. Furthermore, the authors highlight that, in their setting at least, changes in the productive capability development trajectory led to the worsening of some important elements to the coherence of the overall system which increased pressures for reintegration, compounded by exogenous shocks.

Cacciatori and Jacobides (2005), moreover, suggest that it is often the aggregate performance of the vertical architecture of a product market that may sustain it, or push the system back towards re-integration. When the vertically-specialised structure is subject to demand changes or exogenous shocks, it is possible that the aggregate performance of the vertical architecture becomes mismatched and the pressures for reintegration emerge. Furthermore, firms may begin to reintegrate when *“threatened with the spectre of commoditisation, or intensified competition [and] firms will pursue aggressive strategies of reintegration to protect their profitability”* using their existing productive capabilities and scope as the thin edge of a wedge to integrated provision (Cacciatori & Jacobides, 2005:1867).

In sum, the central argument of the productive capabilities perspective is that capabilities are path-dependent and tend to become more specialised over time as firms internalise those activities in which they hold a comparative advantage, and outsource those activities in which they do not. The heterogeneity of productive capabilities distributed along a value chain will, therefore, become less correlated as productive capability development tends to follow a particular trajectory. The presence of heterogeneity along a value chain reveals gains from trade and/or specialisation from transacting with external, specialised firms, transaction costs permitting. However, this self-reinforcing process is, again, not without its limits. First, far-reaching productive capability heterogeneity can lead to ‘gaps’ appearing in the overall product system which may cause a push towards reintegration as new superior knowledge bases emerge that are misaligned with the existing specialised vertical structure, increasing transaction costs and thereby reducing the extent of gains from trade and/or specialisation. Second, the threat of imitation and commoditisation inherent in modular product markets can cause a shift back towards reintegration in order to capture gains from integration. And, thirdly, exogenous shocks to a product market can cause technical discontinuities which may lead to increased asset specificity and existing knowledge and capabilities being misaligned.

2.1.9.3. Knowledge-based view

According to Williamson (1985), there are four types of specificity: (1) site asset specificity, (2) physical asset specificity, (3) dedicated asset specificity and (4) human asset specificity. TCE usually advocates that asset specificity is associated with the delineation of firm boundary decisions. However, the knowledge-based view of the firm posits that ‘firm-specificity’ (or ‘human asset’ specificity) is a key determinant in firm boundary choices. The key principle of the knowledge-based view is that firms generate and exploit knowledge (Kogut & Zander, 1992), and that firm boundary decisions will be based upon the relative efficiency of coordination of complementary activities (Connor & Pralahad, 1996; Kogut & Zander, 1992). Within firm boundaries, firm specificity generates a shared language, shared knowledge and routines, and enhances coordination efforts (Kogut & Zander, 1992). Investments in firm-specific routines, language and skills are critical to firm performance (Barney, 1991) and productive capability development.

Where the knowledge-based view significantly departs from TCE is that specificity, rather than damaging governance efficiency, actually increases it (Monteverde, 1995, Nickerson & Zenger, 2004). According to Monteverde (1995:1628-9), as an activity becomes more firm-specific, it increasingly develops a ‘common organisational communication code’ and provides ‘one single organisation-specific dialect’ which enables the dissemination and codification of knowledge within the firm. Thus, according to Poppo & Zenger, 1998:857), “...relative to markets, firms simply possess advantages in generating firm-specific language and routines that efficiently yield valuable capabilities. In other words, a firm’s routines and productive capabilities are based on knowledge, which is distributed across individuals and can be assembled in various ways (Connor & Pralahad, 1996), without the problems of information asymmetry and opportunism (Wolter & Veloso, 2008).

When intensive knowledge-based transactions are attempted in the intermediate market, however, they are often over-burdened with transaction costs. This is because, as described in the work of Conner and Pralahad (1996), knowledge is often firm-specific and tacit, that is, un-coded and embedded in firm routines and language. As a consequence of its tacit-ness, it is often difficult to write-down, duplicate or transfer, and individuals seldom find it easy to describe his or her know-how to others, increasing information costs (Schilling & Steensma, 2002) and creating information asymmetries across firm boundaries (Jacobides & Hitt, 2005). The advantage of knowledge tacit-ness, on the other hand, is that it more easily transferred within firm boundaries and cannot be easily appropriated by competitors (Teece, 1986). Therefore, despite the possible advantages of collaborating with external firms, the knowledge-based view posits that increased knowledge specificity permits a firm to make more efficient use of its firm-specific language, knowledge and routines, and hence is likely to result in product development within firm boundaries. In other words, while the knowledge-based view departs from TCE in its theoretical underpinnings, its prediction is the same; firms will internalise activities in the presence of specificity.

However, too much firm-specific knowledge is not without its risks, especially under conditions of technological uncertainty (Wolter & Veloso, 2008). For example, Poppo and Zenger (1998) studied the information services sector and highlighted that in dynamic or fast-changing product markets, internal routines and shared language and knowledge can easily become a core rigidity that negatively affects firm performance, as *“unique language, while efficient, may quickly become the wrong language”* and *“groups of individuals governed by markets are more likely to directly benefit from the formation of new knowledge”* (Poppo & Zenger, 1998:873). In other words, technological change can render existing firm knowledge and capabilities obsolete (Henderson & Clark, 1990) if the skills and knowledge required do not build upon existing ones. In addition, Baldwin (2008) notes that whenever a firm’s routines, capabilities or competencies change this may cause its knowledge boundaries to shift, and so the scope of its knowledge will change over time as required by its changing activities. As a consequence, vertical integration may become an inflexibility that hinders a firm in acquiring new knowledge and innovative potential (Affuah, 2001, Wolter & Voleno, 2008). Integrated firms may, under conditions of technological uncertainty, miss out on innovative ideas as a result of vertical integration (Langlois & Robertson, 1989).

Furthermore, external firms in the intermediate market may have superior, specialised productive capabilities and “*this heterogeneity of available technological trajectories may confer evolutionary advantages to disintegrated firms*” (Wolter & Volenso, 2008:591).

The prediction of the knowledge-based view is, therefore, different to TCE; whereas TCE may infer that technological uncertainty increases transaction costs, and hence promotes internalisation, the knowledge-based view posits that technological uncertainty enhances the benefits of acquiring access to external sources of knowledge that are available in the intermediate market, but there is a trade-off against the need to protect valuable knowledge.

Wolter and Volenso (2008) helpfully explore the effects of Henderson and Clark’s (1990) innovation typology of incremental innovations, modular innovations, architectural innovations, and radical innovations on the degree of vertical integration in a product market. The authors contend that incremental innovation is unlikely to affect the degree of integration as transaction costs and existing knowledge is largely undisturbed. However, for modular innovations the degree of vertical integration will decrease as the need for new productive capabilities and falling transaction costs offsets the additional coordination needs. Finally, architectural and radical innovations may result in increased vertical integration owing to higher transaction costs and coordination needs. The impact of this work is that it highlights the fact that the type of product innovation, whether incremental, modular, architectural or radical, has profound implications for product market structure.

Product modularity requires knowledge to be codified and articulated. Decomposition of a product architecture into sub-tasks is conceptualised as a key mechanism in making knowledge less tacit. According to Ulrich (1995), product partitioning reduces cognitive complexity (Ulrich, 1995) and it is often a gradual and recursive process (von Hippel, 1990). Once knowledge is explicated, it can then be codified in the form of manuals, databases, and other systems and shared with other product development teams within the firm, and/or with other external firms, facilitating the creation of an efficient intermediate market (Jacobides, 2005).

The codification of knowledge is often likely to happen initially within firm boundaries, before permeating across firm boundaries. For example, Hoetker (2006) cites Putterman (1995) in arguing that “*communication routines and a common language for discussing technical issues arise more quickly within the firm because intra-firm communication is normally more frequent and intense than with outsiders*”. Tee (2011) also argues that knowledge is articulated within firm boundaries initially, as a form of ‘intra-firm’ modularity (Chuma, 2006). However, MacDuffie (2006) highlights the inherent difficulty of maintaining transactions within firm boundaries once product modularity has emerged.

The notions I have discussed are similar to the idea of structured or unstructured technical dialogue described initially by Monteverde (1995) and later by Christensen, *et al.*, (2002). To structure technical dialogue, Christensen, *et al.*, (2002) argue that an activity must be able to be specified and measured and the parties must be able to understand the interdependencies between the specific activity and the architecture, thereby reducing transaction costs. Furthermore, Christensen, *et al.*, (2002:959) assert that for transactions to occur across firm boundaries, three conditions of ‘modularity’ must be met; (1) specifiability, (2) measurability, and (3) predictability, as without these conditions present, the potential for an efficient intermediate market to exist is limited. Once the conditions are met, a product component and its interface is modular, across which structured dialogue can then happen. Unstructured technical dialogue, associated with high transaction costs, in contrast, continues to occur at the boundary of interdependent product components.

Jacobides (2005) suggests similar conditions for an efficient, functioning intermediate market. The first condition is that coordination must be simplified and that interactions along a value chain must be minimised and production modularised (Baldwin & Clark, 2003). Without simplified coordination, Jacobides (2005) argues, it remains impossible to engage with external firms in an intermediate market. Complex coordination, on the other hand, necessitates the use of a firm and its formalised authority mechanisms (Langlois, 2003). The second step is that information-exchange between parties must become standardised so that information asymmetries are reduced. Jacobides (2005:482) suggests that intermediate markets become more efficient when there is an ability to “*specify what would be transacted across firm boundaries regardless of the presence of...transactional hazards*”.

In other words, the specification of standard templates is viewed as critical for market procurement (Baldwin & Clark, 2003, Tee, 2011). Specification often makes explicit what was previously tacit, and standardised methods of exchanging information replace each firm's own shorthand and particular ways of describing things. As Jacobides (2005:482) states, *“if information on what is desired is unclear or difficult to articulate, outside procurement arrangements cannot be set up, as desired items cannot be identified ex-ante”*.

Jacobides (2005:483) also highlights a further possibility that simplified coordination and standardised information exchanges can also lead to ‘inter-organisational partitioning’. Inter-organisational partitioning leads a firm to create autonomous sub-units which may ultimately lead to market-based exchanges. Gains from trade, according to Jacobides (2005), motivate this process, and when coupled with the emergence of ‘internal shadow prices’ that govern inter-organisational relationships, this can push organisational units to consider make or buy decisions, forcing transactions with external firms where gains from trade are appropriable.

The modularisation processes associated with the simplification and exchange of knowledge may underpin the formation of modular product designs and ‘modular’ industries as an emergent phenomenon (Boisot & Sanchez, 2010). As Langlois and Robertson (1995) assert, the emergence of new intermediate markets therefore requires modes of interaction which enable newly specialised buyers and sellers to interact through a market interface. If this occurs, firms may ‘go modular’, forcing other firms to follow or be selected out (Sanchez & Mahoney, 2013). This firm-level evolutionary perspective emphasises that new product and firm architectures often emerge in a product market as a consequence of endogenous managerial action that may lead to a self-reinforcing mechanism whereby the incentives to adopt modular designs increase (Sanchez & Mahoney, 2013).

Fixson and Park (2008), however, in their case study of re-integration in the bicycle industry, present the opposite view. They suggest that designers seek value in fully integrated designs because they can be better than modular designs along some dimension. The authors argue architectural knowledge or a broad scope of product component knowledge permits reintegration. According to Fixson and Park (2008:1310), “*knowledge across several segments appears to have been a necessary ingredient for maintaining competitiveness in the wake of [a] architectural shift*”. Broader product component scope, they contend, can help firms avoid the modularity trap (Chesborough & Kusunoki, 2001). In other words, a broad scope of knowledge, a la Brusoni, *et al.*, (2001), may assist firms in identifying new architectural innovations.

In sum, knowledge plays a key role in the modularisation process. Firms are often conceptualised as knowledge exploitation entities, however too much firm-specific knowledge can lead to information asymmetries, increasing transaction costs, and it may ultimately also lead to the risk of core rigidities. Joined with the productive capabilities perspective, it also suggests that, over time, firms may be out-fought by value chains with highly-specialised knowledge and productive capabilities putting the firm at a comparative disadvantage in relation to more specialised participants. Partitioning a product architecture is a key modularisation process that pushes a firm to ‘structure’ its technical dialogue in the form of standard templates or specifications, initially within firm boundaries and then eventually across firm boundaries. As a firm decides to use standardised templates and specifications as the basis for market exchange, ex-post transaction costs may fall and gains from trade and/or specialisation may become appropriable. However, discontinuities, perhaps triggered by exogenous or technological shocks may render existing knowledge and productive capabilities obsolete which may act as force back towards integration, increasing information asymmetries and tacit-ness, with increasing ex-post transaction costs. Or, as gains from trade and/or specialisation reach their limits, designers may seek value from appropriating the architectural or broad scope of product component knowledge that underpins a new, discontinuous integrated design.

Section summary

The supply-side processes that may underpin architectural co-evolution is linked to the intended contribution associated with RQ1 – “What supply-side processes are associated with the co-evolution of architectures?”. Our understanding of how such processes combine in different ways over time to propel architectures either towards or away from product modularity and industry specialisation is not empirically well-advanced. In this research study, I aim to examine whether and how such processes combine. Furthermore, prior work assumes such processes act at a systemic level, but how such processes act upon decisions at the product component level to evolve product components towards being more modular or more integrated is less well-understood.

2.1.10 Summary

In section 2.1 of this chapter, I have addressed the concepts of modularity as a design principle and as a process. Firstly, I drew extensively upon the modularity literature to propose that product architectures can be conceptualised along two continuums: as integrated or modular and as open or closed. I argued that the 2x2 matrices proposed by Sanchez (2008) and Shibata, *et al.*, (2005) had weaknesses and failed to reflect the evolvability of product architectures. In turn, I embed the notion of a hybrid product architecture and formalised a stylised product architecture typology. Lastly, I discussed how prior work has conceptualised architectural co-evolution by examining the literatures of transaction cost economics, and the productive capabilities and knowledge-based perspectives.

The aim has to been to argue that architectures are not static, and co-evolve in response to a number of processes that combine to reshape architectures over time. These combined forces are often overlooked in static empirical studies and can only be uncovered by examining the processes, and their interaction, across time.

I have orientated the literature review on architectural co-evolution towards two significant gaps in the literature, (1) connecting product architecture types to industry reintegration to propose more nuanced stylised hybrid product architectures types and (2) to examine how supply side modularisation processes combine across time to propel architectures towards or away from product modularity and industry specialisation. Moreover, I examine how these processes combine at the product component level within a given product architecture.

I now turn to the second section of the literature review and examine the theoretical origins of architectural correspondence, the hypothesised correspondence between product component, task, knowledge and firm boundaries. I then examine prior empirical work and discuss recent contingent perspectives.

2.2 Architectural correspondence

2.2.1 Introduction

The notion that firms should be designed to reflect the nature of the tasks they perform has a long tradition (Galbraith, 1977; Lawrence & Lorsch, 1967). To the extent that different product architectures require different types of tasks to be performed, either within- or across firm boundaries, it is conceivable that the architecture of firms and products should reflect ‘discriminating alignment’ (Leiblein, 2003), in which the governance structure of a transaction is matched with the underlying characteristics of the transaction (Williamson, 1996). Langlois and Robertson (1992) recognised how modular product designs may affect industry structure and in their 2003 article “The Vanishing Hand: the Changing Dynamics of Industrial Capitalism” arguing that *“by allowing specialist producers...to concentrate...on particular components, a modular system thus enlists the division of labour in the service of innovation”*. Sanchez and Mahoney (1996:64), however, were one of the first scholars to formalise a broad proposition suggesting a relationship between product architecture and firm architecture. For example, *“product designs composed of tightly-coupled components will generally require development processes carried out in a tightly-coupled organisation structure coordinated by a managerial authority hierarchy, an organisation design typically achieved within a single firm”*, and, in the case of a modular product architecture, *“the tasks within a multi-divisional firm are intentionally designed to require low levels of coordination so that they can be carried out by an organisational structure of quasi-independent divisions functioning as loosely-coupled sub-systems”*.

Subsequently, labelled the ‘mirroring hypothesis’ by Colfer (2007) and formalised by Colfer and Baldwin (2010), architectural correspondence suggests that the architecture of a firm may often correspond to the architecture of the product it designs. With intuitive appeal, empirical research has flourished in recent years, such as aircraft engines (Brusoni, *et al.*, 2001; Sosa, Eppinger & Rowles, 2004), computer notebooks (Hoetker, 2006); bicycles (Fixson & Park, 2008; Galvin & Morkel, 2001); software (MacCormack, *et al.*, 2012), semiconductors (Funk, 2008); motor vehicles (ie, MacDuffie, 2013; Sako, 2003), and air-conditioning systems (Cabigiosu & Camuffo, 2012; Furlan, Cabigiosu & Camoffo, 2014).

2.2.2. Theoretical foundations

Despite being criticised as being technologically-deterministic (Sako, 2003), the general idea of architectural correspondence is that the architecture of a firm corresponds to the product architecture it designs to the extent that an integrated firm is necessary for developing an integrated product architecture, whereas a modular firm is necessary for developing a modular product architecture (Colfer, 2007; Colfer & Baldwin, 2010; Fine, *et al.*, 2005, Sanchez & Mahoney, 1996). In other words, ‘products design organisations’ (Sanchez & Mahoney, 1996).

While the mirroring hypothesis examines the architectural correspondence between one structural layer of a system and another, there are different targets for correspondence. Colfer (2007:5-6) identifies four types of architectural structure that are amenable: (1) the product; (2) the product development firm; (3) the division of labour and (4) the division of knowledge. In her essay, Colfer suggests that the target for modularisation is not always the same as it depends upon the structural layer under examination. For example, (1) in the case of the product architecture, the target for modularisation is the interdependencies between product components, (2) in the case of the firm, the target is the information-sharing and communication patterns between the organisational actors who develop the product (such as individuals, teams, firms), (3) in the case of the division of labour, the target is the technical product interdependencies among the product development tasks of different firms, and, (4) in the case of the division of knowledge, the target is the technical interdependencies among the information/skill sets required to perform the development tasks (Colfer, 2007:6).

In sum, Colfer (2007) suggests that we can talk of a ‘within-firm’ mirroring – the extent of architectural correspondence between the product and the product development firm – and an ‘across-firm’ mirroring – the extent of architectural correspondence between the product and the division of labour and/or division of knowledge across firms. In more specific terms, I elaborate further that correspondence is hypothesised to be present when the partitioning of product component, task and knowledge boundaries correspond within one firm (in the case of within-firm correspondence) or across two or more firms (in the case of an across-firm correspondence)²³. In other words, an across-firm correspondence encompasses the additional partition of firm boundaries, as shown in Figure 6.

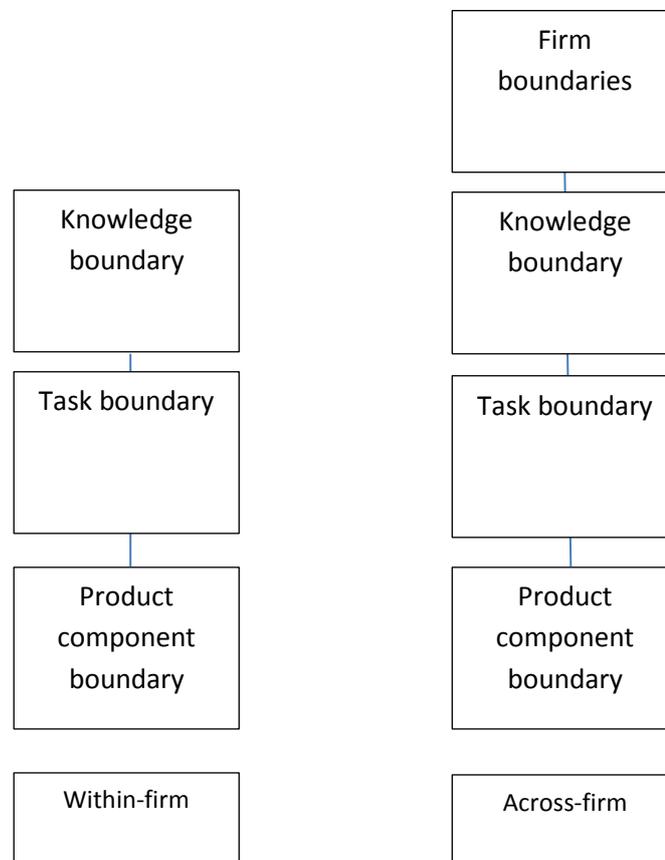


Figure 6: Within- and across-firm correspondence

²³ Colfer (2007) and Colfer and Baldwin (2010) had assumed a simple integral-modular continuum in making their assertions. Hence, the description I offer accommodates the notion of closed and modular and hybrid product architectures

In order to examine the hypothesised correspondence between product component, task, knowledge and firm boundaries, many scholars have used a proxy measure. Colfer and Baldwin (2010:6) argue that an ‘ideal test’ for architectural correspondence is one that analyses how product architectures may relate to other layers and would seek to examine organisational ties such as (1) firm co-membership, (2) geographical co-location, or (3) communication and information-sharing patterns, suggesting that it is then possible to see whether “*a technical dependency was correlated with the presence (or absence) of a given type of organisational tie*”. Colfer and Baldwin (2010) argue that the type of organisational tie used as a proxy for correspondence is not uniform across empirical studies – one possible explanation of the inconsistent results.

In this research study, I highlight that the lack of uniformity in how correspondence or non-correspondence is assessed in prior work can lead to mixed results, and correspondence and non-correspondence at the same time. In order to overcome the limitations of conflicting results, I aim to propose a ‘weak test’ and a ‘strong test’ of correspondence that may provide much-needed uniformity in future scholarly work.

In the *within-firm* literature, individuals or teams within the boundary of a single firm share co-membership of the same firm and therefore an appropriate organisational tie is often (1) information-sharing and communication patterns, or (2) co-location or geographical proximity (Colfer & Baldwin, 2010). In the *across-firm* literature, product development activities take place across two or more separate firms, and product development teams lack firm co-membership. As such, across-firm correspondence has usually focused on communication and information-sharing patterns.

Despite a range of possible organisational ties to assess the degree to which correspondence is present or absent, communication and information-sharing patterns associated with product component, task, knowledge and firm boundaries is foregrounded in many prior studies. Its centrality can be traced back to the ideas of Parnas (1971; 1972) who stressed the potential benefits of ‘information-hiding’. Parnas’s (1972) observation was that high levels of information-sharing can overwhelm a product development project, and it would be more efficient to hide or encapsulate information within a product component boundary so that it cannot affect other parts of the system.

Under Parnas's (1972:1056) approach, every product component *"is characterised by its knowledge of a design decision which it hides from all others. Its interfaces or definition was chosen to reveal as little as possible about its inner-workings"*. The advantages of information-hiding may therefore extend to the reduction of development time through concurrency of product development tasks (Gomes & Joglekar, 2008:445). Baldwin (2008:166) eloquently describes why transactions between actors cause information to be hidden, *"an economical transfer of a good from its producer to a user constrains the surrounding transfers of information quite dramatically. The user cannot know everything about how the thing was made: if that information were necessary, the user would have to produce the thing himself, or at least watch every step of production. The efficiency of the division of labour would then collapse. By the same token, the producer cannot know everything about how the thing will be used, for then she would have to be the user, or watch every user's action. Thus, fundamental to the efficient division of labour is substantial "information hiding". In other words, product development actors need to have high levels of knowledge about their own domains, but little knowledge about each other's"*.

In terms of *integrated product designs*, the usual logic is that there is unlikely to be a clear correspondence of product component, task and knowledge boundaries and, as such, high-levels of information-sharing would be required to progress its design and development. Turning to the alternative organisational tie of geographical proximity v. dispersion, it is likely that under such circumstances the product development team would need to be co-located and therefore tasks encompassed within a single firm boundary. Whereas, in the case of *modular product designs*, the usual logic suggests that a clear correspondence of product component, task, and knowledge boundaries is possible and, as such, high-levels of information-hiding between product component teams (ie, low levels of information-sharing) are beneficial. Turning once more to the alternative organisational tie of geographical proximity v. dispersion, it is likely that under such circumstances the product development teams would be geographically-dispersed and therefore product development tasks spread across two or more different firms.

However, it is important to distinguish between ex-ante and ex-post information-sharing needs (Cabigiosu & Camuffo, 2012). Ex-ante information-sharing is often required to partition product component boundaries and negotiate interface standards which then permits low-levels of ex-post information-sharing, as product development proceeds according to the agreed standards. In Sanchez and Mahoney's (1996) terms, ex-ante information-sharing to agree standards acts as an embedded coordination mechanism ex-post. In contrast, non-correspondence occurs where high-levels of ex-post information-sharing may remain necessary for co-development or to resolve unexpected design issues. In other words, high levels of ex-post information-sharing may persist for a number of reasons, despite the presence of modular product components and interface standards. For instance, (1) firms often need to keep access to product component knowledge even where product component design is outsourced (Brusoni, *et al*, 2001; Zirpoli & Camuffo, 2009), or (2) firms may wish to engage in information-sharing and knowledge integration in order to develop or maintain systems integration capabilities (Brusoni, *et al.*, 2001).

2.2.3. Within-firm correspondence

Within-firm correspondence describes the degree to which product component, task and knowledge boundaries correspond within the boundaries of a single firm (Colfer & Baldwin, 2010). Sanchez and Mahoney (1996:65) were the first scholars to formalise within-firm correspondence and asserted that *"...processes for developing tightly-coupled component designs require intensive managerial co-ordination, since the change in the design of one component is likely to require extensive compensating changes in the designs of many inter-related components. Thus, product designs composed of tightly-coupled components will generally require development processes carried out in a tightly-coupled organisation structure coordinated by a managerial authority hierarchy, an organisation design typically achieved within a single firm"*. In other words, *"since developers must communicate to address the interdependence between their efforts, there should be an isomorphic relationship between the organisation's communication (and information processing) structure and the architecture of the product under development"*.

In the case of integrated product development, the fundamental problem is characterised as one of co-ordination and communication (Colfer, 2007:7) due to unintended information asymmetries and information-processing problems (March & Simon, 1958) as a high number of interactions and interdependencies between tasks increases the demand for information. The notion of a relationship between product component interdependencies and high levels of information-sharing is also hinted at in the earlier works of Galbraith (1973), who argued that task interdependence requires a greater requirement for information-sharing between design teams.

Within-firm correspondence has found broad conceptual and empirical support in the literature (ie, Henderson & Clark, 1990; Sosa, *et al.*, 2004). In a recent literature review of empirical studies, Colfer and Baldwin (2010) find that it received support in 77% (or 17 studies) of the 22 ‘within-firm’ studies examined with 23% (or 5 studies) not supportive. Incidences of non-correspondence were classified into two types. Type 1 (numbering 4) exceptions were cases where a co-located product development team within a single firm designed a modular architecture made up of largely independent components (ie, MacCormack, *et al.*, 2012). The sole Type 2 exception was in the software design industry, where integrated product design was undertaken by geographically- dispersed developers (Srikanth & Puranam, 2007, 2008). However, Argyres (1999) examined the development of the B2 stealth bomber as another example of non-correspondence where an integrated product was developed by a group of firms. Argyres found that the exacting technical requirements of ‘stealth’ had the effect that no-one single firm held all the capabilities and knowledge to develop the product independently, and hence high-levels of information-sharing was used across multiple firms by establishing a set of standardised rules, procedures and systems; in other words, according to Colfer (2007:21), “*the group did not simplify or standardise the product’s architecture, but rather, how it communicated about that architecture*”.

2.2.4. Across-firm correspondence

Modularity as a design principle has been suggested as a mechanism that facilitates task partitioning (Baldwin & Clark, 2000) and knowledge partitioning (Sanchez & Mahoney, 1996). An emerging theme in the literature is the extension of within-firm correspondence to encompass product development activities across two or more firm boundaries, whereby “*the technical architecture, division of labour and division of knowledge will ‘mirror’ one another in the sense that the network structure of one corresponds to the structure of the others*” (Colfer & Baldwin, 2010: 13). Sanchez and Mahoney (1996:65) asserted that “*in essence, the standardised component interfaces in a modular product architecture provide a form of embedded coordination that greatly reduces the need for overt exercise of managerial authority to achieve coordination of development processes, thereby making possible the concurrent and autonomous development of components by loosely-coupled organisation structures*” In sum, an across-firm correspondence contends that the product component, task, knowledge and firm boundaries correspond, and therefore low-levels of ex-post information sharing and geographical dispersion will be likely observed.

Colfer and Baldwin (2010) found that 74% (46 cases) offer support for an across-firm correspondence within the sample of 62 empirical studies and 26% (16 cases) do not. Where it was supported, it largely consisted of specialised industries developing modular products, including stereo systems and computers (Langlois & Robertson, 1992), bicycles (Galvin & Morkel, 2001), semiconductors (Funk, 2008), and mortgage banking products (Jacobides, 2005). The cases not supportive were all examples of independent and specialised firms developing an integrated product.

Where across-firm studies received only partial or mixed support, these consisted of cases of, on the one hand, where product component and task boundaries correspond, but exhibited a non-correspondence with firm boundaries. In other words, forms of closed modularity, where modular product development tasks were partitioned but did not move across firm boundaries²⁴ (Hoetker, 2006; Jacobides & Billinger, 2006; Parmigiani & Mitchell, 2009), which suggests that a firm may create modular product architectures, but still keep all organisational units under one corporate umbrella.

²⁴ Such closed modular examples may be re-considered to correspond under my framework. Colfer and Baldwin (2010) consider these examples as ‘partial’ support due to their assumption that modular product architectures are always associated with movements of tasks out of hierarchy

On the other hand, the second type of mixed results is where product component, task and firm boundaries correspond, but knowledge boundaries exhibit non-correspondence, and firms rely on high levels of information and knowledge-sharing to resolve architectural issues or coordinate their efforts (ie, Brusoni & Prencipe, 2006; Brusoni, *et al.*, 2001).

2.2.5 Contingent perspectives on architectural correspondences

Much of the conceptual and empirical research concerning architectural correspondence has focused upon whether the mirroring hypothesis holds, and only recently has it turned its focus to the contingent conditions under which it may hold (ie, Cabigiosu & Camuffo, 2012, Furlan, *et al.*, 2014). This is important, however, because as Sako (2003:230) highlights, “*there is no simple deterministic link between the type of product architecture and organisation architecture*”. Some initial empirical work that examines the conditions under which correspondence may hold have primarily focused on how product complexity (ie, Brusoni, *et al.*, 2001; Cabigiosu, Zirpoli & Camuffo, 2013; MacDuffie, 2013; Zirpoli & Camuffo, 2009) and the rate of product component change (Cabigiosu & Camuffo, 2012; Furlan, *et al.*, 2014) may lead to a non-correspondence of knowledge boundaries, and hence, high-levels of ex-post information-sharing. I will now turn to a discussion of prior empirical work that examines these contingencies.

In modular product architecture design, the usual logic is that product component, task, knowledge and firm boundaries can be partitioned ex-ante, leading to low-levels of ex-post information-sharing. Sanchez (2003:382) writes that “*the standardising of component interfaces based upon the firms current architectural knowledge largely decouples architectural knowledge-based processes from the component-level knowledge used to develop specific component design during product development*” and, as such, the contention is that manufacturers/assemblers of modular products can specialise in architectural knowledge and outsource product component design (and the knowledge underpinning it) across firm boundaries.

However, according to Henderson and Clark (1990:2-11), successful product development requires *both* product component knowledge, knowledge about each of the core design concepts, and architectural knowledge, knowledge about the way in which the components are integrated and link together. The relative importance of each type depends upon the stability of the product architecture; architectural knowledge is preferred when discontinuities occur or in dynamic environments, and product component knowledge is preferred in more stable phases of incremental or modular innovation. Therefore, maintaining access to a broader knowledge boundary than task boundary may help firms overcome strategic issues arising from architectural innovation. In other words, successful product development may necessitate a non-correspondence of knowledge boundaries.

A similar line of enquiry is followed by Galvin and Rice (2008:5) who suggest that firm's need both architectural and product component level knowledge combined in a kind of 'information structure' that acts as an "*agglomeration of technical knowledge that includes both component-level and architectural-level knowledge*". Information structures are bundles of architectural and product component knowledge that corresponds to product functions and acts as a unit of analysis to help firms to decide which bundles of architectural and product component knowledge to protect within firm boundaries, and which can be diffused via open standards (Galvin & Rice, 2008:6). In their study of the mobile telephone industry, the authors argue that the product architecture consists of three information structures: (1) air interface protocol, (2) connectivity and (3) internal workings. For example, the internal workings information structure must contain architectural knowledge of how it relates to the other information structures, but also product component knowledge related to the menu structure, data storage and circuitry.

These contributions have important implications for the notion of architectural correspondence and its theoretical assertion that knowledge boundaries will correspond to product component and task boundaries. If firms require both architectural and product component knowledge in the form of information structures that correspond to product functions, it seems logical to assume that product component and task boundaries may correspond, but that knowledge boundaries will not.

Unsurprisingly the largest body of empirical work on the non-correspondence of knowledge boundaries comes from scholars who have challenged the assumption that product component and task partitioning is a good map for the division of knowledge (Brusoni & Prencipe, 2001; Brusoni, *et al.*, 2001; Sosa, *et al.*, 2004; Zirpoli & Becker, 2007; Zirpoli & Camuffo, 2009). Zirpoli and Camuffo (2009), for instance, studied air-conditioning systems in the motor vehicle industry and noted that product component and task boundaries correspond. However, they noted a knowledge-task boundary overlap in both manufacturer/assembler firms and suppliers, and that firms' strategy (*vis a vis* value creation), knowledge and capability determined task allocation more so than product component boundaries. Zirpoli and Becker (2007) also argue that product component boundaries are often an inadequate map for task and knowledge partitioning, and highlight a number of reasons why this thesis may have limits, (1) product architectures are rarely fully decomposable and, as such, product component knowledge needs to be retained and assimilated by a manufacturer/assembler firm in order to minimise any remaining product component interdependencies, (2) dynamic product market environments often necessitate new partitioning schemes which may also require a manufacturer/assembler firm to retain product component level knowledge, and (3) technological newness raises uncertainty which also may necessitate a manufacturer/assembler firm to keep product component knowledge within firm boundaries. Cabigiosu, Zirpoli and Camuffo (2013:672) offer a slightly different perspective. In their study of air-conditioning systems they found that where the manufacturer/assembler firm held high-levels of product component knowledge, it was able to successfully partition the product and task boundaries. However, where the manufacturer/assembler firm had insufficient product component know-how, it nurtured co-development practices with external suppliers and engaged in high-levels of ex-ante and ex-post information-sharing to acquire product component knowledge.

Brusoni and Prencipe (2001) and Brusoni *et al.*, (2001) studied the aircraft engine and chemical engineering industries and found that it can be difficult for a group of firms to develop a modular product architecture without a ‘systems integrator’ or ‘lead firm’ who can orchestrate the involvement and activities of others. Accordingly, modular product design is associated with the correspondence of product component, task and firm boundaries, but not with the partitioning of knowledge boundaries as “*systems integrators know more than they do*” (Brusoni & Prencipe, 2001:202). In other words, the “*decisions to outsource technological knowledge differ from decisions to outsource component production....*” (Brusoni, *et al.*, 2001:608).

The non-correspondence of knowledge and task boundaries, according to the authors, arises from the fact that underlying product components evolve and change at different and uneven speeds which creates ‘imbalances’ in the technical architecture that necessitates keeping product component knowledge in-house, even when the product component design and/or production is outsourced. Systems integration is conceptualised as a kind of ‘third-way’ governance mode between vertical integration and market contracting that may help firms solve the apparent knowledge puzzle. Systems integrators are conceptualised as loosely-coupled firm structures that lead and coordinate a group of external firms, while maintaining in-house architectural knowledge, and a systems integration core capability in order to orchestrate the activity of others. However, the notion of a systems integrator is more than simply an ‘orchestrator’, it is a firm who has a core expertise in knowledge integration in order to improve existing product architectures or create new architectural types (Brusoni, *et al.*, 2001:957).

Systems integration capabilities, however, may not be appropriate in every product market context. For example, Ceci, Masini and Precipe (2014:4) argue that systems integration capabilities are contingent upon the particular product design solution and distinguish between design modes optimised for (1) breadth, (2) customisation, or (3) modularity. In the case of product design modes characterised by breadth and customisation, systems integration capabilities were found to be complementary, whereas, under conditions of product modularity this seems to defy the complementary effect.

The main argument is that varying breadth and/or the level of customisation is a dynamic, ex-post integration mechanism and in such circumstances dynamic systems integration capabilities are of value. Whereas, product modularity is conceptualised as an ex-ante integration mechanism where firms require ex-ante systems integration capabilities to define the modular product architecture, but that then the ex-ante modular partitions make further dynamic systems integration capabilities redundant.

Many of these contingent studies are situated in industries characterised by a high degree of product complexity (ie, motor vehicles, aircraft engines) and complexity may hinder the degree to which a successful product component, task, and knowledge partitioning can be accomplished. Luo, Whitney, Baldwin and Magee (2011:28), also argue that modular partitioning is difficult to sustain in complex environments. However, even when the product architecture is characterised as ‘simple’, modular partitioning may also be difficult to sustain. For instance, when underlying product components are subject to fast rates of technological change or product markets are characterised as ‘dynamic’. Langlois (2002:25) observed that modularisation can often be more difficult in dynamic product market settings because “*the problem of defining boundaries of encapsulation becomes even more challenging in a dynamic setting*”; in other words, in dynamic contexts modular partitioning is limited.

Furlan, *et al.*, (2014) examined the air-conditioning industry and noted that non-correspondence can occur in the presence of dynamic product component level change. Drawing upon a range of theoretical perspectives, the authors suggest that product component change increases information asymmetry and asset specificity and this leads to more frequent information-sharing, regardless of the degree of product component modularity. In addition, firms retained access to product component knowledge in order to keep up with fast rate of technological change.

These contributions suggest that the extent to which product component, task and knowledge boundaries are partitioned is a strategic choice; correspondence may be beneficial in stable and/or modular product market settings (Sanchez & Mahoney, 1996) or detrimental in more complex or dynamic settings (Chesborough & Kusunoki, 2001).

In complex or dynamic settings, MacDuffie (2008:42), for example, concludes that firms developing complex products should “*experiment and move activities back and forth across organisational boundaries*” and this helps firms “*...move towards greater product modularity where it increases positive outcomes and avoids negative outcomes – and otherwise to embrace persistent product integrity*”. In other words, the key puzzle for firms is how to “*...reconcile specialisation in the production domain while remaining integrated in the knowledge domain*” (Brusoni, *et al.*, 2001:609).

In sum, in the case of simple product architectures with stable product component technologies, the extant literature suggests that ex-ante partitioning of product component, task, knowledge and firm boundaries in order to benefit from lower-levels of ex-post information-sharing may be an appropriate governance mode. In other words, architectural correspondence may be beneficial. However, on the other hand, when product architectures are complex, or subject to fast rates of product component change, firms maintain a broader knowledge boundary than task boundary and maintain high-levels of information-sharing ex-post.

Architectural correspondence has an intuitive appeal, however recent contingent perspectives have called into question the construct. Many of the instances of non-correspondence relate to knowledge boundaries, under conditions of product complexity and/or fast rates of product component change (even where the product architecture is stable). However, a weakness of many prior studies is the static and cross-sectional nature of their mode of enquiry which may lead to inconsistent results. Moreover, with many product markets characterised as ‘complex’ or ‘dynamic’, the strong support in the extant literature for correspondence appears perhaps rather puzzling.

However, few empirical studies have examined correspondence and non-correspondence at the same time. What contingencies determine correspondence and non-correspondence? This research study aims to identify further contingencies that may mist the mirror and contribute to the literature by moving beyond the intuitive appeal that ‘complexity’ mists the mirror. Moreover, few, if any, empirical studies have examined correspondence longitudinally or retrospectively across time. Baldwin (2008:180), however, hints at such a phenomenon by observing that *“the modular structure of the task network at a particular point in time results from the interplay of firms’ strategies, their knowledge and the physical constraints of specific technologies”*. Furlan, *et al.*, (2014:790) have also hinted at how correspondence may change over time dependent upon the rate of technological change, however highlight that *“one aspect that most empirical studies do not address is the co-evolution of product and industry architectures”*. In addition, Furlan, *et al.*, (2014:804) contend that future research *“will require the investigation of other, more dynamic industries characterised by radical and architectural changes, the adoption of a longitudinal perspective, and innovative measures of modularity”*. Put succinctly, Fixson and Park (2008:1298) suggest that correspondence between changes in product architecture and changes in industry architecture *“demands a detailed understanding of the change processes themselves that occur within the technical domain and industry domain. This, in turn, requires a longitudinal research design”*.

This research study responds to those calls.

3. Research Methodology, Design and Methods

This chapter provides a detailed account of the research design and processes that underpin the research study. I begin with a summary of my subtle realist philosophical position, a discussion of the qualitative research methodology and the methods enacted. Of particular importance to the discussion is the framing of a ‘subtle realist’ epistemology combined with the use of matrix and thematic analysis (template analysis) *across time* which underpins the data analysis phase of this research study. The rationale of employing interviews as the data collection instrument is explained, as well as the criteria used to enhance the quality of the research study. I end the chapter with a summary of how I considered researcher positionality and a section on ethical considerations.

3.1 Research question and objectives

The two over-arching research questions are:

- RQ1: What supply-side processes are associated with the co-evolution of architectures?
- RQ2: Are the design characteristics of product architecture associated with the design characteristics of firm and/or industry architecture?

To answer the research questions, the following process was adopted:

1. To review and synthesise a range of extant theoretical and empirical literatures across the domains of industrial economics, innovation and strategic management
2. To design and apply a research methodology and associated methods to explore what supply-side processes are associated with the co-evolution of architectures, and whether shifts in the design characteristics of product architecture are associated with shifts in the design characteristics of firm and/or industry architecture
3. To synthesise and position the findings within the existing literature in order to advance the current body of knowledge

A visual representation of the study follows:

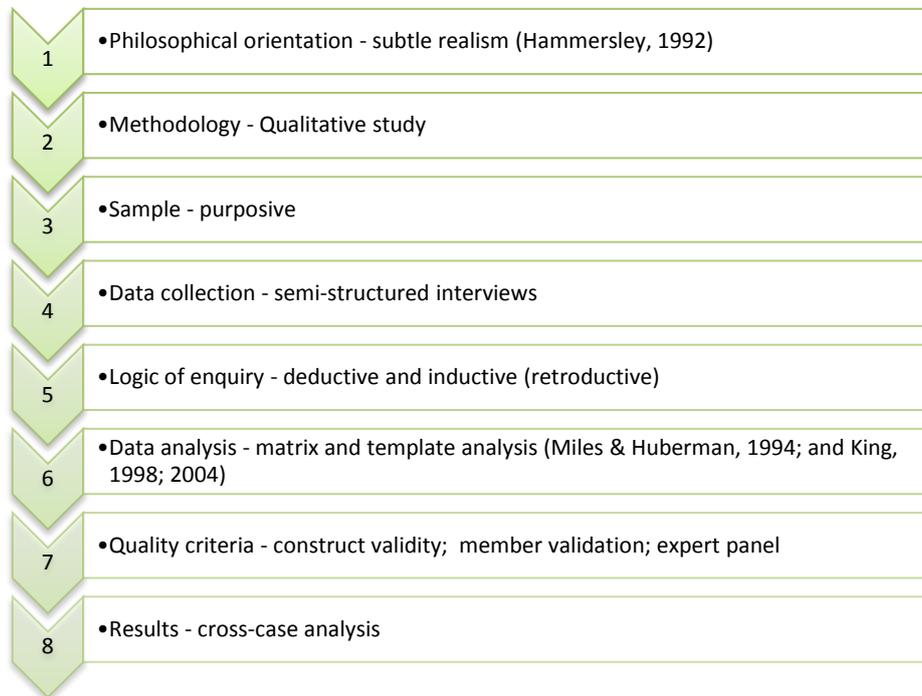


Figure 7: Outline of the research methodology

3.2. Philosophical approach

It is of crucial importance when undertaking research to make explicit and justify the philosophical stance which underpins what it means to be in the world (ontology) and how you know what you know (epistemology).

3.2.1 Ontological and epistemological perspectives

Ontology has been described as the claims a particular type of social enquiry makes about the nature of social reality (Blakie, 1993), whereas Crotty (1998) argues that ontology is the study of ‘being’ and Easterby-Smith, Thorpe and Jackson (2012) state that ontology is the philosophical assumptions about the nature of reality.

Ontological perspectives are usually described on a continuum from realist to relativist. A realist ontology assumes that there is a 'real' world out there that exists independently of us, and that the real world has real objects and structures. A realist ontological frame "*seeks to generate knowledge that captures and reflects as truthfully as possible something that is happening in the real world*" (Willig, 2013:15). It is often accompanied by a set of underpinning assumptions such as the existence of a set of social processes which exist independently of us, and that there is a relationship between our view of the world and the world 'out there'. Often, realist ontology is associated with the 'natural' sciences and quantitative research methodologies that assume that cause-and-effect relationships can be revealed by scientific data collection and analysis.

In contrast, a relativist ontological perspective rejects the assumption that reality is external and objective and posits that reality is socially-constructed and meanings are assigned through our own specific social frames of references, open to a range of interpretations (Easterby-Smith, *et al.*, 2012). 'Truth', therefore, is described as neither purely objective nor subjective (Crotty, 1998) but is constructed through social interactions, the assumption being that there are multiple realities and that all knowledge is context-dependent (Madill, Jordan & Shirley, 2000). In fact, Collins (1983), quoted in Easterby-Smith, *et al.*, (2012:20), says that "*what counts for the truth can vary from place to place and from time to time*". Often, qualitative research has its foundations grounded in a relativist ontological perspective (King & Horrocks, 2010).

In ontology, the issue of the existence of a 'reality' and whether we can access it unmediated is highly contested. Whether there is one 'reality' that can be uncovered or revealed, or multiple socially-constructed realities, and how these beliefs align with different ontological perspectives, continues to be open to debate.

Epistemology is often characterised as “*how we know what we know*” and as a “*means of establishing what counts as knowledge*” (King & Horrocks, 2010:8). Epistemological concerns are central to any methodological approach and its ‘epistemological integrity’ (Marshall & Rossman, 2006). Two main approaches to epistemology are characterised in the literature. One main approach is positivism. King and Horrocks (2010:11-12) define positivism (or ‘naturalism’) as an approach which “*assumes an ontological view that human beings are part of nature and can be studied in the same way as other objects in the physical world*”. The underlying aim of research in the positivist tradition is to generate objective knowledge, a correspondence between ‘truth’ and ‘reality’. In other words, a positivist approach aims to develop patterns, laws or principles that can be rationalized and generalized. Positivism has been described as nomothetic (Easton, 2010; King & Horrocks, 2010) which ultimately leads to positivists believing they can make causal claims. Research in a positivist tradition is therefore often quantitative and works best with large data sets as it tries to uncover ‘reality’ through statistical techniques and replication. However, there are variations within the positivist perspective. For example, a post-positivist may argue that an absolute reality cannot be fully understood and may only be approximated (ie, Denzin & Lincoln, 2000).

Positivism and post-positivism are in contrast to an epistemological tradition of interpretivism, which King and Horrocks (2010:12) suggest is a “*bridge between representations of particular phenomena – what we claim is occurring – and the actual world out there*”. The interpretivist tradition assumes that objectivity is illusory as knowledge always requires interpretation that is influenced by the social world. In general terms, interpretivists deny that there is an unmediated access to what is real and reject the idea of causality. Interpretivism is often associated with the collection of qualitative data. The aim of qualitative research is (often) not to try to uncover patterns, laws or principles, but to study beliefs, behaviour or attitudes which may be influenced by historical experiences or a range of social and personal factors (Easterby-Smith, *et al.*, 2012) that are in a constant state of flux. This means that knowledge is always temporary and contextual (Madill, *et al.*, 2000).

Between these two polar epistemological opposites, a combination of both paradigms may be argued by some scholars on a pragmatic basis (ie, Duncan & Nicol, 2004; Murphy, Dingwall, Greatbach, Parker, & Watson, 1998). One such pragmatic approach is subtle realism (Hammersley, 1992) to which I now turn.

3.2.2 Subtle realist commitments

The polarization of philosophical perspectives, and ultimately of methodology, has been subjected to extensive critique (ie, Kirk & Miller, 1986; Murphy, *et al.*, 1998). Proposed by ethnographer Martyn Hammersley (1992), a subtle realist approach is an example of a so-called ‘third-way’ philosophical stance (Murphy, *et al.*, 1998). Hammersley (1993) argues that adopting a naïve realist or naïve interpretivist approach is problematic. On the one hand, there is no basis which is beyond all reasonable doubt to check the correspondence between an account of reality and reality itself – the ‘criterion problem’, and on the other hand, Hammersley (1992) begs the question that if the job of social science is to produce multiple accounts on the basis of a similar research experience, what value can it have? Hammersley (1992) adopts a position that neither naïve realism nor naïve interpretivism offer a sound philosophical basis for the social sciences.

This research study adopts a ‘subtle realist’ perspective as proposed by Hammersley (1992). Subtle realism rejects the ‘naïve realist’ perspective that maintains a correspondence theory of truth. Subtle realism concedes that it is impossible to have certainty about knowledge claims as researchers inhabit the social and cultural world (Hammersley & Atkinson, 1995), and, as such, researchers can never be absolutely certain about, or stand apart from, the validity of any knowledge claim. However, Hammersley (1992) argues that it is possible to be *reasonably confident* about them, and knowledge claims will usually be based around judgements of their plausibility or credibility. Subtle realism argues that there are knowable, independent phenomena, but it departs from naïve realism by acknowledging that knowledge relies on social and cultural assumptions. It is always possible that cultural or social factors may play a role in either leading us towards or away from the truth (Hammersley, 1992). Subtle realists, then, aim to represent reality, rather than reproduce it (Hammersley, 1992). This approach, Hammersley argues, allows us to accommodate some elements of an interpretivist approach, without abandoning a commitment to a knowable truth.

Following on from this distinction, subtle realism suggests that reality can be represented from a range of perspectives, which may be treated as either true or false. Murphy, *et al.*, (1998:69) point out that within a subtle realist approach there is the possibility of “*multiple, non-competing valid descriptions and explanations of the same phenomena, however, it excludes the possibility of multiple, competing valid descriptions or explanations of the same phenomena*”. Subtle realism, then, offers a middle position between the polar extremes of naïve realism and naïve interpretivism, but importantly “*subtle realism is distinct...in its rejection of the notion that knowledge must be defined as beliefs whose validity is known with certainty*” (Hammersley, 1992:52). Subtle realism accepts that phenomena can be represented from multiple perspectives, which is not the same as saying that there are multiple realities. Murphy, *et al.*, (1998:70) conclude that subtle realism is “*an alternative to the ontological and epistemological dichotomies*”. Although Hammersley (1992) was writing in the context of ethnography, Seale (2002:107) points out that it is “*capable of application to others forms of qualitative research*”, and Hammersley (1992:53) also suggests that subtle realism is a useful perspective from which to elicit accounts to “*provide us with information about events that we could not ourselves witness (for example that happened in the past or in settings to which we do not have access)*”. In terms of its application in research settings, Murphy, *et al.*, (1998), Duncan & Nicol (2004), and Lewis (2015) commented upon the value and usefulness of the approach in health care and social research.

A subtle realist perspective is, of course, not without its critics. In general terms, the combination of epistemological perspectives is highlighted as unsustainable by scholars such as Hill Bailey (1997), quoted in Duncan and Nicol (2004). Furthermore, Seale (1999) argues that subtle realism has no true ontological basis, an observation elaborated upon by Banfield (2004). Banfield argues that subtle realist is ‘ontologically-shy’ because it fails to explain *why* we understand certain descriptions or representations as more valid than others and becomes “*a kind of smorgasboard approach to the production of theory*”. Moreover, Banfield describes subtle realism as a flat ontology which collapses ontology and epistemology to the extent that all knowledge claims are treated equally. Ultimately, however, Banfield proposes a critical realist perspective in order to overcome, what he suggest as, subtle realism collapsing positivism and interpretivism and structure and agency, and ignorance of a Marxist view of history.

Banfield also raises concerns with the problem of relevance. Hammersley's idea of an adequate explanation of a phenomenon is one that "*depends on the particular purposes for which and context in which the explanation is being developed*" (Hammersley, 1992:39), which according to Banfield, offers nothing but "*ambiguity with endless possibilities for the epistemological shopper who is free to select abstractions and generate explanations of the social world to fit fashionable or practical purposes*" (Banfield, 2004:57). In other words, according to Banfield, subtle realism is 'conflationist' and vague.

Despite these criticisms, I believe that my subtle realist commitments are appropriate for this research study. For one, the research study has been structured so as to illicit knowledge claims that we can be reasonably confident about. As such, I believe that it is possible to "*combine a commitment to the social construction of reality with a concern for truth as a regulatory ideal, through studying the different constructions people make of reality, without accepting that particular beliefs are true*" (Murphy, *et al.*, 1998:69). Moreover, the aim of the research study is neither to adopt a critical Marxist perspective nor to generate 'fashionable' explanations, but rather than to generate multiple, valid descriptions and explanations of events and processes across time.

There are a number of reasons for adopting a subtle realist perspective in this research study. Firstly, one of the fundamental implications of subtle realism is that it is well-suited as a potential companion for research that focuses on phenomena in context using techniques such as ethnography or semi-structured interviews (Hammersley, 1992). Therefore, a subtle realist approach is well-suited to studying complex relationships and heterogeneous knowledge claims. Secondly, a subtle realist approach is well-suited to a longitudinal or retrospective research study across time, as the history of events lead to the possibility of heterogeneous knowledge claims. Thirdly, I intend the research outputs to be relevant to practitioners and policy-makers alike. In my opinion, practitioners and policy-makers are unlikely to consider an interpretivist research study as credible due to its ontological and epistemological commitments to competing multiple realities.

Fourthly, I have had a prior professional involvement in the research setting which may exert a level of unintentional influence on the research; therefore, a naïve realist perspective associated with the notion of objectivity would be problematic (a discussion of researcher positionality is discussed in Section 3.7). Overall, I have sought to achieve ‘epistemological integrity’ (Marshall & Rossman, 2006) throughout the research design.

Section 3.4 illuminates my decisions in terms of sampling, data collection, data analysis methods, quality criteria and how they resonate with a subtle realist perspective. Firstly, however, I turn to the logic of enquiry.

3.3 The logic of enquiry – deduction and induction (retroduction)

The logic of enquiry for any research study is often concerned with deduction or induction, or in other words the place that theory takes in the research design. Deduction begins with a theory, operationalises concepts and is concerned with gathering data to test the theory. This type of enquiry is most closely aligned to quantitative methodologies, with the use of statistical analysis techniques and the hypothetico-deductive method (Murphy, *et al.*, 1998). In contrast, in an inductive study a researcher rejects the imposition of *a priori* theory upon the data, and posits that hypothesis should be built from the ground up – as is often said it should ‘emerge’ from the data (ie, Glaser & Strauss, 1967). One of the primary distinctions between inductive and deductive enquiry centres on theory generation versus theory testing, since induction is grounded in the data, whereas the former seeks to examine observed behaviour with reference to existing theory. As with many concepts with dichotomous positions, the inductive-deductive continuum represents an over-simplification as much research represents characteristics of both induction and deduction (Hammersley, 1992; Strauss & Corbin, 2008).

The notion that qualitative research is often closely associated with an inductive enquiry method, where the researcher brings to the table no pre-conceived ideas, theories or frameworks, has also been challenged (ie, Miles & Huberman, 1994). They argue that any researcher always has some preconceived orientating ideas, as without an *a priori* orientation it is almost impossible to embark upon research. Consequently, it is common for some initial theorising to be done prior to the data collection, and consequently very few studies are purely inductive in nature. Miles and Huberman (1994) suggest that, in many instances, a more 'structured' or 'tight' research design is preferable as it helps bring clarity and focus to a research study, but there is always the present risk of over-imposing *a priori* concerns on the research setting.

The logic of enquiry process adopted in this research study closely follows the recommended approach set out by Yin (2014). In the first instance, a draft set of research questions were developed based upon the field of enquiry and professional knowledge of the research setting. The draft research questions were then refined in light of an extensive literature review. From the literature review and gap analysis, I then deductively created a product architecture typology that served to structure the data collection and data analysis phases. Miles and Huberman (1994) also argue that if the aim of research is to describe or analyse a pattern of relationships, then it is imperative that there exists a set of analytical categories, deductively generated in order to guide what aspects of the data to attend to. From there, during data analysis, I used an iterative process of moving back and forth between theory and data, because, according to Emerson (1983) the movement backwards and forwards between theory and data is characteristic of qualitative research and a retroductive logic of enquiry (Easton, 2010). This research study has run along these lines.

How the research design has been operationalised is now discussed in section 3.4.

3.4 Research Design

3.4.1 Sampling approach

In qualitative studies, King and Horrocks (2010:29) suggest that respondents are usually those able to “*throw some light on meaningful differences in experience*” or who “*represent a variety of positions in relation to the research topic*”. This kind of sampling is often referred to as ‘purposive’ sampling. In this research study purposive sampling was used to initially identify suitable respondents.

The total purposive sample was generated in two parts.

First, (1) I generated a purposive sample from my own professional contacts via social media platforms, such as Linked-In. Twenty-four potential respondents were contacted by email and sixteen agreed to participate in the research study. Eight did not respond. However, I was aware that this type of purposive sampling could generate some unwanted bias in the sample. For instance, the age, professional experience and firm-membership profile of the sample may be similar to my own professional profile when employed in the industry. For example, the sixteen respondents who replied had a professional experience ‘skew’: in other words, the longevity of their professional experience roughly matched my own, and covered the period from the very late-90’s to 2014 (in other words, I generated a sample that ‘looked like me’). As I wished to explore the retrospective longitudinal period between 1984 and 2014, this presented a problem. In addition, the sample contained respondents that were all known to me in a professional capacity which could limit the study’s ‘plausability’ and ‘credibility’. As such, (2) the sixteen respondents were all invited to recommend further potential respondents.

The sample suggested a further thirty-six potential respondents, with differing ages, firm backgrounds and professional experience. As the objective was to acquire additional respondents with *continuous* professional experience dating back to 1980s, I decided to only approach potential respondents with continuous professional experience dating back to 1984 (or even earlier). Twenty-six of the thirty-six potential candidates were identified as potential respondents (via desk research using LinkedIn and Google), and fifteen of the twenty-six agreed to participate in the research study, thereby making a total purposive sample of thirty-one. This is shown in Table 2.

Purposive sample	Contacted	Accepted
Phase 1	24	16
Phase 2	26	15
	50	31

Table 2: The construction of the total sample

There are advantages and disadvantages with the sampling approach outlined. King and Horrocks (2010) argue that respondents may suggest other respondents who share a similar viewpoint. However, I believe that this risk was minimal, as I selected from within the additional twenty-six potential respondents to identify those who had dissimilar characteristics.

The screening process for part one of the sample was, therefore, as follows:

- employed in the UK personal pensions product market for *any period* between 1984 and 2014
- employed in a senior product development or strategic management role

To ensure ‘dissimilar’ characteristics in the remainder of the sample, the screening process for the second part was tightened as follows:

- *continuously* employed in the UK personal pensions product market between 1984 and 2014
- employed in a senior product development or strategic management role

The total purposive sample of thirty-one was considered appropriate for this research study for a number of reasons. First, initially I had estimated a total sample of between twenty-five and forty as being adequate to illicit a number of different explanations across multiple time periods. As the interviews and creation of the total sample ran alongside each other, I had thirty-one interviews scheduled by the time I interviewed the twenty-fifth respondent. No new alternative explanations of the phenomenon were recorded beyond interview_25 and, therefore, I decided not to extend the total sample size beyond thirty-one. Second, I believe that the breadth of respondents in terms of professional experience and firm background enhances plausibility, credibility and internal validity of the research study (Yin, 2014). Third, the *total population* of potential respondents that meet the screening criteria is unlikely to exceed seventy or thereabouts. Therefore, the purposive sample of thirty-one (with nineteen having a 30-year professional track record), represents a sizeable proportion of the total population available. To illustrate this, a LinkedIn search of the UK personal pensions product market, filtered with ‘product development’ yields one-hundred and seventy-three possible key personnel. However, many of the matches will not have continuous professional experience dating back to the 1980s.

In Table 3, I show how the purposive sampling approach was successful in locating suitable respondents. In the end, nineteen of the thirty-one respondents had continuous professional experience dating back to 1984 or earlier.

Purposive sample	Professional experience commencing in				Total
	1980s	1990s	Noughties	2010 onwards	
Phase 1	4	8	4	0	16
Phase 2	15	0	0	0	15
Total	19	8	4	0	31

Table 3: The length of continuous professional experience of interview respondents

In addition, the purposive sample was spread across respondents employed by six different firms and four job role types and interviews were conducted between 2 October 2014 and 16 December 2014. Interview respondent backgrounds are shown in appendix 1.

A pilot study was conducted between 2 October 2014 and 10 October 2014 in order to facilitate an early draft of data analysis (see section 3.5.2.1). To construct the pilot sample, I decided to simply use the first three interviews that were scheduled and this generated a pilot sample as shown in Table 4:

Respondent Reference	Organisation Reference	Job Title	Experience commenced
Interviewee_1	Firm_1	Head of Product Development	1980s
Interviewee_2	Firm_2	Head of Product Development	Noughties
Interviewee_3	Firm_3	Director of Product Development	1990s

Table 4: Pilot sample

3.4.2 Interview methods

3.4.2.1 Recruitment process

Potential respondents were informed of the research study, its purpose and benefits with an invitation email. The covering email highlighted the research purpose and importance of the study and stressed the confidentiality and anonymity of participating individuals and their respective organisations. Participants were promised an executive summary of the final report. Personal contact details were given for further enquiries on the research, and the ethical procedures were stated and agreed. A similar covering email was used for all subsequent email reminders targeting non-responders. The process adopted is shown in Table 5.

Communication	Purpose	When	Responses
Email notification of study to purposive sample (phase 1)	To inform potential respondents of the study, its purpose and benefits	30/9/14	13
1 st Reminder to purposive sample (phase 1)	To inform potential respondents of the study, its purpose and benefits	12/10/14	2
2 nd Reminder to purposive sample (phase 1)	To inform potential respondents of the study, its purpose and benefits	1/11/14	1
Email notifications to purposive sample (phase 2)	To inform potential respondents of the study, its purpose and benefits	Between 5/10/14 and 1/12/15	15

Table 5: Respondent recruitment process

3.4.2.2 Interview schedule

Qualitative interviewers often use an interview guide instead of a set of structured questions (King & Horrocks, 2010). The interview guide can be based upon a number of inputs and extant literature or theoretical frameworks are such common inputs. This research study closely followed the advice of King and Horrocks (2010) as I decided to use a minimalist interview schedule that remained flexible enough to allow respondents to share their own experiences and different perspectives, which opens up the possibility that the interaction could lead off in unexpected, but interesting, directions. The interview schedule is shown as Appendix 2.

As suggested by King and Horrocks (2010), I made significant attempts to reflexively consider how my own interview practice developed over the course of the research study. For example, the data collection and data analysis phases occurred in parallel and this meant that in later interviews I was able to orientate the follow-up questions to more deeply explore explanations and experiences that offered a different or new perspective to what had come before in earlier interviews. Changing the follow-up questions posed across interviews, and remaining aware of your own interview practice, is an advised approach as it allows respondents to share relevant aspects of their experience (Easterby-Smith, *et al.*, 2012; King & Horrocks, 2010) and remains sympathetic to the subtle realist frame of accounting for different perspectives.

3.4.3 Data collection

All 31 interviews were 1 to 1 in-depth interviews, and audio recorded on an i-Phone audio recorder. I transcribed interviews 1-3 (pilot interviews). Interviews 4-6 were professionally-transcribed. However, I immediately felt that not being immersed in the transcription process was an issue (Langdrige, 2004) and the accuracy of the transcription was less than anticipated due to the use of industry-lingo. As a consequence, I transcribed interviews 7-31.

The audio files were transcribed in full, word for word (verbatim), and were tidied-up as part of the transcription process. Some interview features such as pauses, hesitation and coughing were not transcribed as these are not relevant to this research study. The average duration of the interviews was sixty minutes, with the shortest being forty-two minutes and the longest being eighty-four minutes. The interviews took place at a location specified by the respondent – twenty-one took place in a private meeting room or office in the respondents' place of work, and ten took place in a quiet area of a public environment such as a café or restaurant. However, recording quality for transcription was temporarily lost on one or two occasions due to background noise, although I don't believe this has had a material effect on the data. It is worth noting that all respondents were keen to be interviewed and seemed to enjoy the experience, most offering more time than they had initially indicated.

3.4.3.1 Interview structure part 1 – establishing a product timeline

The structure of the interview was divided into two distinct parts, part 1 and part 2.

In part 1, my aim was to establish construct validity (Yin, 2014) for the product architecture typology and to invite respondents to create a product architecture timeline.

However, before I progress to describe in detail the process I followed, I will briefly review how other scholars have approached the issue of product architecture construct validity. The approach adopted in the empirical literature differs, of course, to the degree to which the authors embed a positivist or interpretivist paradigm. According to Fixson (2005), however, constructs of product modularity have often encompassed a range of mathematical models, engineering-based approaches, or conceptual perspectives.

In general, mathematical models aim to simplify the product architecture into an overall 'average assessment' of its degree of product modularity. For example, the product architecture is often taken as a given and it, along with its main associative variables of interest, are conceptualised as a range of dependent and independent variables (ie, Argyres & Bigelow, 2007; Park & Ro, 2013). Other scholars have used a proxy measure for 'product architecture', such as design interfaces strength (Sosa, *et al.*, 2004).

Another set of measures come from an engineering background. These are often constructed as a set of indices that aim to compare product architectures along various dimensions of interest, often based on qualitative inputs. Fixson (2005:350), for example, cites a number of empirical studies that have developed indices based on degrees of coupling and component compatibility. In addition, Cabigiosu and Camuffo (2012) and Furlan, *et al.*, (2014), develop indices that measure product component modularity based upon the relationship between product functions and interfaces and Mikkola (2003, 2006) develops indices based on the number and type of product components, the degree of coupling and a substitutability factor. Each of these types of indices aims to measure the overall degree of product modularity.

An alternative approach to indices is to present the information in the form of a design structure matrix (DSM) based upon the work of Steward (1981), whereby the rows and columns of the matrix represent a functional interaction between two hierarchical levels of interest. For example, Fixson (2005) and Fixson & Park (2008) develop a DSM based upon function-component allocation schemes and their relationship to interface standardisation, strength and reversibility. The DSM approach has also been used to compare the degree of correspondence between architectural layers by collecting data at different layers of the hierarchy, plotting in multiple DSMs, and contrasting the results. For example, Colfer and Baldwin (2010) highlight the use of comparing a product DSM and a labour-knowledge DSM and suggest that this is an ‘ideal test’ of architectural correspondence.

Empirically, DSMs have been used by in several architectural correspondence or ‘mirroring’ studies, such as McCormack, *et al.*, (2012) in the software industry, Sosa, *et al.*, (2004) in the aircraft engine industry, Gomes and Joglekar (2008) in software development, and Furlan, *et al.*, (2014) in air-conditioning systems. While DSMs are a popular approach in the extant literature, they often rely upon qualitative inputs to classify the type and number of product components, as well as the degree to which a product component is substitutable, or whether an interface is standardised and open (Fixson & Park, 2008). Given the retrospective nature of this research study, a DSM method was not considered appropriate due to access to reliable input information over the entire period.

A third set of measures, called ‘conceptual’ measures, has been suggested by Fixson (2005) and have tended to focus at the level of either the product architecture, product components, or interfaces. Conceptual measures tend to appear in qualitative studies, and take two distinct forms: (1) the product architecture is either ‘assumed’; or (2) the product architecture is assessed by reference to secondary data sources and subsequently member-checked (Horrocks & King, 2010) prior to qualitative data collection. For example, Zirpoli and Camuffo (2009) interviewed twelve managers across two comparative projects following a review of company data. The product architecture was assumed to be modular, although this assumption was corroborated with some of the respondents prior to interviewing. Whereas Henderson and Clark (1990) interviewed over one-hundred managers in the photolithographic industry and the authors constructed a product and industry history, which was then member-checked ex-post.

In their ground-breaking work on systems integration in the aircraft and chemical engineering industries, Brusoni and Prencipe (2001) and Brusoni, *et al.*, (2001) developed a mixed method approach examining secondary sources such as product-level and patent data in order to create an ‘average assessment’ of a product development project. Following member-checking, interviews then were held with forty-five key personnel across two industries. Cabigiosu, Zirpoli and Camuffo (2013) took a different approach by asking key personnel to rank the degree of modularity in a given set of product components. This ranking approach was also adopted by Furlan, *et al.*, (2014) and the authors then used the ranking information to create a DSM.

Other conceptual approaches have been used. For example, Shibata, *et al.*, (2005) examined secondary sources in order to write a product and technical history of the numeric-controllers product. This was then member-checked prior to the qualitative data collection phase. MacDuffie (2013), on the other hand, took the opposite approach, and interviewed a number of managers across a number of firms in the motor vehicle industry to develop ideas of how product modularity had developed.

Measures of industry architecture are also wide and varied. For example, in quantitative studies scholars have often used established definitions in the literature to define survey constructs (Schilling & Steensma, 2001; 2002) and some scholars have examined secondary data sources (Brusoni & Prencipe, 2001; Brusoni, *et al.*, 2001) in order to characterise the industry structure. In studies such as Jacobides and Hitt (2005), firms in an industry were either characterised as simply ‘vertically integrated’, ‘fully specialised’, or ‘mixed mode’, based upon secondary sources or, in the case of Argyres and Bigelow (2010), industry structure was conceptualised as an independent variable: vertical integration yes or no. Alternatively, Fixson and Park (2008:1303) use a set of indices which they refer to as a ‘vertical integration measure’ based upon firm concentration and market share. In qualitative studies, inductive approaches to characterising industry architecture are more common. For example, Cacciatori and Jacobides (2005) pursued an inductive approach, interviewing sixty key personnel and embellishing it with secondary sources, such as trade journals and industry surveys, and Jacobides (2005) utilised a mixed method approach to inductively determine how the mortgage banking industry had become more specialised.

Retrospective longitudinal studies are also a feature of the extant literature. For example, Zirpoli and Becker (2007) examined the motor vehicle industry over a 10-year period, Cacciatori and Jacobides (2005) examined the building and construction industry from the late 1800s to around the 1970s, Brusoni and Prencipe (2006) and Brusoni, *et al.*, (2001) examined the aircraft and chemical engineering industries over a retrospective 12-year period between 1985 and 1997, Jacobides (2005) and Jacobides and Winter (2005) studied the US mortgage industry between 1981 and 1989 and the Swiss Watch industry between 1980 and 1992, Henderson and Clark (1990) examined the photo-lithiographic industry between 1962 and 1986, Hoetker (2006) examined computer notebooks between 1992 and 1998, MacDuffie (2013) examined a number of firms in the motor vehicle industry between the 1970s and early 2000s, Shibata, *et al.*, (2005) studied numeric controllers between 1962 and 1997, and Fixson and Park (2008) examined the bicycle industry between 1980 and 1990.

With the idea of retrospective longitudinal studies established in the extant literature, I chose to adopt a ‘conceptual’ approach to the product architecture construct, and while slightly different to the approaches discussed above, it shares similar characteristics to the process employed by Brusoni and Prencipe (2001), Brusoni, *et al.*, (2001), and Cabigiosu, *et al.*, (2013). Rather than using secondary sources such as product data or trade journals to develop ‘product architecture’ constructs, I developed a stylised product architecture typology, grounded in the extant literature (as discussed in 2.1.8) and checked its construct validity *ex-ante* with respondents. In line with the assumption that products design organisations (Sanchez & Mahoney, 1996), the product architecture typology provided a deductive structure to the data collection phase and subsequent data analysis phases. However, an inductive logic was used to derive any themes related to the firm and industry level units of analysis.

Part 1 of the interview proceeded as follows: (1) each respondent was invited to read and review the key characteristics of the four stylised product architecture types, derived from the typology. The descriptions that were given to respondents are shown in appendix 3. All respondents confirmed that the theoretical descriptions had considerable resonance with the *actual* product types available in the product market and saw no difficulty in relating the descriptions to the actual products that existed in the product market. Next, (2) respondents were invited to make any amendments or alterations to the descriptions, however none of the respondents deemed this necessary. Next, (3) respondents were asked to create a ‘product timeline’, an idea borrowed from the idea of a ‘lifeline’ that is often used in biographical research, where each respondent depicts key phases or events in their life in a chronological sequence (Brannen & Nilsen, 2011). Lastly, (4) the product timeline was noted and recorded as an aide memoir before proceeding to part 2 of the interview. An example output of a product timeline is shown below in Figure 8:

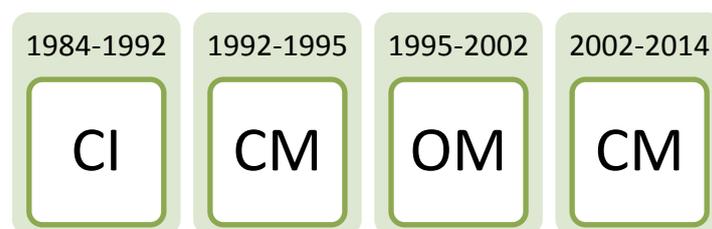


Figure 8: Example ‘product timeline’

The three pilot interviews, however, drew my attention to a few areas where the process described could be improved, and the feedback I received was incorporated into all other twenty-eight interviews that followed. First, one pilot respondent highlighted that I had used a naming convention of product architecture 1, product architecture 2, product architecture 3, and product architecture 4, perhaps leading some respondents to a false assumption that one product architecture type naturally follows another, and that there may be a ‘right’ answer. To remedy this, I revised the naming conventions to closed and integrated (CI), closed and modular (CM), hybrid (H), and open and modular (OM) and I frequently changed their order as they appear in appendix 3 to further minimise this risk. Another pilot respondent suggested that future respondents may perhaps infer that all four of the product architecture types should be used to create the product timeline, or that a product architecture type can only be used once, as opposed to multiple times. To overcome this, I offered a verbal instruction to all respondents prior to beginning the part 1 of the interview, confirming that a product architecture description could be used in the timeline as many or as few times as appropriate, or not at all.

As Yin (2014:154) states, chronological sequences are a major strength of case study style research and are useful, not just as descriptive devices, but as analytical devices that enable the examination of cause and effect relationships because they cannot be “temporally inverted”. Yin also suggests that ‘longitudinal’ studies often explore a theory of interest and how its underlying processes change over time. Yin suggests these can often be across specified time periods and “*reflect the theoretical propositions posed by the study*” noting that “*the desired time intervals would presumably reflect the anticipated stages at which the changes should reveal themselves, and the processes being observed should again reflect the theoretical propositions posed by the case study*” (Yin, 2014:53). According to Yin (2014:45), construct validity can be achieved if the constructs use multiple sources of evidence, establish a chain of evidence and have key informants review the draft. As such, I believe that the process I adopted has underpinned the construct validity of the product architecture typology as a foundation for the inductive phase of the interview, and the discussions that followed in part 2 of the interview.

3.4.3.2 Interview Structure Part 2

The product timeline created in part 1 of the interview served as a structure for part 2 of the interview. The product timeline and interview schedule served as an aide memoir enabling me to ask open questions directed towards particular time periods or events that were meaningful to each respondent.

For example, where a respondent had created a product timeline for, say, 1984 to 2014, I was able to direct questions to discreet blocks of time in the respondents' timeline, such as 'what was going on in this time period? Furthermore, I was able to direct questions to particular transition points from one product architecture type to another, such as 'what led to this change?' However, rather than rigidly sticking to the interview schedule, I often followed up with a number of probing questions to more deeply explore any different explanations that offered a new insight. Yin (2014:154) reminds us the advantage of approach like this is "*history is likely to cover many different types of variables*".

3.5 Data Analysis – Thematic (Template) Analysis and Matrix Approaches

Data analysis methods using textual data are often categorised into those methods that are epistemologically-committed such as IPA, or those that are generic methods (Braun & Clarke, 2006). This research study is concerned with the latter. Thematic analysis is one approach to data analysis which is often framed as a 'realist' method (Braun & Clarke, 2006:78) however it is a method that can be used independent of philosophical stance. Thematic analysis provides a "*common-sense way to refer to patterns in the data that reveal something of interest regarding the research topic at hand*" (King & Horrocks, 2010:159) and involves identifying recurring and distinct 'themes' in the data, usually themes of concern to users of the research. Ultimately, identifying themes means looking for similarities and differences across the entire data set and organising the themes in terms of how they relate to each other. This often results in hierarchal as well as linked relationships with main themes, sub-themes and so on. The thematic analysis process usually has three stages of coding: descriptive, interpretative and over-arching themes (King & Horrocks, 2010). The process of coding involves a number of iterations of reading and revisiting the data and highlighting data extracts either manually or with the use of computer software.

Template analysis is a distinct and flexible type of thematic analysis, first described by Crabtree and Miller (1992), later developed by King (1998, 2004) and as a method has gained traction in management studies, psychology, sociology and healthcare (ie, as discussed in Lewis, 2015; Waring & Wainwright, 2008). It is a method that can be used with many different philosophical positions (King, 2004: 256-7), and Brooks and King (2012:2) suggest that it lends itself to a ‘subtle realist’ approach because “*such an approach can thus make claims as to the validity of a representation arising from the research, while recognising that other perspectives on the phenomenon are possible*”. It differs from other types of thematic analysis in that an initial template, perhaps based on some initial interviews or pilot data, is applied to the data and revised as the research develops until the template captures a better picture of the data. As King and Horrocks (2010:166) point out, the “*reiteration of applying, revising and then reapplying the template continues until the analyst feels it is clear and thorough enough to serve as a basis for building an account of the findings*”. This reiteration between the template and the data, and the data and the template, I believe, lends itself to the iterative deductive-inductive enquiry logic used in this research study.

Another feature of template analysis is the flexibility it offers. Unlike Braun and Clarke’s (2006) blueprint for conducting thematic analysis, there are no limits to the number of levels of hierarchical coding. Importantly for a subtle realist, template analysis does not differentiate between descriptive and interpretative coding, believing that themes are descriptively grounded in the data, but acknowledging that data description cannot be untouched by human interpretation (King & Horrocks, 2010).

Template analysis has been used in this research study for a number of reasons: (1) it is useful “*when the aim is to compare the perspectives of different groups...within a specific context*” (King, 2004:257), (2) it is a method that lends itself to cross-case thematic analysis (King, 2004:258), (3) it lends itself to a subtle realist stance which is interested in valid representations of the ‘reality’ of respondents while acknowledging the interpretive role of the researcher (King, 2004:256), (4) it lends itself to a retroductive moving back and forth between themes and data (King & Horrocks, 2010:166), (5) it permits parallel coding (King, 2004:258); and (6) it is often associated with the use of *a priori* theorising and coding (but this is not a requirement) (King, 2004:257).

3.5.1 Matrix analysis

Template analysis is often used as a method in research studies located between the inductive and deductive approaches to enquiry logic. The *a priori* themes can often relate to important theoretical concepts that inform the research design. In this research study, I decided *not* to use *a priori* codes in the sense envisaged by King (1998; 2004). Rather, I followed an approach suggested by King and Horrocks (2010) in combining an *a priori* matrix and deductive template analysis method.

I created an initial *a priori* matrix and then thematically coded the interview data to cells in the matrix. Therefore, the themes were inductively derived from the interview data but coded to an *a priori* matrix, which served as a structuring device to analyse the data. The use of an *a priori* matrix to structure (rather than code) the data is similar to the matrix approach pioneered by Miles and Huberman (1994) where data is tabulated to different units of analysis so as to facilitate comparisons and connections both between and across different levels of data. The initial *a priori* matrix structure I created reflects the units of analysis of interest to the research questions, and a fourth cell – ‘mirroring’ – aims to capture the degree of correspondence between the architectural layers. This is shown in Table 6.

Mirroring	Product level	Firm-level	Industry level
•Themes	•Themes	•Themes	•Themes

Table 6: Initial *a priori* matrix

King and Horrocks (2010:173) suggest that matrix-based approaches tend to be used in research projects of this size, for example, with thirty hours or more of textual data, whereby the matrix is “*usually defined based on a priori concerns – sometimes these are theoretical but more often pragmatic*”. King and Horrocks (2010), moreover, suggest that a researcher should always allow for the matrix to be modified as a research study progresses.

The initial matrix shown in Table 6 was modified in response to the product timelines created by respondents in part 1 of the interview. To do this, I synthesised the thirty-one individual product timelines into a ‘consensus’ or ‘best fit’. Naturally, across the total sample there was some divergence in the product timelines which emerged due to factors such as individual firm strategies, or respondent ability to re-collect dates with perfect accuracy (this is discussed in section 3.6), however, there was a strong convergence across the total sample. This is shown in appendix 4.

The synthesised product timeline shown in appendix 4 illustrates five sub-periods that were meaningful to respondents which I have labelled as follows:

- “The mid to late-1980s”
- “The early to mid-1990s”
- “The mid 90s to mid-noughties”
- “Mid-noughties to 2012”
- “2012-2014”

Moreover, product architecture types could be mapped onto the 5 sub-periods as follows:

- “The mid to late-1980s” – closed and integrated
- “The early to mid-1990s” – closed and modular
- “The mid 90s to mid-noughties” – hybrid
- “Mid-noughties to 2012” – open and modular
- “2012-2014” – hybrid

Next, I then incorporated the synthesis into the matrix, as shown in Table 7.

Time period	Mid to late-1980s	Early to mid-1990's	Mid-1990s to mid-noughties	Mid noughties to 2012	2012-2014
Product type	Closed and integrated	Closed and Modular	Hybrid	Open and modular	Hybrid
Product themes	Template				
Organisational themes	Template				
Industry themes	Template				
Mirroring	Template				

Table 7: Final matrix

The final matrix illuminates how the product architecture descriptions were mapped to the actual product architecture in the product market at given periods of time.

The method I used combines the advantages of matrix analysis, to structure and tighten the data collection and analysis phases, with an inductive approach to template analysis. The method also allows themes to be coded to different time periods and different units of analysis. In the data analysis phase, I was able to generate multiple templates that can then be compared and contrasted. As King and Horrocks (2010:173) conclude, “*in some studies, the optimum analytical study can involve a combination of matrix and template approaches*”, by carrying out a matrix approach to uncover broad areas of interest, and then followed by a more finer-grained template analysis on elements of rich interest (ie, King, *et al.*, 2007). As described, this research study has run along these lines.

Figure 9 provides a summary of the matrix and template analysis approach used in this research study:

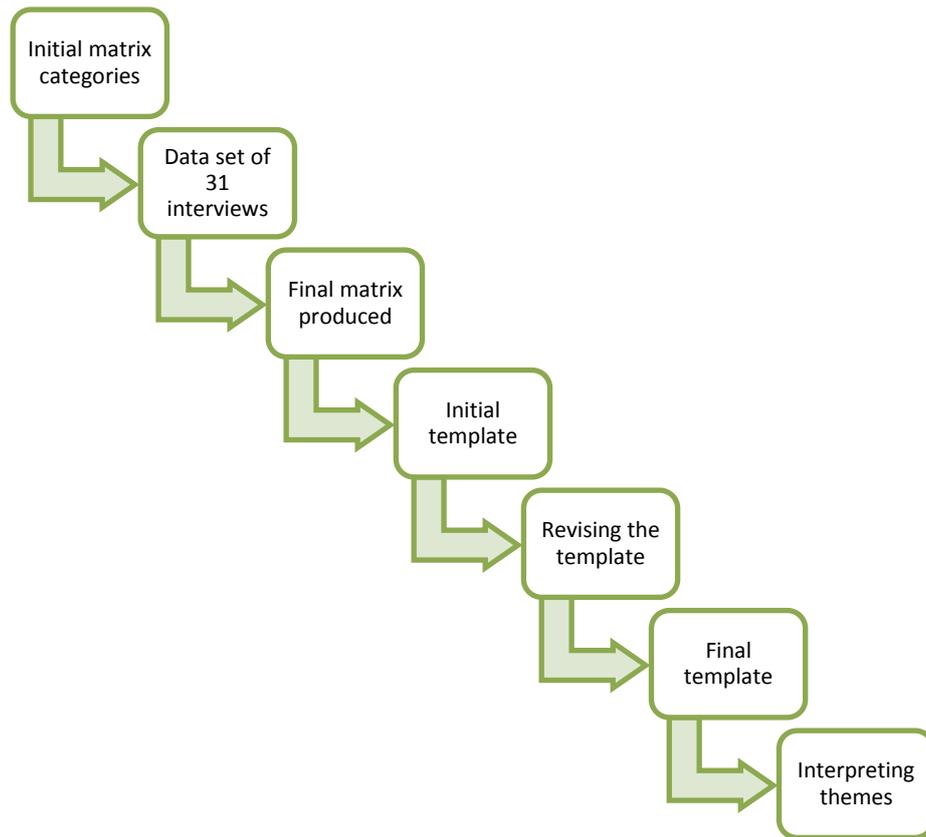


Figure 9: The data analysis process

Next, I will discuss how the template analysis method was operationalised.

3.5.2 Template analysis

3.5.2.1 Initial template

In line with King (2004), an initial template was developed after the three pilot interviews within the framework of the matrix. An initial template should not contain too many themes, for fear of blinkering the analysis that follows (Brooks & King, 2014), or too few codes for this may lead to a lack of direction. I decided to create an initial template after the three pilot interviews because it helped ensure that the process I was adopting had resonance with respondents, and was yielding meaningful themes that could be coded in the way I had anticipated. As a novel approach²⁵, I felt that it was important to have an early indication that the method could be successfully operationalised before proceeding too far into the data collection process.

An example of how the data has been coded is shown in Table 8. The example relates to the time period “mid to late-80s”. The column titled ‘level’ refers to the architectural layer, in this case ‘product’, the column titled “theme” refers to how the theme was named, the column titled ‘meaning’ refers to the definition of the theme, and the column titled ‘textual data’ is the verbatim textual data from one of the transcripts. The textual data highlighted in yellow shows the relevant verbatim quotations. As shown in the example to follow, I have used, where possible, theoretical constructs derived from the extant literature as themes, such as “component interdependence”.

²⁵ Combining matrix and template analysis is a relatively new approach in the literature. King, et al., (2007) is a notable exception

Level	Theme	Meaning	Textual data
Product	Component interdependence	Components of the products are linked and cannot be easily separated	“what you had was the pricing, the commission structure. It was all intertwined, interlinked”
Firm	Integrated fund management	Fund management as an activity was done in-house	“Well my impression is they were much more vertically integrated. I mean the investment houses were by and large in-house”
Industry	Introduction of personal pensions	The Finance Act 1988 that deregulated personal pensions	“Then in 1988 we had the introduction of personal pensions. We had the Government advert “Breaking the Chains”.
Mirroring (example of parallel coding)	Team co-location Multi-disciplinary teams	Product development team located together Product development teams made up of cross-functional expertise	Product development was done by multi-disciplinary teams within the firm so you'd have an actuary; you'd have a marketing person; you'd have an IT person; you'd have a finance person. And they were all sitting in the same building.

Table 8: Example of coding textual data to themes

The initial templates generated following the three pilot interviews are shown below in Tables 9 to 18.

Product	Firm	Industry
Component interdependence	Gains from integration	Deregulation
Integrated fund management		Changes in distribution structure
Fund unitisation		Unit-linked entrepreneurs
Integrated advice		
Product complexity		

Table 9: Initial template product, firm and industry themes: mid to late-1980s

Mirroring
Team co-location
Multi-disciplinary teams
Sequential development process

Table 10: Initial template mirroring themes: mid to late-1980s

Product	Firm	Industry
Component interdependence	Gains from integration	Emergence of IFAS
Component independence	Gains from trade	Direct sales regulation
Fund components	Capabilities	Unit-linked entrepreneurs
Product complexity		Industry consolidation
		Pensions mis-selling

Table 11: Initial template product, firm and industry themes: early to mid-1990s

Mirroring
Team co-location
Multi-disciplinary teams
Sequential development process
Component structure mapping

Table 12: Initial template mirroring themes: early to mid-1990s

Product	Firm	Industry
Component interdependence	Gains from integration	Industry consolidation
Component independence	Gains from trade	Horizontal integration
Platforms	Capabilities	IFA channel
SIPP architecture		Pensions mis-selling
Drawdown		
Product complexity		
Stakeholder pensions		
Standards		

Table 13: Initial template product, firm and industry themes: mid-1990s to mid-noughties

Mirroring
Team co-location
Component structure mapping
Geographical dispersion

Table 14: Initial template mirroring themes: mid-1990s to mid-noughties

Product	Firm	Industry
Component interdependence	Gains from integration	Australian influence
Component independence	Gains from trade	Horizontal integration
Platforms	Capabilities	D2C channel emergence
Product complexity	Governance inseparability	Industry consolidation
Standards		

Table 15: Initial template product, firm and industry themes: mid-noughties to 2012

Mirroring
Team co-location
Component structure mapping
Geographical dispersion
SBU's
Integrator roles

Table 16: Initial template mirroring themes: mid-noughties to 2012

2012-2014			
Product	Firm	Industry	
Component interdependence	Gains from integration	Industry consolidation	
Component independence	Gains from trade	Horizontal integration	
Hybridity	Capabilities		
Guided architecture	Governance inseparability		

Table 17: Initial template product, firm and industry themes: 2012 to 2014

Mirroring
Team co-location
Component structure mapping
Geographical dispersion
SBU's
Integrator roles

Table 18: Initial template mirroring themes: 2012 to 2014

The initial templates illustrate how I thematically coded interview data according to the time period and unit of analysis. For the purposes of the initial templates, textual data was only coded to one hierarchical level to avoid putting too much structure around the themes that may hinder further data analysis (King & Horrocks, 2010). At this point, it should also be noted that data analysis software NVivo 10 was used to aid with organising and coding the data set. King (2004) suggests that hierarchical coding within template analysis is well-suited to computer-assisted analysis. I also found that Nvivo 10 was a useful tool to organise the templates within the matrix structure. NVivo does not analyse the data or make any interpretation of the data analysis on its own, however its organising capabilities made NVivo very useful when comparing themes to data, and reiteration through the data on a number of occasions led to various codes being inserted, deleted, combined, and order changes within the hierarchy (King, 2004). Using Nvivo made this process a lot simpler and more efficient. An argument against the use of software such as NVivo comes from researchers from an interpretivist or social constructionist stance who may be wary of the 'structure' that NVivo adds to the data analysis process, however I did not feel the use of NVivo conflicted with my subtle realist commitments.

3.5.2.2 Revising the initial template

The next stage of the process is to revise the templates as through the course of the data collection process weaknesses in the initial template may become evident. I transcribed and thematically coded the data as I went along, revising the templates as 'live documents'. King (2004) identifies four types of template revision, which I highlight with examples from this research study.

The first type of template revision is the insertion of a new code – new codes can be inserted into the template as a new higher order code, or as a new sub-code. An example of a new higher-order code would be the insertion of the “firm boundary determinants”. This is a higher-order code that falls within firm-level factors and encompasses a number of second-order codes that affect firm boundary decisions. As an example of the insertion of a new sub-code, ‘fund modules’, ‘fee modules’ are sub-codes of the initial higher-order code of ‘component independence’. Secondly, a researcher can delete a code. In this study, no codes from the initial template were deleted because they were recurring themes throughout, however King (2004) suggests that deletion is usually associated with a code that turns out not to be useful, or a code that is subsumed into other codes due to an overlap. Thirdly, codes can change scope as they may be either too narrow or too broad. In this study, the initial template had a firm-level code entitled ‘gains from integration’. This code was retained as a second-order code, but its definition was too broad. Two separate aspects of gains from integration were present in the data; integration gains arising from superior rents, and integration gains arising from superior productive capabilities. While linked, these two third-order codes were created that fine-tuned the definition of ‘gains from integration’. Changing higher-order classification is the fourth type of revision. A lower-order code may be linked to a different higher-order code. In this study, this type of revision was not used.

As revisions were made to the template, prior interviews were re-examined in order to review the data set in light of the revised templates. As King (2004:263) notes, this process can continue ad infinitum, and it is difficult to know when to stop. No template is ever considered final, however, I reviewed all interview transcripts frequently in light of the changing template.

The final templates are shown below in Tables 19 to 28 as follows:

Product	Firm	Industry
1. Component interdependence	1. Firm boundary determinants	1. Product market factors
1.1 Integrated fund components	1.1 Gains from integration	1.1. Market stability
1.2 Integrated advice	1.2 Governance inseparability	1.2. Here come the unit-linkers
1.3 Integrated IT mainframes	1.3 Knowledge specificity	2. Deregulation
2. Optimised	1.4 Absence of intermediate markets	2.1 PEPs
3. Emerging fund components	1.5 Gains from trade	2.2 Tax incentives
3.1 Reduced charges interdependence	1.5.1 Capabilities	2.3 SERPS
		2.4 PP regulation
		2.5 FSA Act 1986

Table 19: Final template product, firm and industry themes: mid to late-1980s

Mirroring
1. Within-firm
1.1 Actuarial-led
1.2 Co-location
1.3 Communication problems
1.4 Emerging component structure

Table 20: Final template mirroring themes: mid to late-1980s

Product	Firm	Industry
1. Component interdependence	1. Firm boundary determinants	1. Regulation
2. Component independence	1.1 Gains from integration	1.1 Pensions mis-selling
2.1 Fund component	1.1.1 Rents	2. Industry structure
2.2 Charges component	1.1.2 Capabilities	2.1 Unit-linked rate of adoption
2.3 Advice component	1.2 Gains from trade	2.2 Traditional provider consolidation
2.4 IT components	1.2.1 Rents	3. Changes in distribution structure
3. Specialised interfaces	1.2.2 Capabilities	3.1 Direct sales regulation
	2. Simplified information exchange	3.2 Pace of intermediation
		3.3 Demand for component variety

Table 21: Final template product, firm and industry themes: early to mid-1990s

Mirroring
1. Within-firm
1.1 Component structure
1.2 Geographical dispersion
1.3 Systems integration
1.4 Co-location

Table 22: Final template mirroring themes: early to mid-1990s

Product	Firm	Industry
1. Stakeholder pensions	1. Firm boundary determinants	1. Regulation
2. Intermediate SIPP	1.1 Gains from integration	1.1 PEPs and growth of ISAs
3. Unit-linked pensions	1.1.1 Rents	1.2 Charges disclosure
4. Product platform aggregation	1.1.2 Capabilities	1.3 Pensions mis-selling
5. Component interdependence	1.1.3 Control	1.4 Stakeholder
6. Component independence	1.2 Gains from trade	1.5 Tax changes for incumbents
6.1 Charges module	1.2.1 Rents	2. Industry structure
6.2 IT module	1.2.2 Capabilities	2.1 Vertical integration
6.3 Fund module	1.3 Intermediate markets	2.2 Horizontal integration
7. Specialised interfaces		2.3 New entrepreneurial entrants
8. Open architecture & SIPP		3. Changes in distribution structure
8.1 Drawdown		3.1 Networks
9. D2C channel		3.2 Commission wars
10. Industry interfaces		3.3 Intermediation growth
		3.4 Bancassurance ties

Table 23: Final template product, firm and industry themes: mid-1990s to mid-noughties

Mirroring
1. Within-firm and across firm
1.1 Component structure
1.2 Geographical dispersion
1.3 Systems integration
1.4 Co-location
1.5 Communication patterns

Table 24: Final template mirroring themes: mid-1990s to mid-noughties

Product	Firm	Industry
1. Component interdependence	1. Firm boundary determinants	1. Regulation
2. Specialised interfaces	1.1 Gains from integration	1.1 RDR
3. Component independence	1.1.1 Rents	1.2 A-day
3.1 Fund modules	1.1.2 Capabilities	2. Industry structure
3.2 Charges modules	1.1.3 Control	2.1 Vertical integration
3.3 Drawdown module	1.3 Gains from trade	2.2 Horizontal integration
3.4 Self-investment module	1.3.1 Rents	2.3 Entry of low cost TPAs
3.5 Tools modules	1.3.2 Capabilities	3. D2C channel
3.6 Fees modules	1.4 Intermediate market	3.1 Guided investment solutions
4. Product aggregation platforms	1.5 Governance inseparability	3.2 Orphan clients
5. Product outcome homogeneity		3.3 Cost of advice
6. Industry interfaces		4. Distribution
		4.1 Simplified investment proposition
		4.2 Movement to high-net worth
		4.3 Need for operational efficiency
		4.4 Professionalism

Table 25: Final template product, firm and industry themes: mid-noughties to 2012

Mirroring
1. Within-firm and across firm
1.1 Component structure
1.2 Geographical dispersion
1.3 Platform/Systems integration
1.4 Co-location
1.5 Communication patterns

Table 26: Final template mirroring themes: mid-noughties to 2012

Product	Firm	Industry
1. Component independence	1. Firm boundary determinants	1. Shifting value chain dynamics
1.1 Charges module	1.1 Gains from integration	1.1 TPA consolidation
1.2 Funds module	1.1.1 Rents	1.2 Value shifting to advice
2. Industry standards	1.1.2 Capabilities	1.3 Squeeze on fund management
3. Component interdependence	1.1.3 Control	1.4 Squeeze on product providers
3.1 Advice modules	1.3 Gains from trade	1.5 Emergence of power blocks
3.2 Tools modules	1.3.1.Rents	2. Regulation
3.3 Funds modules	1.3.2 Capabilities	2.1 Auto-enrolment
4 Specialised standards		2.2 RDR
5. Vertical specificity		3. Industry structure
6. Guided modularity		3.1 Vertical integration
7. D2C channel		3.2 Horizontal integration
7.1 Advice modules		
7.2 Funds modules		
7.3 Tools modules		

Table 27: Final template product, firm and industry themes: 2012 to 2014

Mirroring
1. Within-firm and across firm
1.1 Component structure
1.2 Geographical dispersion
1.3 Autonomous SBU
1.4 Co-location
1.5 Communication patterns

Table 28: Final template mirroring themes: 2012 to 2014

3.5.2.3 Interpreting the template

King (2004) offers four suggested approaches to interpreting the templates. Without a process of interpretation, template analysis achieves very little, as interpretation is a step in the process that should attempt to do justice to the richness of the data set. The first approach to interpretation is the listing and counting of codes. King suggests that the counting of codes can often be useful in some studies, however, as King (2004:266) warns “*the frequency of codes per se can never tell us anything meaningful about textual data*”. As the aim of the study is to represent similarities and differences in accounts of phenomena over time, I have not adopted a listing and counting approach to interpretation.

The second and third approaches are selectivity and openness, where a researcher attempts to interpret every code with equal detail. I have heeded this advice however my aim is to build a narrative around the themes that are of central importance in building an understanding of events, and inevitably this may mean less of an emphasis on themes that are considered less central or peripheral to the narrative. The fourth approach is to explore the relationships between themes. The relationships may reveal some ‘integrative themes’ which permeate much of the template, possibly within a particular time period, or across all time periods in the study. This is the primary approach to interpretation that is adopted in this research study, and I will use a range of templates to explore relationships across time periods and across different architectural levels. The aim is to identify common ‘integrative themes’ across different time periods that may help us understand the phenomena in a richer detail. There is also the potential to use a multiple case-study approach (Yin, 2014) and explore the similarities and differences across different organisations. This may offer an opportunity to extend the research at a later point, as highlighted in chapter 6.

3.5.2.4 Writing-up

King (2004) suggests that in writing up a research study there are three common approaches:

- A set of individual case studies, followed by a discussion of the similarities and differences
- Structured around the main themes, with illustrative examples drawn from the textual data
- Thematic presentation, using a different case study to illustrate each of the main themes

I am drawing upon the advice of King (2004:268) in that he observes that an account structured around the main themes, with illustrative examples *"tends to be the approach which most readily produces a clear and succinct thematic discussion"*.

The section that follows examines issues of quality and how retrospective longitudinal research may be judged.

3.6 Issues of Quality

This section has two aims. First, I discuss quality as it relates to retrospective longitudinal research designs. Second, I discuss how I have used member-checking and expert panels to further enhance research credibility and plausibility (Hammersley, 1992).

Although retrospective longitudinal research designs have been used in a number of disciplines such as psychology, sociology, and organisational and management studies (Snelgrove & Havitz, 2010), and is typical of the style of research found in peer-reviewed journals such as *Academy of Management Journal* (Cox & Hassard, 2007), some criticism has been directed towards the validity and trustworthiness of retrospective and longitudinal research designs.

This section aims to comment upon the appropriateness of a retrospective longitudinal research design in light of the study's research aims and objectives. It is worth, however, recalling my subtle realist commitments. This sub-section focuses on the steps I have taken to enhance the plausibility and credibility of this research study. It has not been my aim to eliminate bias or somehow get to a single 'truth' in line with positivist commitments, rather to reflexively consider how issues around memory and recall may influence the findings, as well as how to improve the quality of the research.

One of the most frequent criticisms of retrospective longitudinal studies is the fallibility of respondent memory and recall (Cox & Hassard, 2007; Nisbett & Ross, 1980). However, Snelgrove and Havitz (2010) argue that memory fallibility can be reduced if respondents are asked to recall information about 'critical incidents' rather than general information, and I have used this approach in the interview setting. The interview process was divided into two parts: in part 1, respondents were asked to match product architecture types to time periods in order to construct a product timeline. The product architecture types acted as a 'specific cue' (Snelgrove & Havitz, 2010) to improve recall from long-term memory. The resulting product timeline can be conceptualised as a series of critical incidents that were then discussed in part 2 of the interview. The criticality or distinctiveness of past events has a positive impact on recall; quoted in Snelgrove and Havitz (2010:340), Yarrow *et al.*, (1970:71) suggest if respondents are able to "*recall specific events, they will have been extreme or unusual events in their experience*".

Another common criticism of retrospective longitudinal accounts is that respondents may reconstruct history, perhaps with intention or unconsciously. For example, respondents may attempt to project a socially-desirable image or orientate past decisions within a light of rationality, and such attributional bias may be problematic in strategic management contexts because of the social commitment to the work-place. In addition to the use of critical incidents and specific cues to reduce reconstruction bias, a technique I used was to ask respondents to only respond to a question within their own direct professional experience and not to elicit a response if they were uncertain or could not remember.

This approach draws upon the research of Koriat and Goldsmith (1994) where such techniques were associated higher accuracy. A second technique was to construct the purposive sample with multiple respondents for each job role, organisational background and across time periods (Cox & Hassard, 2007), and Miller, *et al.*, (1997:201) suggests that if such guidelines are followed “*scholars could be truly comfortable with the idea that retrospective reports are not fiction*”. An alternative approach I rejected was to seek methodological triangulation and combine data from mixed methods. The reasons for this are both pragmatic and methodological. For example, the availability of secondary data or published sources across the 30-year time period is limited, particularly the further back in time you go. To illustrate this point, until the mid to late-90s product specifications did not exist in many firms. From a methodological perspective, quantitative-orientated methods would be inappropriate for this research study as the total population of potential respondents is estimated to be less than 70 or thereabouts and therefore statistical inference would be highly problematic.

Nonetheless, retrospective longitudinal research can remain problematic. For example, it remains possible that interviewees, intentionally or unconsciously, may be lying, telling only part of the ‘truth’, or because they wish to conceal particular mechanisms. It is also possible that respondents are simply unaware of particular aspects of their social world (Gardner, 2001). However, De Meuse, *et al.*, (2001) justify a retrospective longitudinal research design in the field of human resources management as appropriate because “*to track employee perceptions of the psychological contract over 50 years in a truly longitudinal fashion would have been virtually impossible. Consequently, a retrospective methodology was utilised in this study. The authors recognise that there are problems associated with this approach.....Despite justifiable concerns about the accuracy of retrospective designs, this study supports the contention of researchers who assert that these designs can be useful in identifying patterns indicative of dynamic processes*”. I adopt a similar perspective.

To chart the co-evolution and correspondence of architectures over a period of 30 years would have been virtually impossible by any other approach. Consequently, a retrospective longitudinal methodology was utilised. Recognising that biographical knowledge remains fuzzy and ambiguous, however, means the retrospective designs “*should not be used as an excuse for lack of rigour in interrogating these data*” (Gardner, 2001:198). With this in mind, I now turn to discuss issues of research quality.

There is often no general agreement about what constitutes ‘good’ research or what quality criteria to use (Hammersley, 2007), however, as Madill, *et al.*, (2000:2) argue “*it is still entirely appropriate that qualitative research be open to scrutiny and that the credibility of findings rest on more than the authority of the researcher*”. Given the heterogeneity of methods in qualitative research, it is important that any criteria to assess quality are contingent upon the epistemological tradition, what Johnson, *et al.*, (2006) referred to as a “contingent criteriology”. Some scholars within a radical constructionist tradition may resist any attempts to judge the quality of research (Madill, *et al.*, 2000). However, within a ‘realist’ tradition, some scholars have suggested that the notions of objectivity and reliability are possible to transfer from a positivist tradition (ie, Miles & Huberman, 1994), while yet others have proposed alternative quality criteria (ie, Hammersley, 1992; King & Horrocks, 2010; Lincoln & Guba, 1985; Madill, *et al.*, 2000). There are, therefore, a number of criteria that can potentially be adopted to assess the quality of qualitative research, contingent upon the philosophical tradition of the research and its design. In this research study, in addition to the steps I have outlined in respect of the retrospective longitudinal design, I will now outline the steps and procedures adopted in terms of member-checking and the use of expert panels (Fereday & Muir-Cochrane, 2006; King & Horrocks, 2010).

3.6.1 Member-checking

In order to strengthen the face validity of the research (Paton, 2002), before I progressed to the main sample, each of the three pilot interviewees was emailed the following for review:

- Interview transcript
- Matrix
- Initial template
- Example of coded interview transcript

Feedback suggested that a small number of ‘theoretical’ themes were unclear to respondents and therefore some themes were renamed to reflect industry-lingo.

Following each interview, all thirty-one respondents were emailed the transcriptions of the interviews. The covering email asked respondents to make amendments to the transcript if they deemed it appropriate. Twenty-six of the thirty-one respondents confirmed by email that the transcript was an accurate reflection of the interview. Five respondents did not reply. As such, interview transcripts were member-checked by twenty-six respondents prior to data analysis.

3.6.2 Expert panel

In line with Hammersley (2008), it is important to test the validity of one's own interpretations, and so I decided to use an expert panel to review my data analysis ex-post, heeding the advice of King and Horrocks (2010:163) that "*external views are generally required to give a process credibility*". The aim of the process was to assess the plausibility and credibility of the matrix and final templates. I selected two persons for the panel, based upon their relevant professional experience in the product market covering the period between 1984 and 2014 and presence in the trade media as 'opinion leaders' and 'experts'. One person had participated in the pilot study (interviewee_1) and the second person had not participated, but knew its background and context. A meeting was held in Edinburgh on April 3rd 2015.

The expert panel were invited to review and discuss:

- Synthesis of the thirty-one product timelines
- The description of the themes, and examples of how verbatim quotations had been coded to a theme
- The final matrix and templates

The expert panel agreed that the synthesis I had identified was plausible and credible. There were no comments in respect of the sample coding I had given them to review. The final matrix and templates raised some discussion, such as whether it would be appropriate to subdivide the sub-period "mid-90s to mid-noughties" into two further sub-periods of "late-90s" and "early-noughties" as the time period encompasses a significant architectural change around year 2000, however it was agreed that the benefits of doing so would be less than the additional complexity that it may bring to the research. On the whole, the two experts both felt that the final matrix and templates were a credible representation of the co-evolution and correspondence of architectures in the product market.

3.7 Positionality

Researcher positionality is an important consideration in this research study as I was employed in three different organisations in the product market between 2000 and 2011 as a product architect. Merton (1972) identified two opposite views on positionality, namely (1) the outsider doctrine and (2) insider doctrine. The outsider doctrine values research conducted by researchers from outside their own communities, where notions of objectivity and neutrality are often sought after. The insider doctrine, by contrast, suggests that to understand a situation a researcher must be a part of it, and therefore that insider researchers are often in a unique position to understand the experiences and perspectives of groups which they are members of. Insider research is often associated with advantages related to the engagement of respondents and their willingness to share experiences and perspectives (Kerswetter, 2012), relational intimacy and a superior understanding of a group's culture (Breen, 2007).

Taylor (2011) also highlights a number of advantages including the deeper levels of researcher understanding through prior knowledge, knowing the lingo of respondents, easier access to and selection of respondents, and quicker establishment of rapport and trust. Taylor balances her argument with some disadvantages, such as researcher positioning in relation to the quality of the data collected and problems arising from the insider position, such as bias and blindness to the obvious (Chavez, 2008). Breen (2007) also discusses the complication inherent in insider research as a loss of objectivity, especially in terms of making assumptions based upon the researchers' own prior knowledge. Much work, however, assumes that being an insider, or at least being perceived as an insider, *"is an advantageous position in which to be since this gives the researcher a privileged position from which to understand processes, histories and events as they unfold"* (Herod, 1999:320)

However, Herod (1999), and others, note that a crude dichotomy between inside and outside research isn't quite so analytically simple and is in practice rather more "messy". As Chavez (2008), Kerswetter (2012), and others argue, often few researchers can be characterised as either complete insiders or complete outsiders, and occupy somewhere in the 'space between' (Dwyer & Buckle, 2009), which often means that researchers will have a responsibility to make explicit where they are positioned within the space and how this may effect research design and outcomes. The distinction between insider and outsider perspectives can often correspond to different and contrasting epistemological positions. For example, interpretivist epistemologies may often be associated with insider research, and outsider research can often be associated with more positivist epistemologies (Chavez, 2008). However, there are undoubtedly strengths and weaknesses to both types of positionality and perhaps a positionality that draws on the strengths and minimises the weaknesses of both may be considered the most appropriate. As such, Herod (1999) suggests that a researcher can manipulate his or her positionality so that, at times, the researcher can create distance, and at other times play down such distances in order to 'manage' the so-called positionality gap (Moss, 1995).

In this research study, I heeded the advice that the outsider-insider dichotomy is too simplistic and adopted an "in the middle" positionality (Breen, 2007), neither fully inside, nor fully outside in order to attend to both the advantages and complications of adopting one or the other. However, I also believed that a static positionality 'in the middle' was also problematic, and so decided to manipulate my positionality depending upon the extent to which each respondent may perceive my insiderness or outsiderness, a strategy referred to by Kanuha (2000:443) as "*researching at the hyphen of insider-outsider*".

What now follows is an account of how I enacted a dynamic 'in the middle' positionality in the context of the research question, data collection and data analysis phases of the research design.

3.7.1. Positionality and research practice

I was employed within a product development and strategic management role for three organisations in the personal pensions product market between 2000 and 2011. In the months following employment as an academic, I began to reflect upon my experience as a practitioner in the context of the strategic management literature. What had piqued interest was that the extant literature on architectural co-evolution and correspondence did not seem to adequately represent what seemed to have happened, or was happening, in the product market under enquiry. In other words, my prior ‘insider’ knowledge could not be fully reconciled with the academic literature within the strategic management field. Out of this interest, and discussion with potential supervisors, I began to review the literature for further information on the subject, which motivated me to explore the topic as a thesis. For these reasons, I became interested in a retrospective, longitudinal study of the UK personal pensions product market. This approach is characteristic of positionality in the middle, in so far as my prior insider knowledge of the product market influenced and guided the literature search and research questions. However, in contrast, the literature review was undertaken at a critical distance from the context. Finally, through a process of iteration, I settled upon some research questions.

Insiders can often face challenges in the data collection phase of a research project – and especially during the interview setting - such as conflict between the interview format and the conversational style of insider relationships, the legitimisation of insiderness, bias interviewing, and how the researchers’ identity may affect respondent accounts (ie, Chavez, 2008, Kerswetter, 2012). In this research study, I was aware of the advantages, but also the challenges, of an ‘in the middle’ positionality at all stages of the data collection and data analysis phases. For example, in gaining access to respondents, I made my academic and prior ‘insider’ status clear to potential respondents, because often academics are not seen as important and worthy of the commitment of time.

My perceived ‘insider’ status was considerably helpful in securing access to and commitment from respondents. Often, practitioners in the product market context are reluctant to discuss experiences or perspectives with ‘outsiders’ because of the intricacies of the product market are often believed to be too complex or idiosyncratic for ‘outsiders’ to properly understand or comprehend. Furthermore, as the research project is currently very topical to what is happening in the product market, the research project was seen by respondents as instantly ‘legitimate’, and respondents were on the whole extremely happy to be involved. In order to manipulate the insider-outsider dichotomy, I omitted any overt references to academic ‘lingo’ in correspondence that I felt may alienate respondents, such as references to ‘Dr’, ‘PhD’, or ‘doctorate’, other than that required by ethical standards. It was felt that this would more quickly engender trust and relational intimacy.

Insider researchers often rely upon respondents with whom they are familiar (Breen, 2007), and this is one reason why the total sample included two parts. The first part of the sample was generated from professional contacts, via social media platforms such as LinkedIn. Therefore, I considered it advantageous to ensure that the second part of the sample contained respondents that not only met the screening criteria, but who were not known to me personally or professionally. As a consequence, the *total sample* consisted of some respondents who may perceive me as an ‘insider’, and some who may consider my role as an ‘outsider’, as shown in Figure 10.

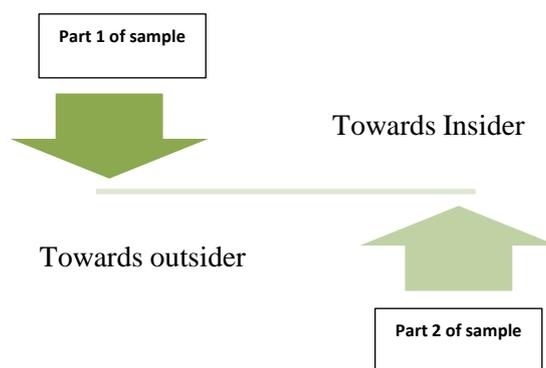


Figure 10: Insider – outsider: balancing the purposive sample

During the interviews, I explained to all respondents the extent of my prior knowledge of the product market and re-confirmed the purpose of the research study. To maintain a critical distance, I reminded respondents that I had no direct knowledge of the product market since 2011. Furthermore, I would only use my existing knowledge of the product market between 2000 and 2011 to orientate follow-up questions or probe for clarification, but, importantly, respondents should assume no prior knowledge on my part so as to make sure the ‘obvious’ was included in respondent accounts. As Asselin (2003) reminds us, it is best for an insider researcher to assume he or she knows nothing about the phenomenon being studied.

There is also the risk that the interview may be shaped and guided by my prior experience, and not the experiences and perspectives of the respondents. This risk was an influence in my decision to adopt a minimalist interview schedule so as to allow respondents to share their own experiences and perspectives without too much structure. Following DeLyser (2001) and Chavez (2008), a researchers’ request to assume no prior knowledge seems for the most part to have worked in eliciting rich detail in respondent accounts and in avoiding comments such as “you know what I mean” and frustration whereby the respondent often assumes the researcher already knows the answer to the question (DeLyser, 2001).

I was also aware of subjective influences that I may have on the reliability of respondent accounts. It is known that interview respondents may react defensively or try to impress the interviewer (Van Heurten, 2004) with the ‘right’ answer. Using decision-making as a theoretical context, Van Heutgen (2004:211) points out that *“respondents may have difficulty remembering the process, and be unaware of what really influenced their decisions. They then may reconstruct what actually occurred according to what they usually do or are supposed to do. They will rationalise what is frequently an irrational process by creating logical stories or saying what they think the interviewer wants to hear”*. Cited in Van Heutgen (2004), Merton (1946) recommended that such types of bias can be overcome by keeping direction to a minimum. Furthermore, other tactics such as building a basis of trust, the use of probing questions, triangulating the data across numerous respondents or with the extant literature can help overcome these issues. This advice was heeded in this research study, as discussed.

Kanuha (2000:444) asserts that where a researcher is known as being a ‘native’, the potential benefits of access to privileged information needs to be balanced if the researcher wishes to avoid blind spots pervading the data analysis phase and that *“for each of the ways that being an insider researcher enhances the depth and breadth of understanding to a population that may not be accessible to a normative scientist, questions about objectivity, reflexivity and authenticity of a research project are raised”*. In addition, Asselin (2003) points out that one of the risks of being an insider in the data analysis phase of a research project is that the researcher may suffer role confusion whereby the data is analysed from a different perspective from that intended, and that the data analysis proceeds with an emphasis on shared factors and that discrepant factors are de-emphasised (Dwyer & Buckle, 2009).

To overcome these sources of positionality bias, I was aware of the need to encourage alternative explanations in order to avoid blind spots and open up my own subjectivity to scrutiny and rigour. As discussed, I have used member validation and expert panels in the data analysis phases in order to imbue ‘confidence’ in the findings. These steps, together with access to respondent accounts unlikely to be available to outsiders, are posited to improve the plausibility and credibility of the research findings. Nonetheless, there are possible unknown impacts that my insidership/outsiderness may have had on the data collection and data analysis processes.

In summary, there are obvious advantages and disadvantages of being an insider or outsider in the research process. I have highlighted how questions of positionality have influenced aspects of the research design and process. DeLyser (2001:442) reminds us that *“in every research project we navigate complex and multi-faceted insider/outsider issues”*. In this research study, I have positioned myself as neither a complete insider or complete outsider, and tried to maximise the potential advantages, and minimise the potential disadvantages of both, occupying a dynamic space ‘in the middle’ of the continuum, following the advice of Acker (2000) in her suggestion that we attempt to find a way to be both. By reflexively manipulating positionality, I hope to have been able to elicit rich respondent experiences and perspectives, and to minimise potential sources of positionality bias. As Dwyer and Buckle (2009:59) argue, *“the core ingredient is not outsider or insider status but an ability to be open, authentic, honest, deeply interested in the experience of one’s research participants, and committed to accurately and adequately representing their experience”*.

3.8 Ethical considerations

There are a number of professional associations that offer guidance on good ethical practice in social research. For example, The Social Research Association (SRA) advises on carrying out ethical social research. Within the discipline of business and management, the British Academy of Management, the American Academy of Management, and the Academy of Marketing offer guidance to researchers working within the field, and similar bodies exist in other disciplines in the social sciences such as the British Psychological Society and the Educational Research Association. Within Northumbria University, the Principles of Good Research Practice and governed by an independent Research Committee in each School (School of Business and Law). This study was granted ethical approval prior to data collection on 29/9/14.

Within qualitative research, Willig, (2001:18) outlines five key areas of ethical practice:

1. Informed consent – this study offered to each interview participant a copy of Northumbria University ‘Informed Consent Form for research participants’ (see Appendix 5)
2. No deception - prior to any data collection, each respondent was informed of the purpose of this research, its aims and outcomes
3. Right to withdraw – respondents were informed in writing of their right to withdraw from the study at any time, for any reason without question (see the ‘Informed Consent Form for Research Participants, Appendix 5)
4. Debriefing - respondents were emailed a copy of the interview transcript and promised an executive summary of the final report
5. Confidentiality - within this study the interview data was stored on an I-phone recorder for no more than 24 hours. This was downloaded to my own private password-protected computer, and saved as password-protected electronic files, only accessible by me. The audio recordings on the I-phone were then deleted. Respondents and organisations were anonymised with a numbering system. The numbering system was maintained in a separate, password-protected electronic file. Any references to locations, such as a building, street, town, or city were also anonymised as I felt that this may reveal the identity of an organisation. Therefore, the study guaranteed confidentiality and anonymity of respondents and organisations. The following wording was included in all correspondence:

All data will be stored securely either electronically on computer or in hard copy version in a locked cupboard. The recordings on the digital recorder will be deleted once they have been saved on my own password-protected laptop. This will be done within 2 weeks of the initial recording. As part of the data analysis process, hard copies of the anonymised transcripts (raw data) may be given to the doctoral supervision team and a small number of other research participants to review. Hard copies will be returned to me. Data obtained through this research may be reproduced and published in a variety of forms and for a variety of audiences related to the broad nature of the research detailed above. It will not be used for purposes other than those outlined above without your permission. Participation is entirely voluntary and participants may withdraw at any time.

Where audio files were professionally-transcribed, the audio files were downloaded onto a CD and subsequently deleted from the device. The audio files were sent to the transcription agency by recorded delivery, and returned in the form of password-protected electronic files by email. The transcription agency was advised of my data confidentiality and privacy policy before starting any work, and the transcription agency used has the University's preferred supplier status.

In addition to the 5 areas set out by Willig (2001), Easterby-Smith *et al.*, (2012) added further criteria for ethical research.

6. Ensure no harm – this study's aim was to ask respondents to share their experiences of events in a product market. As such, no harm was done to respondents
7. Treat respondents with dignity – in this study the respondents were informed in writing of the purpose of the study and were treated with the due respect and formality by me

8. Declare affiliations and conflicts of interest – the study is funded by Northumbria University. I worked in the product market between 2000 and 2011, and is known to some of the respondents. Where I was known to the respondent, this was acknowledged in communication and in the interview. I agreed with each respondent that the interview should proceed on a formal, professional, but friendly, basis, and that the respondent should assume I am not familiar with the product market context. Where I was not known to the respondent, I declared an ‘insider’ status, but again agreed with each respondent that the interview should proceed on a formal, professional, but friendly, basis, and that the respondent should assume that I am not familiar with the product market context. I declare no conflict of interest or affiliation with any of the respondents or organisations
9. Honesty and transparency in communicating about the research – in this study the participants were informed in all forms of correspondence the purpose and outcome of the study, and the data protection and privacy procedures
10. Avoidance of any misleading or false reporting of research findings. The purpose of the study is academic. However, the findings may be of interest to practitioners and policy-makers. Copies of an executive summary will be made available to respondents.

3.10. Summary

This chapter has explored the research design of this study. It has explained and justified epistemological commitments, and discussed methodological decisions. The aim of the chapter was to justify the selection of a subtle realist perspective, and to position the study’s data collection and data analysis methods within that frame of reference.

The next chapter presents the Data Analysis and Findings.

4. Data Analysis and Findings

In this chapter, I will present the data analysis and findings. One of the most common ways of presenting a thematic or template analysis is to “describe and discuss each of the overarching themes in turn, referring to examples from the data and using direct quotes to help characterise the theme for readers” (King & Horrocks, 2010). Rather than a purely descriptive account, this chapter aims to follow the guidance of Braun and Clarke (2006) by making an argument in relation to the data. As King (2004:268) reminds us, “this tends to be the approach which most readily produces a clear and succinct thematic discussion”.

One of the challenges of presenting the data analysis and findings in this research study is to maintain a clear and succinct discussion across multiple time periods, multiple units of analysis, and the relationships between them. To try to overcome these challenges and present a clear and succinct discussion, the overall structure of the chapter will be presented as follows:

- Two main sections that correspond to the two overarching research questions:
 - RQ1: What supply-side processes are associated with the co-evolution of architectures?
 - RQ2: To what extent are the design characteristics of product architecture associated with the design characteristics of firm and/or industry architecture?

Within the RQ1 section, I focus on (1) product themes across time, (2) firm-level themes across time, and (3) industry level themes across time. Within the RQ2 section, the organising principle is time, moving successively through time from the mid-80s to 2014.

Within the templates throughout this chapter, I have chosen to illustrate the full set of themes that emerged from the data. However, given the restrictions of space and word count set out in the Northumbria University PhD regulations, it is not possible to discuss all aspects of the data. Therefore, from the templates I have carefully selected a number of ‘integrative themes’ (King, 1998, 2004) which re-occur across time and which help illuminate the core issues related to RQ1 and RQ2. In other words, the narrative within this chapter is a *sub-set* of the templates, which, by definition, is a summary of the interview transcripts.

Finally, as a context-specific research study, this chapter inevitably uses technical vocabulary and lingo. To enable the reader to understand some of the explanations, footnotes are used to elaborate key issues that impact upon the key contributions discussed in chapter 5.

The templates shown throughout this chapter are diagrammatical representations, using NVivo, of the tabular final templates as shown previously in sections 3.5.2.1 and 3.5.2.2.

In vivo quotations will be shown in *italics* and referenced as follows, for example, (interviewee_28) will be shortened to (28).

I begin, however, with a short narrative of a few significant historical milestones in the history of the UK personal pension product market that sets the context for the starting point of this research study. The mid-1980s is the starting point of this research study as it is widely-recognised that so-called ‘modern’ personal pensions are associated with the personal pension regime announced in the Finance Act 1984 and which received Royal Assent to become law on 1 July 1988. As a consequence, this research study begins at the cusp of an evolution from ‘old’ style pensions to ‘modern’ person pensions.

Individual personal pensions were initially offered in the post-war period of the mid-50s, when “*we had a very simple product range, typically retirement annuity contracts that came in about '55 or '56 and ran till '88*” (27). These products remained largely unchanged until 1988, linked to a single investment fund, known as a ‘with-profits’ fund designed and developed by insurance companies, known as ‘with-profits insurance companies’. A significant development then occurred in “*1960, the first unit-linked insurance companies came along. And what caused them to set up a unit-linked company, there was the Prevention of Fraud Investment Act and that prevented the door-to-door selling of unit trusts*²⁶. *Up to that point they were selling unit trusts door-to-door. But they couldn't do that anymore. But what they could sell door-to-door was life-assurance linked pension policies. So they introduced a life-assurance pension plan which was linked to their unit trusts*” (24). In other words, this event brought the idea of unit-linked investment to the UK for the first time.

Executive Pensions embedded unit-linking as a design concept in the mid-70s, “*in c1975 there was a new piece of legislation which allowed for executive pension plans. There was concern about the fact that smaller corporations didn't have pension plans because running a defined-benefit pension plan*²⁷ *was a big expensive operation. You needed trustees. You needed actuaries. You needed someone to do your investments. Big and expensive overheads. So, for big companies, it was ok, but for small companies, it wasn't. So, they introduced some legislation in '75 for the executive pension plan, which allowed small companies to set up a pension contract for as few as one person. They were often sold initially as directors' contracts. And right from the very beginning they were unit-linked. Unit-linking was now established in a small but growing sector. And many organisations sold these in quite large numbers making full use of the flexibility of the regulation. So, if you look at the mid- to late '70's, there were a lot of unit-linked insurance companies that started in that time*” (25).

²⁶ Known alternatively as ‘funds’, ‘mutual funds’, or ‘fund components’.

²⁷ Known alternatively as a final salary scheme.

In 1979, an innovative organisation entered the UK personal pensions product market with a unit-linked executive personal pension. And, by 1984, *“they introduced effectively a platform. But it was introducing about half-a-dozen investment houses to compete with your traditional life companies. So you're in a market environment where all the life companies were essentially doing the same thing and there was opportunity to do something fresh”* (27). The history of this significant moment is summarised as follows: *“[Organisation] came into the country from a South African base and they had unit-linked methodologies down in South Africa in the 50's. But investment in those days was not allowed outside the UK, more or less. You couldn't invest. There were severe currency restrictions. So, the only funds that they had were a cash fund, a UK equity fund and a property fund. So, unit-linked pensions were very slow to take off”* (24). However, *“[Organisation] realised that unit-linking would deliver a mass market customer need but it also delivered for the shareholder in the sense of not tying up huge amounts of capital to support the guarantees implicit in traditional with-profits. They were saying, actually one of the big components in the product world is the investment solutions and the investment solutions are very opaque. So, they were redesigning one of the core components”* (24).

The announcement of the introduction of personal pensions in the Finance Act 1984, and its subsequent launch on 1 July 1988, became the catalyst for a significant period of change in the product market which continues until today. This research study begins its data analysis from the mid-80s until 2014 and proceeds from micro to macro (ie, product to industry).

4.1 RQ1: What supply-side processes are associated with the co-evolution of architectures?

4.1.1 Product-level analysis

4.1.1.1 Mid to late-1980's

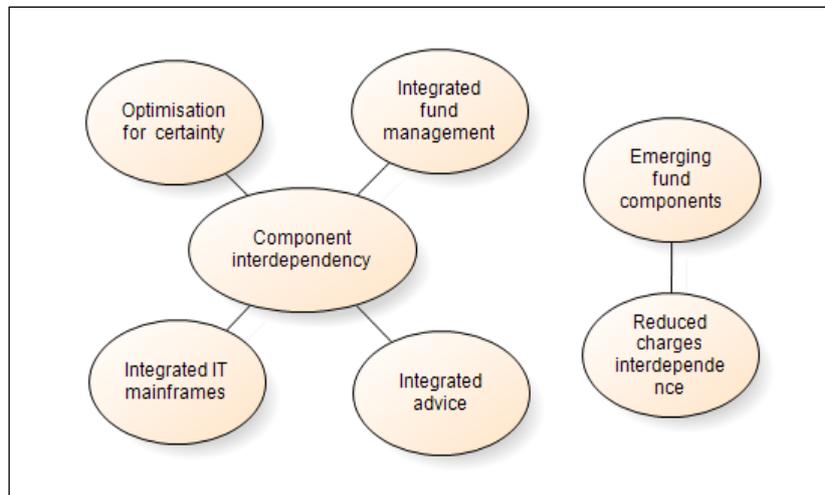


Figure 11: Product template: mid to late-1980s²⁸

In the mid to late-80s, respondents associated the period with a closed and integrated product architecture that represented a ‘with-profits personal pension’, as shown in Figure 12.

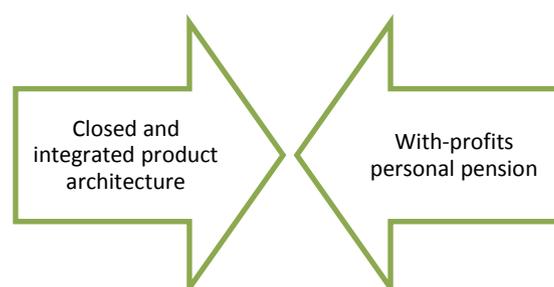


Figure 12: Closed and integrated product architecture = with-profits personal pension

²⁸ As also shown in tabular form in Table 19 (section 4.3.2.2)

The with-profits product architecture had closed and integrated characteristics, as two respondents remarked “*it was all intertwined, interlinked*”(28), and “*most components are interdependent with each other*”(18). The interdependence of product components had significant consequences for the product development process as changes in one product component often had knock-on, and potentially unforeseen, design changes in other product components, “*you tweak one part, other bits need to be looked at*” (28), “*it had to be all our components, and if you changed one thing, the whole thing changed*” (22), “*they're incredibly tough to change because everything's integrated, everything has an impact on everything else*” (15), and “*with-profits were the most internally complex products, and most implications were not foreseen*” (3).

This often led to a complex product development processes that was hindered by numerous re-iterations and re-works. A respondent reflected “*so then another product comes along that's broadly similar but they're all a little bit different, so you've got to re-think it through all again. I would have said the product development process was quite lengthy*” (4) and another reflected “*it was very hard to change, they are tightly-bound. And you couldn't really see how any of those products were going to be de-constructed so you could add a different type of fund to them*” (27). As one respondent recalled, product component change was often hindered and often resulted in a more architectural type innovation “*if you want to change anything, you basically find a new product*” (18).

In the mid-80s, personal pensions were designed on proprietary IT mainframes, supported by a complex set of manual operational processes. The complexity of the product architecture meant that many IT systems were not fit-for-purpose, as one respondent recalled “*they weren't necessarily fit for purpose because most systems were really intense from an actuarial perspective*” (27). One of the additional challenges was that IT systems were used to design and administer multiple product sets, including non-pension products, building in further cross-product interdependence, “*we had an endowment policy; we had a whole life policy; we had a single premium bond; we had a pension; and they were all run with the same system. The only difference between the pension and the bond was that the pension had a piece of paper which said different things on it. The system behind it was exactly the same, but there were no industry standards whatsoever*” (4).

In the mid-80s, a specialist pensions product called an ‘Executive Pension’ was gathering momentum²⁹. An Executive Pension was a personal pension that “*tended to be targeted at either the owners of businesses or some of their key employees*” (26). Executive Pensions offered more investment variety and allowed executives to take more control of their retirement, as one respondent stated “*there was a very big divide between employed people and self-employed people, the types of products that would appeal. And that's again another game changer that materialised through the '80's where people could top-up into an executive personal pension*” (27). A respondent also recalled that this ‘specialist’ product segment was dominated by new unit-linked entrants, “*so the early implementers of unit-linking looked at the Executive Pension market as being their market place*” (26).

By the late-80s, the concept of unit-linking³⁰ had permeated the design of many ‘modern’ personal pensions. Overall, however, many product architectures remain characterised as closed and integrated, but some unit-linked insurance companies began to modularise³¹ the fund component interface in order to increase product variety. Exogenous factors helped it gain traction as “*in 1980, the end of currency restrictions and .all of a sudden ooof*” (24). Therefore, by leveraging Executive Pensions as the thin edge of a wedge, new unit-linked insurance companies extended their unit-linked designs into mainstream personal pensions as a means of competing against incumbent with-profit insurance companies, “*late-'80's, there was an increasing trend of more investment choice becoming available through the unit-linked route*” (26) and personal pension designs launched on or after 1988 tended “*to be unit-linked at the beginning*” (24).

²⁹ Executive personal pensions were launched in the mid-1970s, but it was not until the announcement of a deregulated personal pensions market in the Finance Act 1984 that the product began to gather any momentum

³⁰ The difference between with-profit funds and unit-linked funds is that a with-profit fund is one homogenous pool of money managed on behalf of multiple consumers, “*if you look at pre-88 pensions, there wasn't an obvious correlation between what you were paying in and the fund value of your pension. It was conventional with-profits so it was based on a sum insured you might get at the end plus regular bonuses, plus a terminal bonus that you may or may not get*” (16). The actual value of your pension entitlement was by reference to complex actuarial calculations. Whereas, in a unit-linked structure the fund is split into slices, or ‘units’, and an individual consumer buys identifiable units in the fund so that the units purchased are directly linked to the value of the underlying assets of the fund. For example, a unitised fund with say 1000 units and with a fund value of £1000 has a unit price of £1. A contribution of £100 would buy 100 units. Later, if the fund was worth £2000 the unit price would be £2 - the pension’s unit value has doubled. Unit-linking therefore, can be characterised as a modular innovation of fund components.

³¹ In this context, modularising is the same as ‘unit-linking’

4.1.1.2 The early to mid-1990s

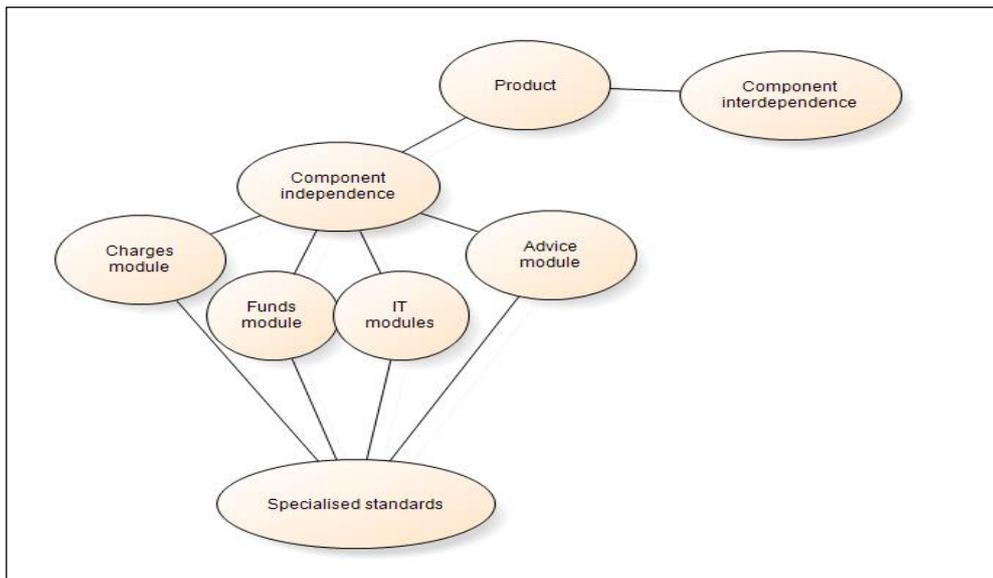


Figure 13: Product template: early to mid-1990s³²

In the early to mid-1990s, respondents associated the period with a closed and modular product architecture, representing a ‘unit-linked personal pension’, as shown in Figure 14.

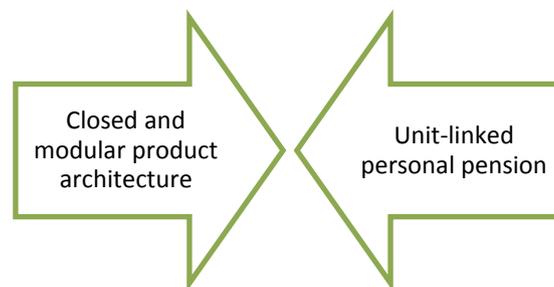


Figure 14: Closed and modular product architecture = unit-linked personal pension

By the early to mid-90s fund modularisation or unit-linking was much more main-stream across the product market. However, the variety of unit-linked funds available was still quite limited and nearly always designed internally within firm boundaries. For example, “*early 90s, you started to get people saying we’re going to have more than just a with-profit fund, and commonly would have four funds or five unit-linked funds*” (17).

³² As also shown in tabular form in Table 21 (section 4.3.2.2)

However, unit-linking required significant changes to the IT architecture. For example, *“I mean in a big monolithic system, it’s not very easy to do because you have to commit major surgery to cut the component out of the system. I can definitely remember that adding funds were made a lot simpler not just externally, but just adding funds internally was made simpler”* (4). The modular innovation process to unit-link fund components resulted in the creation of a specialised interface, known as a ‘mirror fund’³³. The mirror fund interface was a significant design breakthrough that, ultimately, paved the way for fund components to be sourced from external fund management groups. Outsourcing fund components was considered by many new unit-linked insurance companies as a route to differentiation or competitive advantage as *“it emerged as an opportunity to provide choice to customers”* (27). By the early-90s, some new unit-linked insurance companies had extended the range of modular fund components up to around 250 because *“our internal funds had been so incompetently run”* (13) suggestive of a capability gap.

Despite fund modularisation, industry standards had not emerged by the mid-90s, but firm-specific, specialised standards were permeating across firm boundaries as the basis for market exchange. For example, *“there were some company specific standards. You give us this sort of information and we can put your fund into our system”* (4), and in terms of the mainframe IT architecture *“there was also more standards inside the system, one bit talking to another, so I think the companies were building interfaces to try and componentise the system”* (4).

³³ The specialised interface was known as a mirror fund. A mirror fund is a new unit-linked fund that invests into another fund. Say, you have two funds, A and B, each with different management charges. Insurance companies created two more funds, A1 and B1. A1 invests in A and B1 invests in B. A1 is a ‘mirror’ of A and B1 is a ‘mirror’ of B. The design characteristics of A1 and B1 are standardised in order to avoid significant design implications. There were significant advantages for the insurance company in this approach as the assets within the mirror fund sat on the balance sheet of the insurance company even where the underlying fund was manufactured by an external fund management group, and mirror funds could be structured to have common dealing and valuation points, and other commonalities that had operational and efficiency benefits

4.1.1.3 The mid 90's to mid-noughties

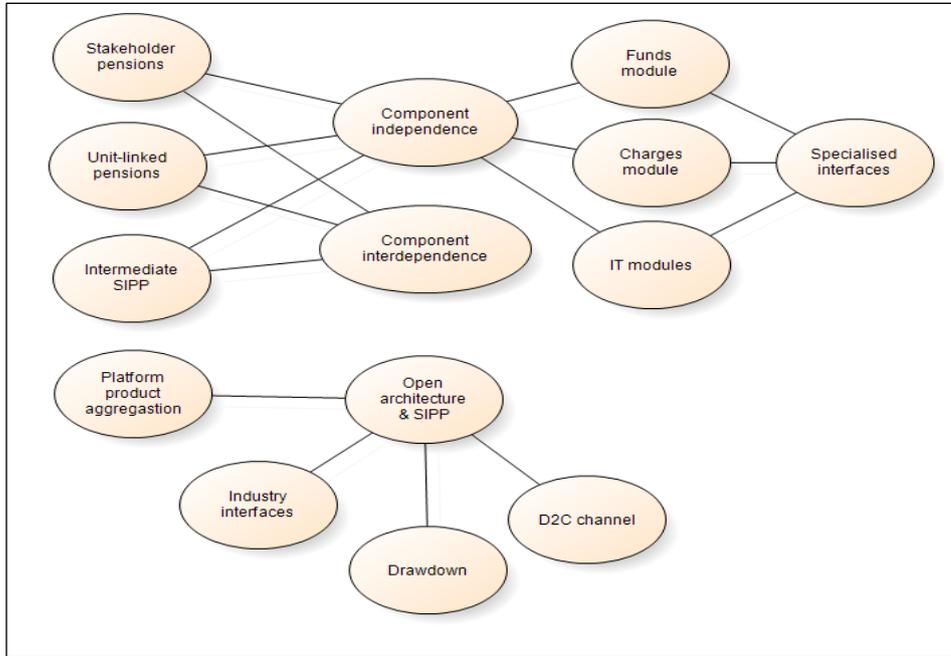


Figure 15: Product template: mid-90s to mid-noughties³⁴

In the mid-1990s to mid-noughties, respondents associated the period with a hybrid product architecture that represented a ‘fund supermarket’, as shown in Figure 16.

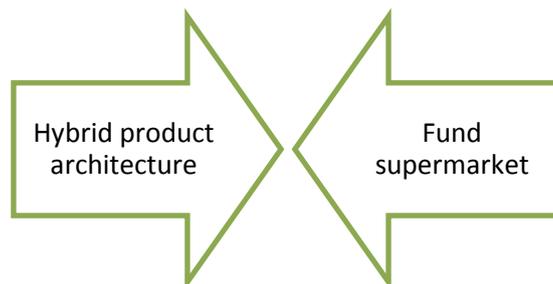


Figure 16: Hybrid product architecture = fund supermarket

³⁴ As also shown in tabular form in Table 23 (section 4.3.2.2)

During the mid to late-90s, the variety of unit-linked fund components increased significantly, *“so, in that time, personal pensions were offering a small range of 5 to 10 external funds. And then that developed and evolved to quasi-open architecture. There was an element of plug-and-play, but within a framework”* (11). Increasing fund component variety became a key basis of competition, as a few respondents recall: *“in '96/'98, products would have 15 or so external fund links”* (29); *“we had 50 funds within our personal pension”* (10); *“the big thing was about actually how do we take the pension and go from having say 10 funds to having 200 mirror funds”* (20); and, *“maybe to a range of 300 funds, where as previously, it would have been more like 20 funds”* (10).

The process of interfacing unit-linked mirror funds remained a firm-specific process rather than an industry-standard, largely due to the idiosyncratic nature of each firm’s mirror fund interface design, *“because what they had was a two-pin plug and you had a three-pin socket. So how did you make that work? You needed to have an adaptor in place. And the adaptor was a whole pricing system. And that was in-house”* (25) and *“you almost had an internal translation mechanism, so a piece of middleware, and that was the mirror fund”* (28).

By the early 90s, commissions paid to IFAs³⁵ were being charged within the product structure as a single up-front payment. In the single stroke of a pen, personal pensions were much more portable and exit penalties³⁶ could be removed from the product design. As a respondent recalls *“I remember around late 97, early 98, we launched a no-penalty pension. And that was the first product that hit the market that people could have flexibility, actually could move things without financial penalty and implication”* (17). The ‘modularisation’ of commission also coincided with the removal of other ‘low value’ integrated product components previously designed into the architecture and hence *“it was a time when you started seeing components like total permanent disability coming out of products. So products became more investment-focused and straight forward”* (17).

³⁵ Independent Financial Advisors

³⁶ Commission was previously paid as an up-front lump sum to financial advisers and collected back over time (usually 5 years) via product charges. As such, insurance companies held a balance sheet risk that the customer would switch providers within the 5-year period. To eliminate this risk, consumers were faced with a 5-year exit charge if they switched providers. As a consequence, exit charges can be conceptualised as an exit barrier that limited customers’ willingness to shop-around or switch providers. By charging commission as a one-off lump sum, no 5-year exit charges were needed. Conceptually, this design change is characteristic of a modularisation of the commission component

However, the core IT architecture remained tightly-integrated, *“there was already an emergence of external fund links, but that's as far as it went really. There was nothing much else which you could unplug. Everything else was run off big heavy mainframes and while inside those mainframes there are commission modules and charge modules and rebate modules and all that kind of stuff, everything else is still interlinked to some extent”* (28).

However, despite these drawbacks, insurance companies remained reluctant to consider modular IT systems architectures emerging in the intermediate market because *“it's not that simple because of interdependence. Some of the systems that we used as the core architecture for personal pensions, also powered other products that weren't ready or needed to move”* (11) and *“it was too hard, too expensive and too risky. No-one was signing up to it. No-one's going to say, what you should do is take out the entire core architecture of your system and replace it with something new that's going to cost you 50 million quid and take three years and be really risky. And in that time you're going to be able to sell nothing. No-one's voting for that”* (11).

Furthermore, because the core IT architecture was incredibly complex, few people understood how to simplify it, *“you've got a system that's probably got 30 years of knowledge on it and all the legacy built into it and trying to unwind that is a very hard task because no-one currently alive understands all the products that were on the system. I mean legacy was biting, there's no doubt about it”* (4) and *“It got to the point where the changes have become almost out of control and unworkable and the IT language that was all built in so archaic that very few people could actually use and work those systems”* (23).

While unit-linking continued to gain traction in the product market, with-profits personal pensions were reaching an implosive end. In an attempt to defend existing technologies, with-profits insurance companies announced higher and higher bonus rates each year without a direct relationship to the underlying performance of the fund, “*up to ‘97 the way that the with-profit declaration process worked was that you started to look at everybody else. You didn’t start looking at the financial health in your fund. What was happening is that it was a willy-waving contest to be honest. So people were declaring bonus rates beyond their financial means, and hoping that everything would all work out in the end. They played a huge gamble with the underlying assets. And all that happened was that it all imploded on it. I recall a strong desire from actuaries to get out because it was destroying capital hand over fist*” (28).

In parallel with the events in the personal pensions product market, a new product market for long-term savings had emerged, PEPs and ISAs³⁷. The PEP and ISA market led to a fundamental architectural innovation across the UK financial services sector. Upstream fund management groups seized upon the opportunity to enter the new product markets and leveraged existing fund management capabilities and modularised IT architectures developed for their core business of fund management. The PEP and ISA regulations enabled the use of CIS funds³⁸ but excluded the use of unit-linked mirror funds, and therefore the ISA ‘standards’ gave fund management groups an ‘architectural head-start’ in entering the ISA product market. As CIS funds were designed according to industry standards, fund management groups were quickly able to design ISA products to provide an almost unlimited variety of fund components, and ISA products quickly came to be available with access to 1000+ CIS funds – subsequently the term ‘fund supermarket’ was adopted.

³⁷ PEPs were initially launched in 1986 by the then Conservative government about the same time as the deregulation of financial services, and ISAs were announced by the then Labour government and launched in 1999 in order to replace PEPs. Both these products gave tax advantages to consumers who wished to hold long-term investments such as collective investment schemes (CIS funds). While PEPs dated back to 80s, it was not until the mid-90s and especially the technology-boom in the late-90s that the product really took off. PEPs and ISAs were not initially in direct competition with the personal pensions product market, but their growth had significant architectural implications for personal pensions.

³⁸ CIS funds are collective investment schemes designed according to industry standards such as EU standards. CIS funds were also the underlying funds linked to unit-linked mirror funds.

The emergence of modular ISA products with access to 1000+ industry-standard fund components provided a significant challenge to all insurance companies. However, at the same time unit-linked mirror funds suffered another (terminal) nail in the coffin³⁹. As a result, fund management groups spotted an opportunity to extend the ‘fund supermarket’ architecture to include personal pensions, *“it then exploded out again because actually those components were entirely interchangeable. And providers were then in a position where they could offer everything, and actually what you very quickly saw was an explosion of a supplier base particularly with use of funds because there were no limitations. So it was demand in the market and no limitations and so very quickly you ranged from having an in-house fund manager to having a small group of fund groups you might deal with and have a relationship with, to then having to deal with hundreds in some cases”* (6). With unit-linked personal pensions on the wane, SIPP⁴⁰ were identified by insurance companies and fund management groups as a ‘modular’ pensions architecture that could be developed within the parameters of a fund supermarket architecture, essentially creating a multi-product technological platform.

However, between the decline in unit-linked personal pensions in c1997 and the emergence and growth of fund supermarkets between 2000 and 2005, the then Labour government introduced Stakeholder Pensions in April 2001 in response to pensions mis-selling⁴¹ and consumer mistrust in the sector.

³⁹ In 1997, the then Labour Chancellor removed pension fund tax credits, and a consequence, perhaps unintended, it removed a tax-based barrier to holding CIS funds within a personal pension. Prior to this, unit-linked personal pension funds were taxed advantageously to CIS funds, and so this removed the final advantage of the mirror fund interface design solution. From now on CIS funds could be held on a tax-comparative basis within a personal pension.

⁴⁰ ‘SIPP’ (self-invested personal pension) was introduced along with personal pensions in 1988 and it has specific permitted investment rules set out by regulation that, for example, permits access to a wide range of investment choices, such as CIS schemes, but also direct holdings in listed stocks and shares, commercial property ownership and other esoteric asset types. It had not really taken off in the product market until adopted by insurance companies and fund management groups c1999-2000. SIPPs are essentially ‘modular’ by reference to the permitted investment standards.

⁴¹ Pensions mis-selling is discussed as an industry level factor later in this chapter

Stakeholder pensions created a challenge for all insurance companies due to the ‘industry standard’ charge cap of 1% per year set out in the regulations (later revised upwards slightly), *“I mean stakeholder came in 2001, but what that really changed for us was just the simplicity of the product – more around the charging structure than anything else”* (23) and *“If you tell people you've got one charge and it is one per cent of your total front value, that's a lot easier for them to understand than some multi-charge product”* (7). Insurance companies, however, had significant problems in making money *“as well as providing a service within those figures, providers had to pay the remuneration to advisors. The early two-thousands was a very advisor-driven market and basically, you started getting commission wars in terms of who would pay the most commission within the price-cap. So, your margins went down. But to still live within your means, you had to bear the pain and reduce your offering”* (7) and *“it did mean that the providers wanted to offer more choice, but they had to within a fixed price. So, therefore, they couldn't give people too much choice, because it would increase the cost if they exercised that choice too frequently”* (10).

Modular unit-linked mirror fund components held within stakeholder pensions could be added, substituted or deleted, provided the price cap was adhered to. This meant that the price cap created an ‘industry standard’ for stakeholder-specific unit-linked mirror funds, however, *“because of the charge cap you were limited to the funds you could offer and this meant that the funds were mainly internal [funds], but some external. I mean the fund range was really quite limited”* (28). As one respondent summed up *“as there was a price cap, insurers had to look at what you could and could not offer within the price cap. So, what you actually found is, in terms of investment flexibility, you were finding products with as few as four funds available, whereby previously a unit-linked personal pension could have had 250. So, basically, it was very much providers having to offer what they could and still make a profit within a price cap. So, as a result, a lot of flexibility and a lot of product features actually ceased to exist”* (28).

With unit-linked personal pensions falling out of favour, and stakeholder pensions proving almost impossible to make money from, insurance companies were stuck between a rock and a hard place: on the one hand stakeholder pensions meant that *“we aren’t going to get any money on this”* (17), while on the other hand, *“the market’s moving more and more towards open architecture and external fund links”* (17). The attractiveness of SIPPs was *“to create a more efficient profitable product base that can you know sell on a lower cost base with fewer people and make more money. You put all of that together, and you get momentum”* (17). SIPPs were therefore considered a product design that could make money and permitted a much wider range of investments that could ‘plug and play’ into the product architecture, *“it’s your mega plug-and-play because you can put any mutual fund in. The charging structure differed by which bit you were in. You could add on commission. It was a real ‘à la carte’ product* (31), and *“so you’ve got a dominant personal pension coming to the fore: which is a full open-architecture product, which again needed a much more stable system in terms of the types of investments that you could hold. And so from that point of view you had product innovation going to the fore; investment breadth going from funds or a wider fund range to proper open-architecture”* (27). So, SIPPs became adopted by insurance companies and fund management groups as an ‘open architecture’ modular pension, considered the natural replacement to unit-linked and stakeholder personal pensions.

Many of the product components within a SIPP can be designed, developed or procured independently of each other. As a respondent recalls *“what we had to decide was how do we do the self-investment bit. So, we thought we need a stocks and shares component, which we went and procured. We needed a discretionary fund manager proposition. So, we basically went to open-market tender and had an initial panel. We needed a fund supermarket hub - we procured it. We needed a commercial property proposition, we had to procure the componentry of that”* (28). What developed was *“a bit of arm’s race to further develop the product”* (7).

Income drawdown⁴² was also a key driver in the development of the SIPP market. Although first launched in 1995, it wasn't until the wider investment choices within a SIPP became available that consumers were attracted to the concept, *“the wedge just got hammered ever deeper with the advent of drawdown. The belief was we needed the SIPP to have drawdown. So ultimately it was about holding shares rather than just funds or property and so on, so it gave you much more freedom. So I think a lot of it is recognition, again a bit like the with-profits to unit-linked transition, it was recognition that we're moving from fund-based personal pensions that have gone from insured funds to more wider range through to open-architecture, where you got much greater band-width”* (17).

By the early-noughties, the 'fund supermarket' architecture had become necessary for any meaningful new product development. As a result, insurance companies licensed IT systems from fund management groups or other third-parties. And, *“the reason for doing that was because as we're launching a quasi-fund supermarket, we thought that their systems would be more appropriate for a fund supermarket”* (12), and *“it was a nightmare trying to build a fund supermarket around the old model and it didn't really work. And that's what drove the decision to get into bed with suppliers”* (12). As the fund supermarket architecture embeds, interface standards begin to emerge across firm boundaries and permeate the industry, *“there was a degree of commonality and we saw that with the development of trading protocols and wanting to be able to trade electronically”* (25), and *“I guess one of the things that changed to make that easier was the introduction of industry standards. Different things depending upon the component, it would've been introduced where everybody would deal with everybody else in the same way”* (1). In some cases, external third-party standard-setting organisations (SSO) emerged to establish industry standards, *“there was an external company which was made up of some of the providers and suppliers who said wouldn't it be a good idea if we could talk to each other in a common language to make things easier”. So they set themselves up as an independent firm”* (1).

⁴² Income drawdown was an innovation that permitted a consumer, in retirement, to leave their pension invested and 'drawdown' an income, as opposed to buying a retirement annuity.

4.1.1.4 The mid-noughties to 2012

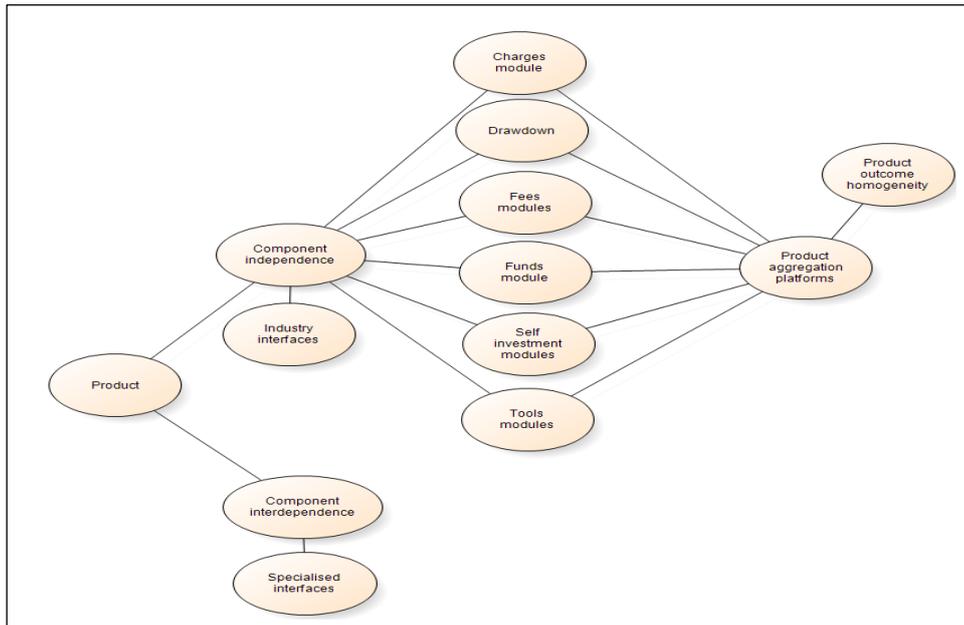


Figure 17: Product template: mid-noughties to 2012⁴³

In the mid-noughties to 2012, respondents associated the period with an open and modular product architecture that represented a ‘product platform’, as shown in Figure 18.

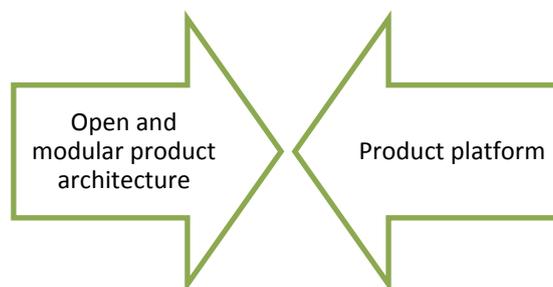


Figure 18: Open and modular product architecture = product platform

⁴³ As also shown in tabular form in Table 25 (section 4.3.2.2)

In this period from the mid-noughties to 2012, the core IT architecture shifted again away from the ‘fund supermarket’ architecture to a ‘product platform’, as a consequence of third-party IT system architects entering the product market, “*we actually procured a new system that allowed exactly that – plug and play. We worked with suppliers, we did due diligence on all those systems, worked out which one we liked the look of, looked at the proposition we wanted to launch, and we wanted our stuff to be able to work with those other systems. It wasn’t, just chuck it to them, but we made our components plug into theirs. We got them in to agree the interface, to see how we administer it, before we cut them loose and off they went. It wouldn’t matter then if they were next door or the other side of the world. The new system had open architecture, it was designed knowing it needed to plug into that system over there, knowing it was future-proofed so that if we wanted to change that with that, we could do it*” (9).

Existing fund supermarket architectures were modularised in respect of fund components, but it was not easy to adapt them for innovation in the product layer or to permit plug and play flexibility to a wide range of other product components⁴⁴. In response, new platform systems architectures were being leveraged from open architecture markets in Australia and New Zealand into the UK and hence “*common components and platforms from Australia came to the UK market; day one it was running and near-ready-to-go. You spent six months customising it; launched it; it was probably quite vanilla in some aspects*” (5). While the new platform systems architecture had a high degree of standardisation, there was already an emerging realisation of risks to differentiation and/or competitive advantage, because “*it was 90% standard. That’s why I think the ‘plug and play’ thing to be honest is too pure*” (5).

⁴⁴ For example, SIPPs permitted access to an extensive range of investment options such as stocks, shares, commercial property, and the market was also demanded a number of decision-based software modules that could not easily be plugged into the fund supermarket-based IT systems

However, the platform systems architecture offered widespread modular partitioning and interfaces to providers of complementary product components, *“one of the things happening across the platform market is platforms offering differing types of investment approaches and investment solutions, and platforms are increasingly saying we are fairly agnostic about the investment choices you make on our platform, we offer a broad range of choices, whether it’s a combination of different funds, a single fund solution, whether it’s using a third party like a discretionary fund manager, whether it’s using an organisation that pulls together a model portfolio, what’s happening is that platforms are offering all of those different investment approaches and solutions”* (2).

In response to demand-side factors, further investment components were added to the platform such as listed stocks, investment trusts, exchange-traded funds, commercial property, etc via standards. For example, *“I think commercial property was a big part of it as well. People were beginning to look at retirement differently. They were retiring but not maybe completely giving up work and you know wanted the flexibility of still investing their pot”* (30). The investment components were also being extended significantly because *“the reason why it was successful, is that investment trends change. So, people were talking investment trusts, exchange-traded funds, target date funds.....no matter how the fad on the market or the regulation changes, it can accommodate those”* (17).

With this explosion of fund and investment component choice, IFAs and consumers began demanding access to a range of decision-based software tools that would help them decide which investments may or may not be appropriate. Whereas in the 80s and 90s, with-profit and managed funds were popular, now many IFAs and consumers wanted more control and choice, but with decision-based software to help them decide between the different options. This led to the demand for a wide range of decision-based software ‘tools’, such as risk-profiling, retirement planning, income drawdown analysis and many more. For instance, *“what’s happening is that functionality and components available in personal pensions has got richer and more diverse, so things like the ability to rebalance a portfolio, the ability to switch in and out of assets, the ability to monitor the performance of assets in the portfolio to a particular risk rating or risk tolerance, all of that has developed over the last 10 years or so”* (2).

The range of fund, investment components and decision-based tools components across a range of pension and non-pension products led to increasing component commonality across the product platform, for instance, *“it does give us a lot of flexibility and so if we wanted to we could offer an incredibly broad range of choice in terms of functionality, in terms of those core components, in terms of investment choice, in terms of different products with different types of options”* (2) and *“most consumers in the UK have lots of stuff they use to save for their retirement and so it’s not just a pension, it could be some ISAs they’ve got, it could be some money elsewhere, or whatever. And, suddenly, here was a proposition that a customer could see all that in one place and make decisions and changes quite easily”* (22).

With increasing commonality, a further number of industry process standards developed, *“Origo, used for measuring service levels”* (5), and fund pricing information *“why would I want to deal with a fund pricer when there are international standards out there? People who do all this sort of thing and they are accepted by the end customer and advisors as a known brand”* (10). In addition, industry standards also developed for process components such as quotation systems, *“there was a growth in industry quotation systems”* (4). Therefore, in the case of modular product and process components, a significant number of industry standards developed, One respondent summed up: *“so, a lot of the basics are standardised. Things like how do I do valuations? How do I do an application?”*(20).

4.1.1.5. 2012 onwards

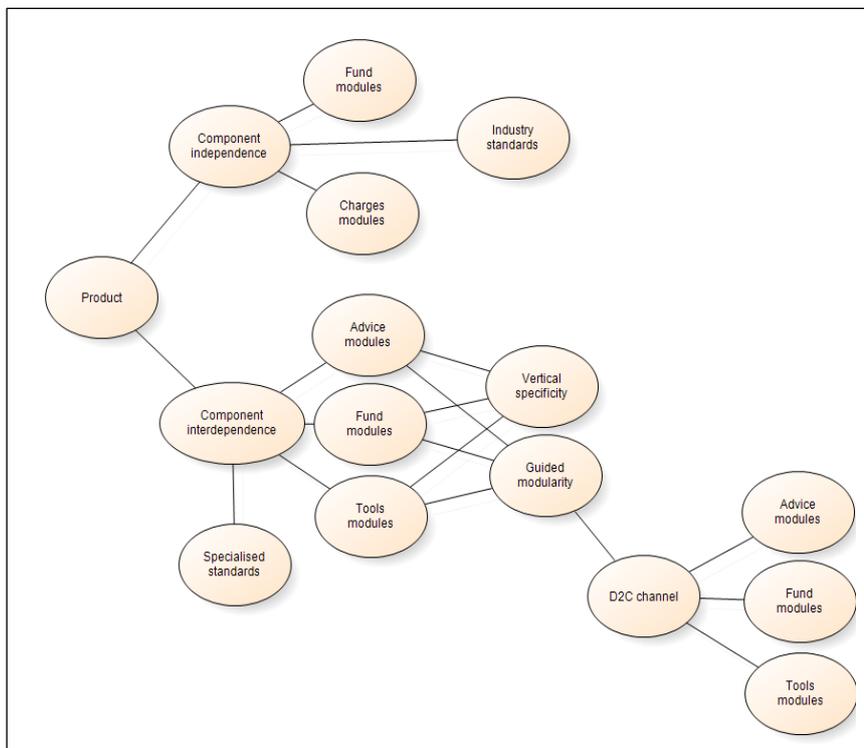


Figure 19: Product template: 2012 onwards⁴⁵

Between 2012 and 2014, respondents associated the period with a hybrid product architecture that represented a ‘product platform’, as shown in Figure 20.

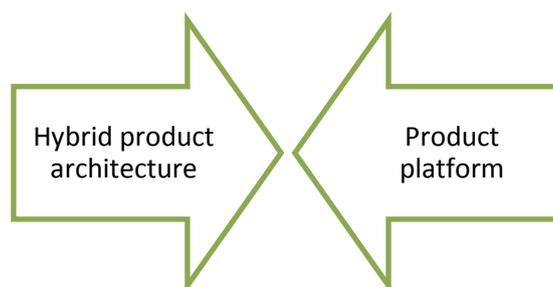


Figure 20: Hybrid product architecture = product platform

⁴⁵ As also shown in tabular form in Table 27 (section 4.3.2.2)

Between 2012 and 2014, the progression towards an open and modular product architecture slowed, and in some cases the product architecture began to evolve back towards the characteristics of a hybrid product architecture. Firm strategy appears to play a role. For example, a few competitors have continued to offer more and more product component choice and variety, but most suggested that an almost infinite product component choice left consumers open to making inappropriate choices, *“the more choice you’ve got, the more like a rabbit in the headlights you become”* (21). As such, there appears to be two emerging positions: *“I think, probably until the last two or three years, firms have been quite nervous of being seen to try and shepherd people one way or another to investment solutions. And some remain like that. You know, some very significant market participants who are genuinely there for folk who understand investments....they will say we don’t care what you buy, it’s totally immaterial to us. Others might say, well no, we do feel some degree of co-responsibility for what’s going on and client suitability”* (29) and *“I think the market tried to go full open architecture. Everybody wants choice – full open architecture. But actually, that’s not what the market wanted at all. Nobody likes complete choice, because you can’t make a choice when you’ve got a choice of everything”* (11).

A re-bundling of product components emerged around the time of RDR⁴⁶ in 2012. For example, some product platforms collaborated with fund management groups to create ‘specific’ versions of a fund component⁴⁷, *“if you look at some firms, they have a different share class now. Same fund but different pricing within their own pension, and consequently you can’t transfer that fund, or plug and play that fund outside. You would have to effectively take the same investment management approach but the pricing would be different elsewhere. You’ve effectively re-bundled the proposition* (21).

⁴⁶ Retail Distribution Review – a regulator initiated review of distribution and platforms

⁴⁷ For example, by creating different share classes with different management charges

Firms are also embedding proprietary fund components within sets of product component bundles, *“an example is where the platform provider or manufacturer has also got a vested interest in having their own manufactured funds sold. So, there are lots of examples where platforms are choosing to make their own fund solutions more attractive, that might be virtue of their price, by virtue of the way they are marketed, by virtue of the prominence their given, and the clear motivation there is for the platform to generate more income for its own in-house fund business as well”* (2). The re-bundling of fund and investment components had knock-on implications for the coupling characteristics of decision-based software components. For example, manufacturer/assembler firms and decision-based software tool providers have collaborated, in some cases, to create decision-based software tools with specific marketing outputs or specific stochastic models that ‘recommend’ specific bundles of proprietary fund and investment components. For example, *“they’ve got a specific risk modelling system that plugs into a specific fund”* (15). The re-bundling of product components appears to be a form of ‘guided modularity’ where the underlying core product components are loosely-coupled and independent of each other, but the manufacturer/assembler firm ‘guides’ the consumer towards a specific bundle of product and process components that, perhaps, have the greatest rent appropriation opportunities or best meet consumer needs.

Guided modularity is considered attractive because *“initially it was thought that these components are entirely independent, you put them together and you take that to market. But actually the putting of them together doesn’t add that much value. I think it then became quite clear that actually the value was in the way you could integrate them for the process of the client and the advisor”*(6), *“full open architecture, it came back from that and then that ‘coming back’ point was driven by demand and attractiveness to the market”* (11) and *“I think here of guided architecture”*(28). The ‘guiding’ remains within the framework of an open and modular architecture because *“even if they don’t use 99% of the funds, they want to retain the flexibility to do that, but you’re seeing a big narrowing down”* (21) and *“so as subsequent waves of legislation have come in, I think providers are starting to think of ways to not narrow the choice – I think there’s relatively few who have narrowed it – but provide an extra guided step”* (28).

In many cases, guided bundles of product and process components have increased in specificity, where a bundle of product components within the architecture is made 'specific' to a particular value chain configuration. For example, a manufacturer/assembler firm may bring together the systems architect, an IFA firm, one or more fund management groups, one or more decision-based software tools providers and possibly others to create a specific product component bundle for that particular value chain configuration. For example, *"[Organisation] are a classic example where they have a very tight proposition, as long as you don't put square pegs into round holes. You've effectively got an administration system, a financial planning system, an investment proposition, tax wrappers, all tightly integrated together and within a single price, effectively"* (21). Whereas product platforms were once competing against other product platforms, the emergence of specific value chain configurations has created *"power blocks in the market, almost vertically integrated fully but through technology and other relationships. So what you end up with is relatively large power blocks. They're not as integrated as the old provider ones, but they're still big power blocks going against each other with the platform at the core of it. So we see ourselves as a competitor set and looking at different power blocks, and they come from different places. So what you are seeing in the market is the platform providers increasingly in concert with the fund groups to build out those power blocks"* (6). The respondent continued *"so my sense of it has been a period probably five years or so where disruptive technology and competitors came in and broke up the old provider power base, and you had a very fragmented market. What you're now seeing are those power blocks trying to reform configuration"* (6)

As a result of these processes, industry standards are also being challenged. For example, *"there's been a tendency for a lot of businesses to be very similar, and almost the economics tried to create common interfaces and common fund ranges, fund agreements and structures, but that's all commoditisation, it's then difficult to have something that looks distinctive"* (18), or a reluctance by organisations to engage further in an industry standard setting process, *"I mean a lot of those things are still progressing now, they've been extremely slow to emerge but whether that has simply been less interest and people not wanting to enable things to happen or efficiently in the industry, or it wouldn't be in their short term commercial interest. I think certainly it made it very difficult to get the players to co-operate and so I think across industry new standard-setting is very slow"* (18).

Because of the need to continually innovate, few firms have committed to prioritise further investment in the standard-setting process, for instance *“if we don’t like the standard user interface, we build it ourselves. So to innovate you have to go down that route, and we realised you could more effectively do that yourself”* (18) and *“industry standards are a barrier. I suppose the reason is that it’s sort of related to people, people develop standards, and implement them, there’s never really a true standard, there’s always someone who works to distort it. There will always be a need for people to talk because you can standardise everything there should always be a next innovation around the corner, there will always be a need to develop new standards”* (20).

4.1.2 Firm-level analysis

4.1.2.1 Mid to late-1980s

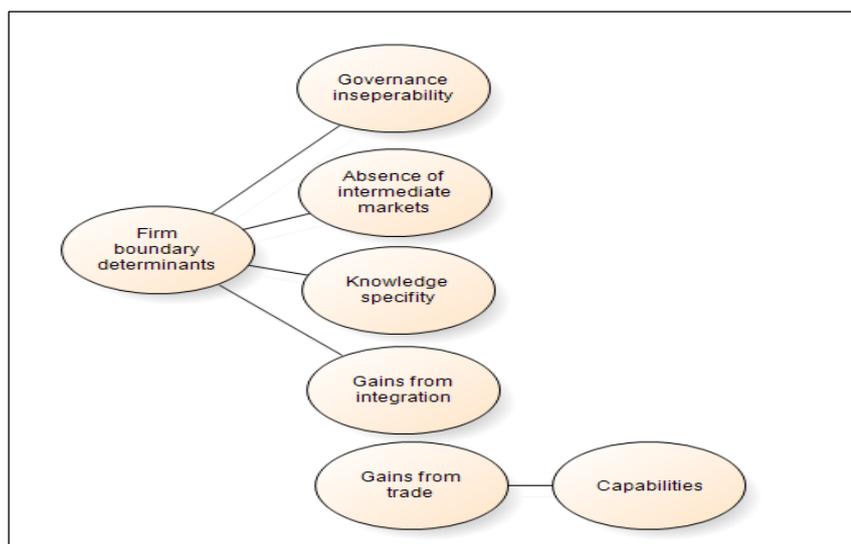


Figure 21: Firm template: mid to late-1980s⁴⁸

In the mid-80s, many insurance companies exhibited vertically integrated characteristics. For example, the “advice” and distribution component⁴⁹ was enacted within firm boundaries, “*at that time direct sales forces were more common, so you were looking at something much more vertically integrated, and more face to face*” (3). One of the advantages of internalised advice and distribution was “*they were expensive to build but you got all of their business*” (25). Integrating advice and distribution within firm boundaries was also used to aggressively sell new products. For example, new unit-linked insurance companies primarily used integrated advice and distribution processes to quickly gain market share, as one respondent recalls: “*aggressive integrated sales forces that were entirely unit-linked. They drove the unit-linked product because that was all they did*” (24).

⁴⁸ As also shown in tabular form in Table 19 (section 4.3.2.2)

⁴⁹ For example, the people and processes who advise consumers and distribute products

However, integrated advice and distribution processes were quite expensive to set-up and maintain, especially in terms of recruiting, training and development, but also in terms of regulatory capital, *“this was going to be a big burn on capital”* (25). In the mid-80s, while integrated advice and distribution was a common governance mode, firms were beginning to explore other modes such as outsourcing distribution to intermediate markets, such as IFAs, for instance, *“the company I worked for had a tied sales-force, but it also distributed through independents”* (25) and *“at the point in time it was either direct sales or intermediated, but independents was a way of doing something different from a distribution point of view* (27).

Fund management was also enacted within firm boundaries, *“certainly for the early part of that period, most of the funds were in-house funds”* (4). In addition, the variety of fund components tended to be limited, often a choice between a with-profit fund and a managed fund, *“quite limited initially, so you tend to have maybe a with-profits fund or managed fund. The investment links were quite limited in terms of the range and they tended initially to be in-house investment management”* (26).

Firm boundary decisions were associated with a number of key factors: (1) governance inseparability, (2) the absence of efficient intermediate markets, and (3) the specificity of the firms' knowledge. In terms of governance inseparability a respondent recalled that *“I think it was just the era of insurance companies, people didn't tend to outsource things in those days. It was just after the sort of like black suit and bowler-hat phase of the City. That's how they'd always done it. And they'd always done it on an in-house basis”* (8). Intermediate markets were on the whole limited for most product components, for example, in the case of fund management, *“there was very little and it tended to be very small”* (26). A final factor that is associated with vertical integration is the firms' architectural and knowledge specificity, as few people within firms understood how the product architecture actually worked, *“if you go back before 1988, you didn't even know what the charges were, it was all in the product and the bonus rate”* (1) and *“you can sell and market on different features, you need to have marketing people who understand it, and you couldn't do that with conventional with-profits. It was a hugely complicated product that people internally didn't really understand”* (1). One respondent summed up that *“in general it was all about internal manufacture, a web of systems with complexity which means it's only Geoff in the cupboard who knows how to fix them”* (20).

In the late 80s, however, while most firm boundaries continued to exhibit vertically integrated characteristics, outsourcing advice and distribution grew, and outsourcing fund components also emerged as some new unit-linked insurance companies outsourced fund design to external fund management groups in the intermediate market. This was primarily driven by the availability of specialised productive capabilities in the intermediate market, “we didn’t do it because we suddenly had this blinding flash of insight – we did it because we had a terrible investment record” (24). As one respondent summed up, the productive capabilities of in-house fund management was so poor that “in the mid and late-80s people started saying maybe in-house insurance company fund management guys aren’t the best people to manage money” (25).

4.1.2.2 Early to mid-1990s

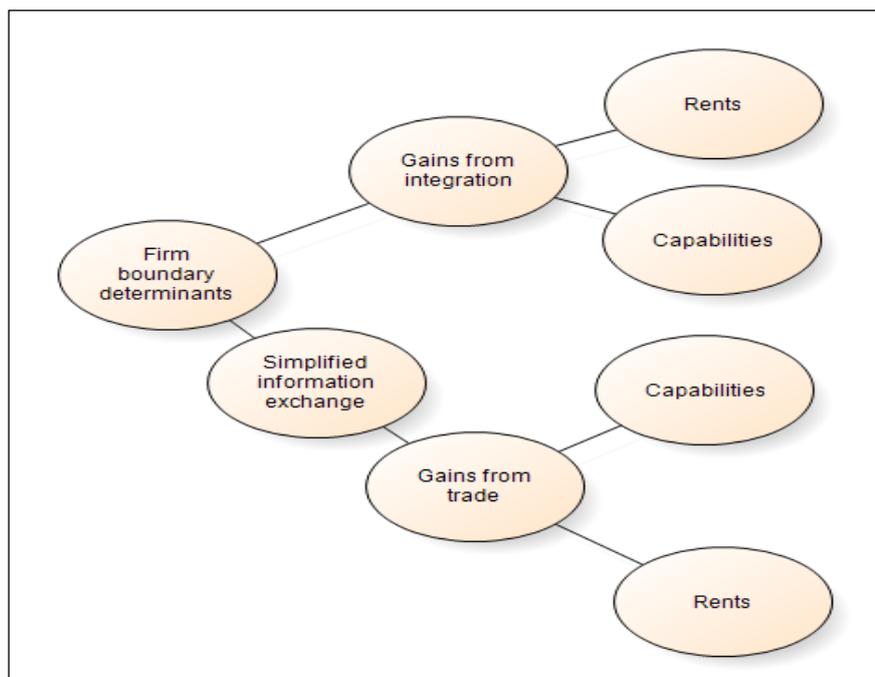


Figure 22: Firm template: early to mid-1990s⁵⁰

⁵⁰ As also shown in tabular form in Table 21 (section 4.3.2.2)

While an intermediate market had grown for advice and distribution and fund components, insurance companies remained vertically integrated in terms of IT systems architecture, customer services, and administration. One of the primary factors associated with insourcing IT is the sheer complexity of the systems and the knock-on implications for product design. For example, *“getting the system to talk to other systems without developing a whole brand new system was impossible. The system was very old – maybe it could have talked to other systems but the people were not around who built it to have the conversation”* (9).

Efficiency considerations also play a role, *“Cost came into it. More cost to use suppliers. It was our product, on our system, it was our sales force selling it, our funds, managed internally, and you can manage the costs better that way. As soon as you start outsourcing different components, a) you’ve got the initial costs of building the things, and allowing everything to talk to each other, and b) you’re kind of exposed to the costs of that third party, you don’t have the same control over those costs (1), and “I suppose it was about control, the lack of control you had, and the risk you were building in, the risk being that if the costs do go up, how easy was it back then for you to change suppliers. Once you’re tied into a supplier, you don’t have the capability of plugging another one in, you were kind of linked in to the heart of the supplier, it was more of a risk really”* (17).

In order to efficiently use intermediate markets, it was recognised that there was a need to simplify information exchange and reduce complexity. For example, the scale of the product development issues are highlighted *“if you try to cut up the pie and see all the effort that went into the product, 80% of it or more would be the systems build and a small amount up front trying to decide what the product looks like in the first place. But the vast majority was on the systems build so that was the critical path on all developments”* (4).

With the introduction of modern personal pensions in 1988, and hence exponential growth of unit-linked product architectures, speed to market became a significant strategic issue and simplifying the product architecture and so facilitate the use of intermediate markets became critical. For instance, *“we don't want it to cost twice as much because you're componentising it, but it's not actually about cost, it's the timescale we're worried about really”* (4), and speed was one of the main factors in simplifying the interfaces for fund components, *“where we've simplified the fund addition process, it was time to market. I think cost and time were embarrassing and there's no doubt about it, you felt like a big clunky organisation, it takes a long time to get something to market, losing market share. So I think time to market was pretty key and the overall cost is pretty key and there were some eye watering costs”* (4). The respondent continued *“Yes, I would have said internally whether its funds or the whole system, the idea of a componentised model would make things easier and we could just link these components together to make the whole development easier, and the battle was how do you change this great big monolithic piece of code into a componentised thing?”* (4).

Accessing external specialised fund management productive capabilities was another key factor in driving insurance companies to adopt unit-linked product designs. As one respondent recalls *“what we'll never be able to do is be a top investment group in every aspect for all scenarios; so what we want to do is to offer expertise that we don't have, necessarily on a wider basis from fund management groups who know how to manage money”* (26). As such, *“the hypothesis was that you would not get as good investment performance as you would if you outsourced to people who are experts in fund management in different asset classes and different countries”* (25). Rather than being specialists, insurance company in-house fund management teams were often ‘jack of all trades’. This lack of specialisation presented a problem, *“they were trying to research Japanese equities from the UK, how on earth do you recommend a buy or a sell of a Japanese equity if you've never been to see the directors of the firm? All you're looking at is the report and accounts. You probably aren't big enough to even pick up the phone and talk to them and have them talk to you because they'll probably go: “who the hell are you?”. So, actually what you need is either local fund managers in the various markets for equities or firms who are experts in a particular asset class. Whereas, you've got this sort of 'jack of all trades' fund management business sat inside the insurance company”* (25).

However, while the unit-linked insurance companies collaborated and outsourced to upstream fund management groups with relish, the traditional with-profit based insurance companies were slow to follow, continuing to focus on their own productive capabilities in ‘with-profits’ fund management. As one respondent suggests, *“in that initial stage, they were the traditional players and therefore unit-linking wasn't seen as being necessarily immediately something they wanted to get into”* (26).

4.1.2.3 Mid-1990s to mid-noughties

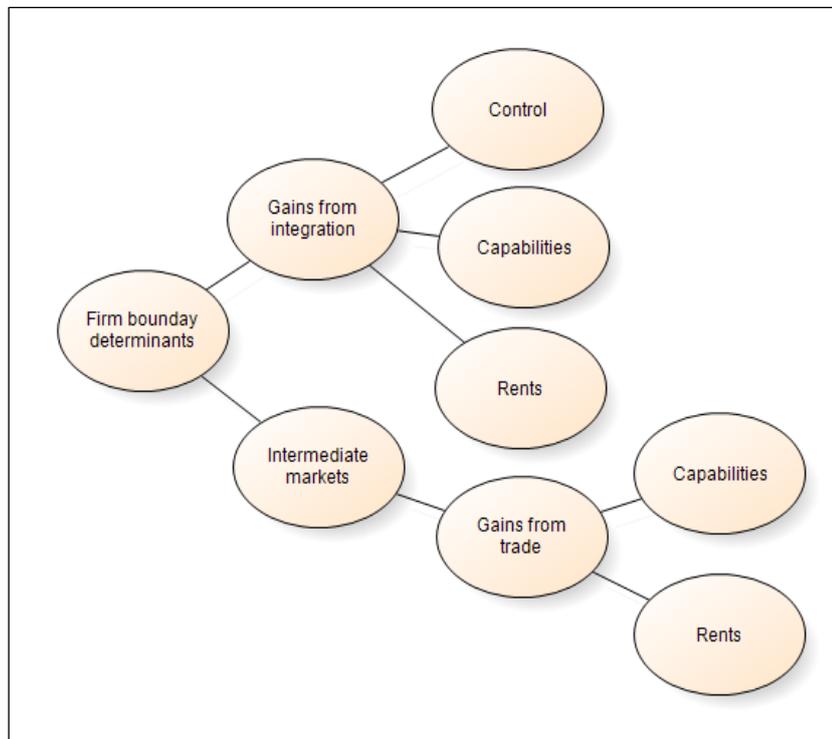


Figure 23: Firm template: mid-1990s to mid-noughties⁵¹

⁵¹ As also shown in tabular form in Table 23 (section 4.3.2.2)

Outsourcing product components such as funds and advice and distribution continued in the period of the mid to late-90s. In terms of outsourcing the fund components, gains from trade were associated with growing scale economies that supported ex-post negotiations with fund management groups, *“the margin that we had to give away was negotiable downwards on a growing basis”* (26) because insurance companies gave fund management groups access to a product market that they could not access directly, *“we’ll want to negotiate and review on a regular basis. You’ll get something, but something that’s less than you thought. But you’re getting 100% of something as opposed to nothing”* (26).

One respondent suggested that *“it wasn’t a cost-driven thing. There’s a marketing opportunity here, there’s an opportunity for us to differentiate what we do as opposed to what other people do, produce some more value for the customer and therefore gain market share so ultimately get a return for the shareholder”* (25). Another respondent recalled that a growing market and scale continued to support ex-post negotiations and hence gains, *“initially, we paid the fund managers too much. We got wise to that and we squeezed them down and down. So we were retaining a very significant part and what we did was we expanded the cake. So it became much more profitable. So we made lots of money during that time* (24).

Insurance companies were also withdrawing fully from advice and distribution and outsourcing to IFAs, driven by increasing bureaucratic and production costs, *“originally, it was just the fund management that was outsourced. What you also saw in the mid-’90’s, you began to see product manufacturers withdrawing completely from distribution. I think that was simply the fact that shareholders were waking up to the fact that these beasts of direct-sales forces, they couldn’t see how they could make money out of them. They were waking up to the fact they couldn’t run effective distribution. So, they began to withdraw from distribution and so that began to get outsourced”* (25) and another respondent associated the process with regulatory costs, *“definitely the economics of running those businesses, then the first wave of regulatory scrutiny”* (18).

By the early and mid-2000s, the process of de-commissioning existing IT systems began which, in turn, signalled the eventual end of the unit-linked and stakeholder personal pension product design. In its place saw the emergence of a ‘fund supermarket’ product architecture based upon an industry-standard fund architecture. In addition, with a pension simplification review⁵² announced by the regulator for completion in 2006, signalling an imperative for all incumbent firms to embrace a new era of ‘flexible’, transparent and modular product designs for the personal pensions market, and speed to market was again essential.

With limited productive capabilities in ‘modular’ IT architectures, one motivation for insurance companies to license-in the fund supermarket architectures was timing - the momentum and volume growth in ‘fund supermarket’ designs between 2000 and 2005 was huge, largely driven by the recycling of existing personal pensions away from insurance companies and switched into more flexible and modular pension architectures, such as SIPP. The unit-linked mirror fund standards perpetuated by insurance companies were now seen as ‘old-fashioned’, expensive and non-transparent. By comparison, with-profits as a concept was seen as a relative ‘dinosaur’ of another era. Being slow to adopt the new fund supermarket architecture, incumbent with-profits insurance companies also had some significant catching-up to do, *“it was pretty much speed to market, because a lot of organisations misjudged what pension simplification would really mean. But I think as the market moved a lot of people realised that they were now 18 months behind a competitor and couldn't make up 18 months without actually buying in the technology”* (7). Cost also played an important role, *“it all depends on the cost. I think originally a lot of companies keep stuff in house because it was perceived to be cheaper. Then because of the rise in technology, and how other companies had embraced it, the life companies then realised that they can do it quicker and cheaper, so it's a commercial decision, we can retain our profit and still get what we want but we just outsource it”* (7).

⁵² This was known as A-day, a pensions simplification date, whereby all existing pension rules were unified or ‘standardised’ into a single set of rules. A day was set for 2006, but was announced as early as 2003/4

The intermediate market was also growing in terms of providers of complimentary product components beginning to converge around the fund supermarket architecture, *“it was very much growing, the necessity of the open architecture, so you take the progression, you started with with-profits, you’ve then got unit linked insured funds which were effectively the insurance company going out to buy the underlying funds to offer to its customers. So that increased the supply base. Then I think we went to the ISA and SIPP world, it then exploded out again because actually those components were entirely interchangeable. And we were then in a position to offer everything, and actually what you very quickly saw was an explosion of a supplier base”* (6).

Another respondent recalled that *“lots of suppliers saw it as an opportunity to get in on the pensions market, whatever you were supplying, whether it was software or funds, everyone wanted a bit of the value chain, and all it did was really was break up the value in there to lots of different bits. I think the supplier market was becoming competitive, as a provider you’ve now got a choice of suppliers and they’ve got to compete with each other, so, whether it was IT supplier or a fund provider, as a product manufacturer I now have a wider choice so they had to compete and their prices come down, and that reduced the risks. So, the fewer product manufacturers there are, the fewer suppliers there are, the less competition there is, the bigger that risk is, and prices change”* (1).

4.1.2.4 Mid-noughties to 2012

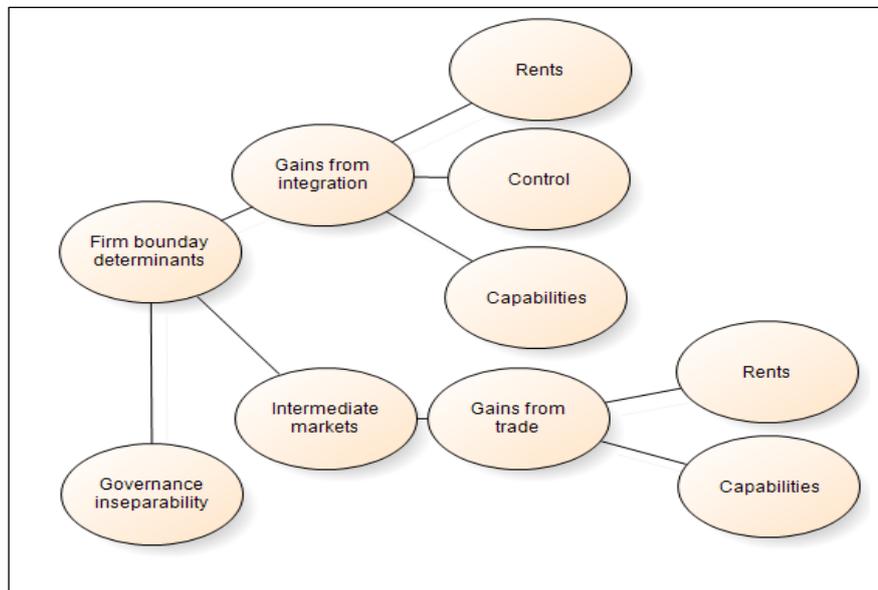


Figure 24: Firm template: mid—noughties to 2012⁵³

By the mid-noughties, the ‘fund supermarket’ architecture had become a dominant design, and yet it was superseded by an IT platform systems architecture which arrived in the UK via the entry of a handful of systems architects primarily from Australia and New Zealand, where multi-product platform architectures were already dominant in home product markets. Owing to its open and modular design, insurance companies “*pretty much just handed everything over*” (9). This was because “*running a core platform technology needs investment, and unless you’re going to hire the developers, you’re better to find them. We outsource a lot, and they spend billions of pounds in core operating infrastructure so scale and quality comes from outsourcing*” (5).

⁵³ As also shown in tabular form in Table 25 (section 4.3.2.2)

The emergence of a new open and modular systems architecture established an efficient intermediate market, and incumbent insurance companies could leapfrog the fund supermarket architecture because *“it costs us a lot of money to keep up with legislation and it's just not efficient anymore to make all of those changes in-house. We are not a specialist technology provider. You know, it makes more sense for us to use the likes of [organisations] where they're looking after four or five different providers. They make one change at one point. If they make the change once and, you know, we pay a fifth of it because we're splitting it between the five people who are using them, how can we compete with that if we're trying to do it all in-house?”*(23).

Efficiency considerations were also a significant motivation, *“technology is really all about scale. The technology is scaled so quickly and that's the challenge. So technology creates massive scale and it creates massive winners and losers really quickly. But also there's going to be casualties there. So I think that greatly influences the level of suppliers”* (5). As the new systems architecture embedded, the growth of complementary product and process components that could plug and play into it also explodes, *“in terms of the component technology, we saw quite a lot of innovation in that space”* (20).

Furthermore, the use of modular and technology-enabled IT architectures meant that *“it is operationally most cost effective for us to do it that way. We are writing a lot more business but driving more transactions online because of cost”* (28) and *“this was a period where everybody was going through lean and Six Sigma and everything and driving costs out of the organisation”* (4) because *“the fact that the with-profits and unit-linked period was well and truly over, but the organisations hadn't actually slimmed down at this stage and now they were going through the pain of slimming down to get to the lower margins”* (4) and *“CEOs were saying “I've now got variable costs and less of a fixed expense company””*(8)

Productive capabilities also played a role. Upstream firms in the intermediate markets had superior and specialised productive capabilities in modular and technology-enabled IT systems. For example, one respondent simply recalls *“we absolutely did not have the expertise to build it. It was getting to the point where it was just not sustainable. This old culture of “we build everything”, actually things were changing so rapidly. Technology was changing and requirements were changing so rapidly, we just didn't have the expertise and it became a massive bottleneck”* (30). Productive capabilities around the systems architecture created a significant challenge for incumbent insurance companies, as one example illustrates *“we didn't have a core competency around systems so we didn't have a new world thinking in that type of technology. The strategy was we knew we wanted to be in the market, we knew we didn't have a core competency around it. We knew it would be expensive to do it ourselves. Somebody else needs to do it for us, so let's pick somebody who's good at it* (22).

New modular and technology-enabled IT systems threw everything up in the air, and incumbent insurance companies and ‘fund supermarkets’ needed to make strategic decisions about what to outsource and what to retain insourced. As one respondent recalls *“so I think dis-intermediation is definitely going on and there's re-carving up of the value chain”* (4). With many advice and distribution and fund components already outsourced in earlier time periods by the mid to late-2000s insurance companies had outsourced the core IT systems architecture to new-entrant IT platform systems architects as well, while seeking to retain within firm boundaries the service and process components that they believed offered some limited differentiation or competitive advantage, for example *“the outsource is almost a given. We wanted to keep the customer experience, the front-end, we didn't want to outsource that”* (21), and *“contact with new customers or current new business customers we'd like to keep within the organisation”* (4).

Some gains were still appropriable through foregrounding the marketing and distribution of proprietary fund components⁵⁴. So, while the fund and investment components were to a significant extent outsourced to hundreds of multiple upstream fund management groups and other providers, manufacturer/assembler firms had a commercial interest in directing consumers and IFAs towards its own proprietary fund components, *“I mean clearly it's in our interest if they invest in our own funds. The margins can be very different” (28)* and *“the big life insurers, with their own funds in the products were getting revenue on the underlying fund in one part of the business and the product wrapper at the other side. And if you were lucky, both parts were making money. And if you were less lucky, the fund charge was probably covering a lot of the expenses overrun on the pension wrapper” (11)*.

4.1.2.5 2012 onwards

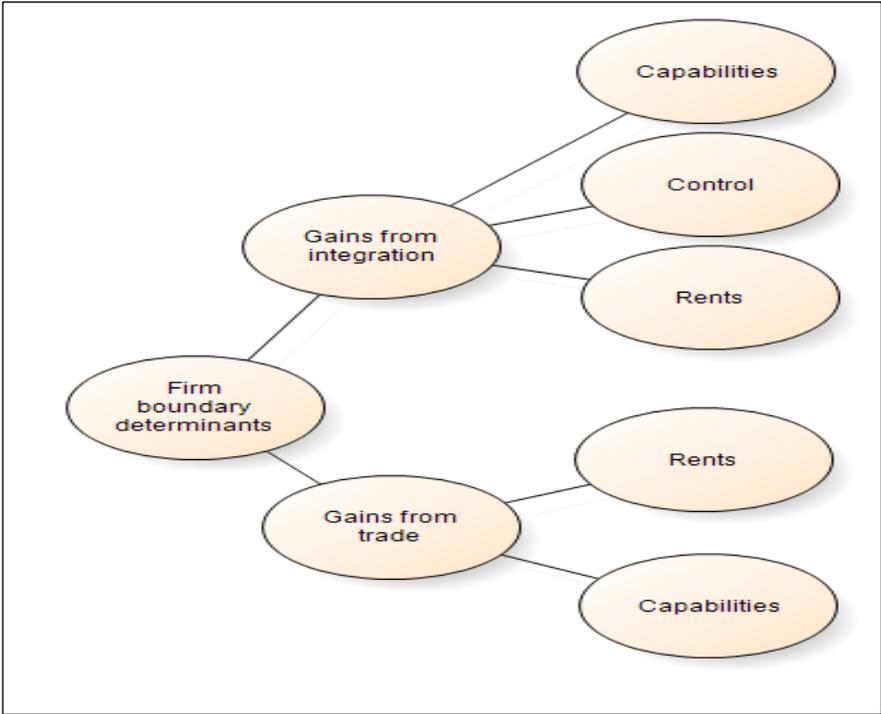


Figure 25: Firm template: 2012 onwards⁵⁵

⁵⁴ For example, because the firm was able to secure ‘manufacturing margin’ from both the product and the fund

⁵⁵ As also shown in tabular form in Table 27 (section 4.3.2.2)

As an open and modular product architecture became dominant by the late-noughties, manufacturer/assembler firms came under increasing commercial pressure to differentiate and seek competitive advantage – gains from trade were diminishing and many plc firms were under pressure from analysts and the investor community. In response to the known threats of commoditisation and imitation, most firms began to consider mechanisms to improve profits *“how do you add more value back? Margins get squeezed so, so tight, that somebody in the value-chain's got to do something different (15) because “there's a footrace to the floor, which is clearly doing damage” (14) so that, “you can certainly see some of those companies taking back bits that they would think offer differentiation, even though that worked out fine with outsourcing, taking some of those back in house to try and find different levels of differentiation there” (5), and “It's all commoditisation, it's then difficult to have something that looks distinctive” (18). As such, “as platforms become bigger, the challenge is why outsource if you can do it internally? Why pay the extra to somebody else? So there's a bit of a balance between getting the market standard, best of breed onto a platform versus can we do it in-house?” (10).*

In terms of which activities to insource within firm boundaries, different manufacturer/assembler firms seem to have different productive capability trade-offs, dependent upon the product component type. For example, one respondent suggested that *“providers recognise there are things like commercial property. Some companies have now specialised in how they manage commercial property in pensions. We currently don't have the expertise, so we outsource it” (14).* The respondent continued *“we say we've got great customer service at the front end, but the back end it can be delivered by those companies that can run it and do it differently to what we do. What you're faced with is “can I do all things brilliantly?” Probably not. “Can I do all things brilliantly and efficiently?” Definitely not. So, you then have to pick and choose what you're good at. So then, you'd go, “we're really good at trading. What we're not great at is running property funds; or what we're not very good at is doing self-investing, or commercial property”. So, I buy that service in” (14).*

Fund components were, again, perceived as one significant source of differentiation, competitive advantage and rent appropriation. While the fund architecture remained largely open and modular, manufacturer/assembler firms took significant steps to foreground their own in-house fund components, for example *“having your own funds, you take a bigger margin”* (24). They also created selective fund lists, including many in-house funds, and these lists were often integrated within the products in order to increase revenues, *“where can we make extra money?” We've got our extra special funds, our elite range or whatever”* (24). This gained *“real impetus 3 or 4 years ago and coincided with a big increase of business for platforms, there was an expectation that they would become easier and cheaper and so what's tended to happen is that margins in that area have been compressed, and one of the ways of the providers overcoming that down-squeeze in margins is to get more fund flow into their own in house funds”* (2).

The product components resourced within firm boundaries, however seem primarily associated with productive capabilities, *“we insourced because we thought we would be better at it”* (22), *“so keep the stuff that you're good at in-house. And decide what you're not good at and stick it outdoors”* (14), and *“if they think that a particular process or component is where you get the majority of your margin, this might influence them on where you insource. But, yes there will be a capability issue there as well. Companies that have the investment capability behind them are naturally going to insource that and say ‘we can build our own risk rated portfolios; we have the expertise, so why wouldn't we use that?’ Obviously, if you haven't got that, but might have software capability, you might insource that side of things. (1).* Furthermore, *“what's been happening really is a lot of providers have realised the sort of things they outsourced originally, from about 2006 onwards, they realised that they could actually do that function themselves now. And now that companies have changed structures and the way they work, a lot of the stuff which was going outsourced is actually being brought back in house because it can be done for a cheaper unit cost. I think it's the mind-set in terms of providers now seem to be more happy to actually spend on infrastructure for the long term rather than trying to do the short-sharp fix”* (7).

It also appears that productive capabilities have been strengthened in this period either through recruitment or through acquiring product component suppliers, one respondent recalls *“We’ve built our capabilities and built out internal capabilities, rather than partnering. I think largely bringing people in from other businesses that have those capabilities, you’re piecing together the capability”* (18) and *“some brought investment firms; they just saw some benefits in buying investment capability. At the other end, some brought distribution back in very expensively. (16).*

4.1.3 Industry-level analysis

4.1.3.1. Mid to late-1980s

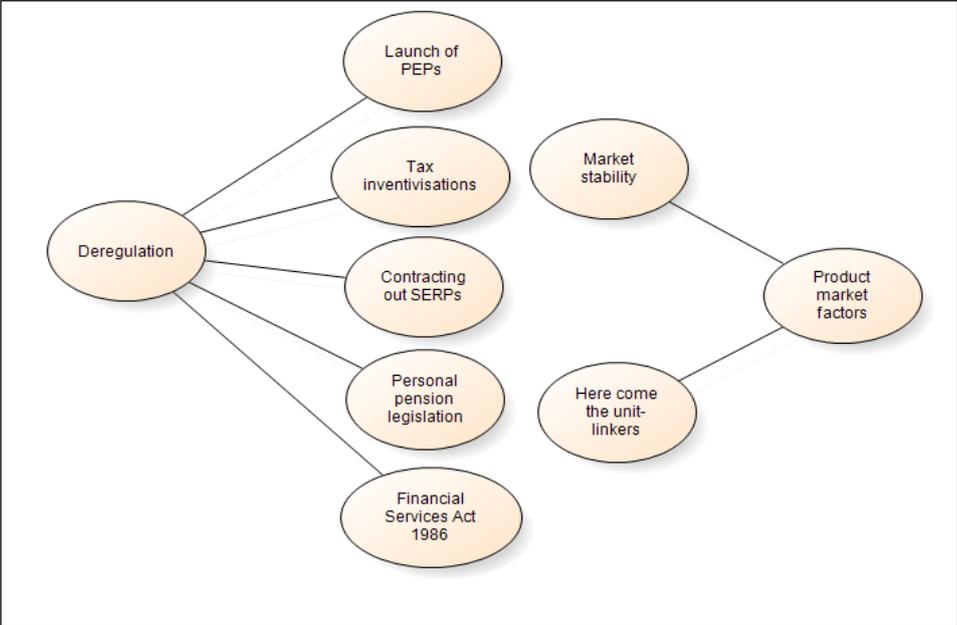


Figure 26: Industry template: mid to late-1980s⁵⁶

Product market stability characterised the industry at this time. In the mid-80s, the personal pensions product market was focused on professional investors and the self-employed. For instance *“the market was quite small, so it was focussed usually on professional types”* (8) and *“I think there were quite a lot of people in the marketplace, but the volumes that were being sold weren't particularly high because the market was really confined to professionals and very few other people - self-employed effectively”* (8).

⁵⁶ As also shown in tabular form in Table 19 (section 4.3.2.2)

The mid-80s saw the market entry of a number of new unit-linked insurance companies into the personal pensions product market, having used Executive Pensions from the mid to late-70s as a thin end of the wedge, “so, I suppose [organisation] broke the mould. They were in executive pensions so they were a challenger-brand” (21). One of the key factors in entering the UK personal pensions product market at this particular time was the low capital requirements for unit-linked insurance companies, “we set up with capital of £60,000. That’s all it took to set it up £60,000. Because there were no solvency regulations” (24).

The deregulation of the personal pensions product market in 1988 plays a significant role in the timing of entry and the growth of unit-linked personal pensions. The then Conservative government played an active role in the promotion of modern personal pensions and ‘contracting-out’ of the state pension scheme (SERPs). One respondent recalls, “in 1984, the legislation was first muted, but eventually in 1988 the personal pension was introduced. From the start was primarily unit-linked. The Conservative government didn’t want to be involved with pensions. They were being crippled by the State pension as it was. So, they said this has got to be done by private industry and we’ll just let it get on with it. And so they introduced the personal pension” (24). Another respondent recalled, “in 1988 we had the introduction of personal pensions. We had the Government advert “Breaking the Chains”. It was also said “get out of your defined-benefit schemes, because they’re rubbish and you’ll be able to understand personal pensions”. The context at the time was that there had been the “Big Bang”; the stock markets had just opened up to the public; people were buying shares, floatations and privatisations were king. And so, everybody was interested in making a fast buck on the stock exchange and the personal pension market effectively got behind that” (8). Or, as one respondent flippantly puts it “people were told ‘you need a personal pension, come out of your all-singing, all-dancing, defined benefit scheme, where you take none of the risk, where your employer takes all the risk, you have none of the downside, you’re gilt-edged pension with inflation-linking for the rest of your life, you don’t want that, you want a personal pension where you’re in control of it’. That’s what the Government said effectively. So, there were even TV campaigns about it and people suddenly had all the responsibility for picking funds and bore all the risks of poor performance” (9).

At the same time, the Conservative government was promoting the idea of opting-out of the State-Earnings Related Pension Scheme (SERPs). Under this initiative, consumers could transfer money built up inside the State pension scheme and transfer the whole amount into a personal pension, and invest it into the stock-market, as one respondent suggested *“you have to remember a lot of them in the market got fired up by SERPs contracting out, so when the market was created, people bought products who wouldn't have necessarily been in the financial services market beforehand”* (26). To help ensure the personal pension product market took off, the Conservative government also offered tax incentives for contributions to a personal pension scheme, for example *“tax relief at source, that was a huge swinger for many customers”* (27). The government also opened up who could offer personal pensions because *“up until 1988, the only people who were allowed to provide pensions were insurance companies. Before 1988, nobody else was allowed to provide a pension. So there was a monopoly for insurance companies in the pension market”* (25)

The Financial Services Act 1986 also came into being on 1 July 1988. However, it *“was unfortunate that at the moment personal pensions were introduced was also the moment when the Financial Services Act came into being. Again the Government didn't want to get itself sullied with regulation of financial services. So, the industry was going to organise how it was all going to be done. And they came up with the idea of self-regulation. And the rather naïve Conservative government swallowed it hook, line and sinker. But what that meant was that the industry captured the regulatory process and wrote it for its own benefit. It was 1998 when it was implemented. They hadn't put in place any monitoring at all. They didn't really know what they wanted to be monitored. So they ensured that all the regulation was biased”* (24). Ultimately, the Financial Services Act 1986 regulated financial advice and distribution, with the result that *“sales-forces started getting smaller and smaller and it went towards independents”* (12).

The motivation for insurance companies to outsource to independents was that *“a tied sales-force automatically carry risk. From that point of view, if you're selling as well as administering as well as running funds, you carry risk in all areas. Whereas, if you're just segmenting the value chain and just focussing on a key component, there's still money to be made by specialising in a certain part of that value chain”* (27). As the product architecture began to evolve towards a unit-linked design, there were also productive capability issues in distribution, *“the tied guys were coming in trying to learn the basics of the new unit-linked pensions product. I can remember quite a lot of pressure coming down from above to get those people through those tests so that they can you know...“no matter how you do it” to get them on the road, because there was money to be made from people selling these products! That's the kind of market it was in those days!”* (8). However, *“there was an awful lot of intermediaries that were previously direct sales, providing more sophisticated investment advice to customers. So the shift is starting to get into investments”* (8). Ultimately, this shift had important implications for product design as, *“if you've only got four funds, what's the IFA got to do? He can't really justify a greater income, if he can only actually recommend from four funds”* (27). In other words, demand for variety from downstream IFA firms also played a role in fund component modularisation.

As consumers were being urged by Government and the industry to take accountability and control for their own personal pension provision, *“increasingly people were attracted to the idea of being responsible for their own futures and taking responsibility for their own financial affairs”* (25). There was also a motivation from consumers to participate in the stock market, *“every week there was a new IPO. But that was the mentality. There was an increasing interest in the population in being responsible for their own wealth management. And I think unit-linking was a reflection of that”* (25). With-profits investments just didn't offer the potential upside of unit-linked funds linked to the stock-market and *“people didn't want to miss out on the upside”* (26), and *“stock markets sort of kept on going up and up and up. So, they could sell on the basis of look at our equity funds – vroom! Fantastic, and it was all going into unit-linked”* (24).

4.1.3.2. Early to mid-1990s

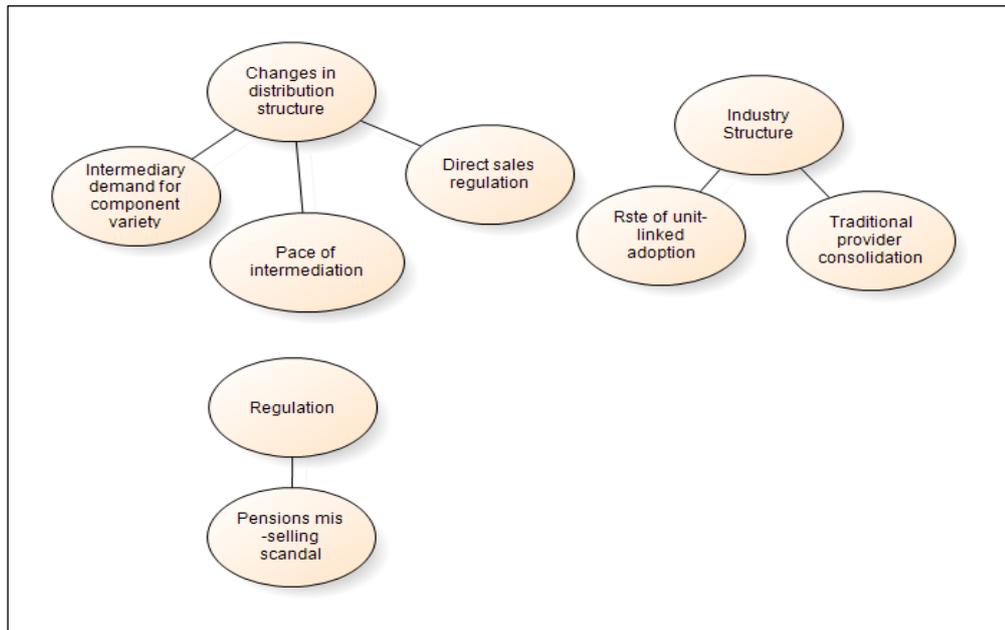


Figure 27: Industry factors: early to mid-1990s⁵⁷

Following deregulation in 1988, it was only a few years before a wave of pensions mis-selling scandals hit the industry and continued for a decade or more, “*you had the advent of personal pensions in the late 80s, have your own personal pension, which then led to the mis-selling scandals about five to seven years later*” (9). In the personal pensions product market, there were “*a lot of high commissions, a lot of scandals – people going to jail, it was a very cut-throat business*” (12) and it was a scandal that “*ultimately cost the industry, 50 billion pounds of compensation. Companies completely disappeared. The compensation was so great that they just couldn't support it. It was a terrible mess*” (24) and “*a lot of the sales people were villains basically*” (4).

⁵⁷ As also shown in tabular form in Table 21 (section 4.3.2.2)

Pensions mis-selling was, therefore, one of the factors that led incumbent insurance companies to review the governance mode of advice and distribution and began at pace to outsource the activity. Pensions mis-selling was effectively “*underpinning the growth of the IFA market. IFA's were directly accountable. So they were accountable for their own advice*” (26). The emergence of IFAs was also facilitated by selection forces, “*the market was evolving fast, I suppose more of the quality, top-end, direct-sales people, who historically could only go and recommend their own company's product could see unit-linking evolving with products that were more attractive, but they couldn't sell it while they were employed by a with-profits company. So if they couldn't compete on options, they were being outsold by the increasing IFA market*” (26).

The pensions mis-selling scandal and changes to the governance mode of advice and distribution are ultimately associated with the fast rate of adoption of unit-linked product designs in the early to mid-90s such that traditional with-profit insurance companies “*suddenly realised they needed to embrace it otherwise they could see market share disappearing*” (26) because “*unit-linking became a very significant part of the market relatively quickly*” (25) and “*the pendulum was swinging towards providers within the unit-linked space, Growing very, very rapidly*” (13). As the product market evolved towards unit-linking at pace, the industry begins to horizontally consolidate, “*so you certainly have the mergers going on in that place so there's a fight for scale going on*” (4) as “*the smaller fish are losing out and the bigger fish are merging*” (4). Some of the traditional insurance companies, however, “*became casualties because of the trend in demand. So there was quite a lot of acquisition and takeover through that period. So, there was an awful lot of consolidation of the provider market place going on as some of the smaller players realised that they were never going to be able to compete by themselves*” (26).

4.1.3.3. Mid-1990s to mid-noughties

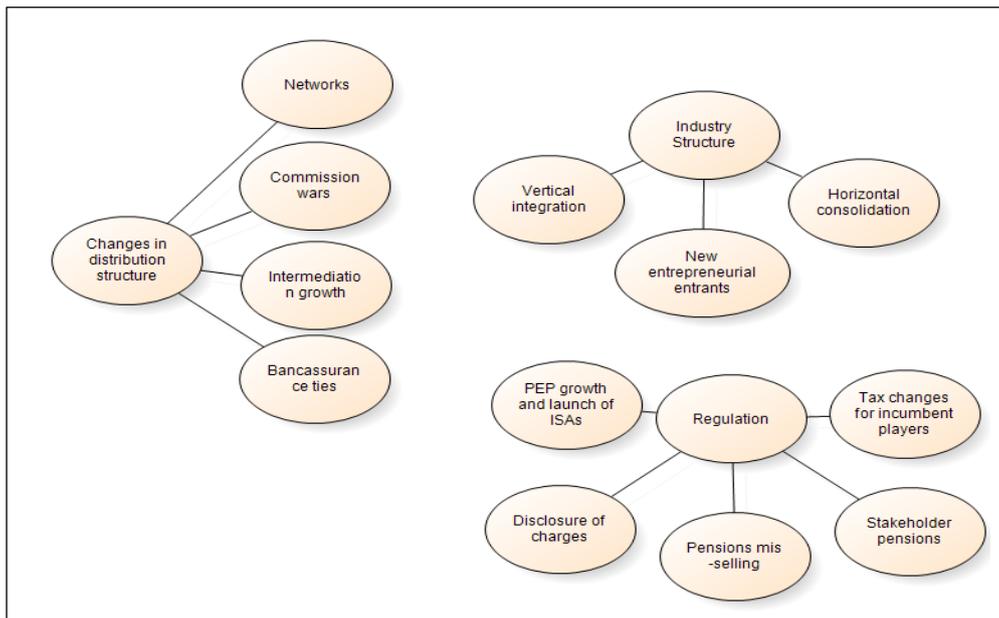


Figure 28: Industry factors: mid-1990s to mid-noughties⁵⁸

By the mid to late-90s, there is continued growth in outsourcing advice and distribution - *“the world had become more intermediated, people going to IFAs”* (22). Part of the attractiveness of running an IFA business was the amount of commission being paid by insurance companies to influence sales in a growing market, *“if you wanted to win business, you paid big sums of commission”* (23) and *“from a financial point of view, it probably started to become non-viable”* (8). However, a brake was applied in 1997 when the regulator introduced ‘standards’ that all personal pension product charges and intermediary commissions had to be disclosed to consumers in a prescribed format known as a ‘quote’, ‘projection’ or an ‘illustration’. The illustration had to use ‘standard’ growth rates to show what consumers might get back in the future, less all product charges and commissions. At an instant, the illustration revealed the extent to which charges and commissions reduced pension returns, forcing downward pressure on both. The emergence of league tables also reinforced downward momentum, *“it wasn't until 97, when your projections had to reflect your product charges, that our quotes didn't come out competitive at all. There was complete panic”* (13).

⁵⁸ As also shown in tabular form in Table 23 (section 4.3.2.2)

With the pensions mis-selling scandal continuing and the new disclosure ‘standards’ biting, “commentators were saying: hold on a minute, the price is higher than it should be. That led on to stakeholder pensions. Give people choice, but protect the customer by limiting what the provider can actually charge” (10). Because the price cap on stakeholder pensions was deemed likely to reduce variety (as it did), the Government permitted insurance companies to continue to offer other types of personal pensions at a price higher than the stakeholder cap providing ‘RU64’⁵⁹ rules were met. As a consequence, insurance companies and downstream IFA firms were attracted to SIPP as the product was deemed a ‘legitimate’ route to comply with RU64 requirements, permitting an almost unlimited range of investment options not available within stakeholder pensions. However, with continued downward pressure on margins, horizontal consolidation among incumbent insurance companies continued, as “there was increasing competition and therefore scale was becoming more and more important. You saw some big mergers” (25) and “there was quite a lot of consolidation and some of them were exiting. It's all to do with margin” (8).

By 2000, the prevailing head winds for insurance companies came together to initiate the next evolution towards a ‘fund supermarket’ architecture. Upstream fund management groups began to forward integrate into product provision as new product opportunities emerged, but also as a defensive measure. One of the key reasons for forward integration by fund management groups into product provision was diminishing gains from specialisation; for example, “if you have a fund that's performing well and has a reputation and track record – happy days. If you don't, unhappy days!” (17). The fund management groups also saw possible disintermediation risks, which were further compounded as soon as other fund management groups forward integrated into product provision. For instance, the fund management groups panicked and said: “oh my God, we're all going to be subject to [first mover organisations] whims and they're going to Hoover up; we've got to defend ourselves”. So a group of them got together, there was six of them and they all put in a sixth and they said we'll build a fund supermarket too. And they had no idea how much it was going to cost them to do what they were doing but you know we've got to protect ourselves. It was initially purely a defensive step” (24).

⁵⁹ RU64 meant that any IFA recommending a non-stakeholder personal pension had to have legitimate reasons, but “with one stroke of the pen, the cost of pensions plummeted” (24). Reasons had to be based upon consumer need, ie, the need for investment options not available within a stakeholder personal pension

Forward integration also had a commercial logic, however, as one respondent suggests “so, originally, where you earned the money was each fund manager's happy selling their funds and just being best at it. So long as you had the best performing funds, but we all know that you might have the best sometimes, but not other times. So distributors were picking and choosing between you all. So if you could get into the product provider food-chain and create a platform for everybody to go through, then you could earn some money off that as well” (15). The fund management groups initially began offering ISAs in 1999. There was a race to acquire market share, “it was all a race, it was a case of we will have this out in weeks” (17). Some of the fund management groups capitalised on existing productive capabilities to enter the market quickly, “they had street-corner outlets in the US, where people would come in and would buy and sell funds. So, they had a system and they just brought the whole system over to the UK” (24). The new entrants “started from the basis of the ISA market and the investment market, attracting assets that might otherwise be going into with-profits. I think also those providers started to nibble at the life office in terms of offering other wrappers like the SIPP pension too” (6).

Despite these shifts, many incumbent with-profits and unit-linked insurance companies were still slow to react to the fund supermarket architectural innovation, but slowly “there was a dawning realisation from the providers that in this possible future world maybe we're not really needed” (4). The reason for slow adaptation may be associated with rent appropriation and the burden of legacy, “it was mainly down to cost and the implication. So, if you're to launch a brand new shiny thing, but you've got say one million people in legacy products, what are you going to do with those one million people? Those one million people are in high margin products. If you then shift those people, you run the risk of losing that profit. Basically, a lot of providers looked at the sheer cost and thought that you could be looking at a hundred million pounds spend and the payback period on that would be a very, very long time for a life company” (7).

4.1.3.4. Mid-noughties to 2012

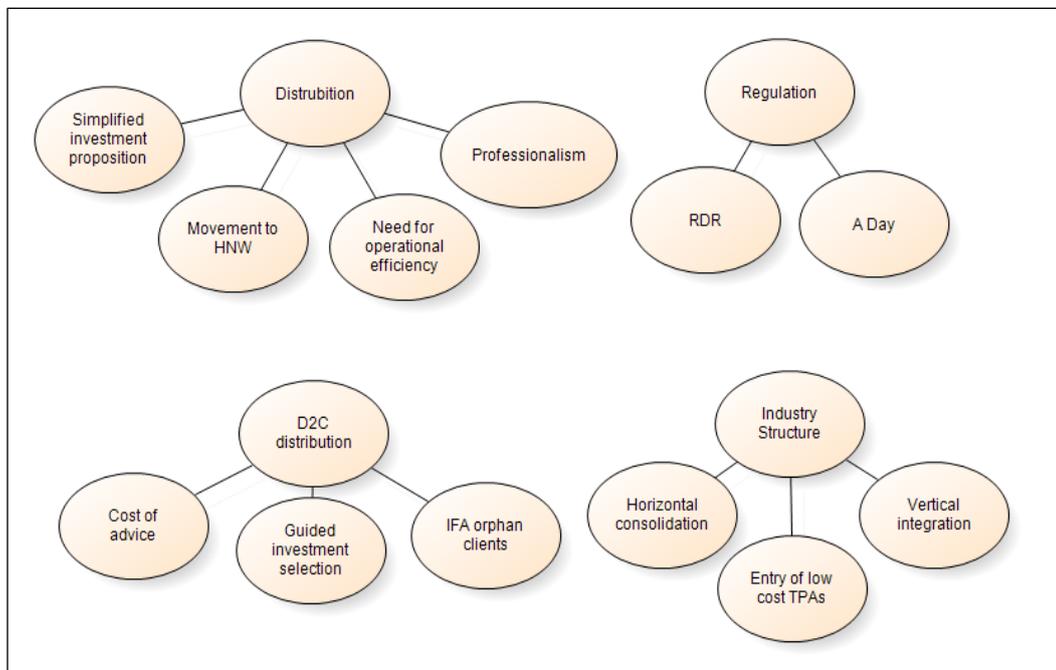


Figure 29: Industry factors: mid-noughties to 2012⁶⁰

The bringing together of ISAs and SIPPs into a multi-product platform architecture is primarily associated with pensions simplification regime (‘A-day’) in 2006. The intention of A-day was to unify the tangled range of pension rules since the 1950s into a single set of ‘standards’. The unification of rules was seized upon by the industry as an opportunity to promote the consolidation of a vast range of existing legacy pension products into a single ‘industry standard’ product type under a single set of rules, and SIPP was positioned as that product, owing to its modular properties and wider permitted investment rules.

⁶⁰ As also shown in tabular form in Table 25 (section 4.3.2.2)

A-day, therefore, was cast as an opportunity for pension consolidation, and this initiated a vast movement of legacy pensions away from incumbent insurance companies and towards new entrants with multi-product platform technologies. For instance, *“A-day in 2006 was another big catalyst. Suddenly the barriers to entry are removed, as getting regulated as a SIPP provider has a lot lower capital adequacy requirements than insurance companies, and that blew open the market at that point, and you started to see platforms gain traction from a standing start in c2003/4 to where by the end of decade pretty much all pensions business was being written on an open architecture platform product”* (21). A-day ultimately forced incumbent insurance companies to react quickly and adopt the new platform systems architecture, as one respondent points out *“from A-Day, it was such a massive change to the market. They had to do something or they would go out of business. You had this massive bubble of consolidation business at A-day that forced people to do something or you died”* (9).

By unlocking access to hundreds of billions of legacy pensions, the volumes available to new manufacturer/assembler firms meant that downward pressure on margin continued and volume and market share were king, *“when stakeholder came along it was a 1% world and it’s a 0.5% world now. When you did all the maths in current new business world, the only way it’s really going to work is six scale players, it can’t be more than six because they won’t get enough volume to actually get on the scale* (4) and this *“shifts the mind set towards scale so when you’re looking at scale and you’re looking at technology that will charge you on the basis of benefitting from that scale”* (20).

A-day signalled the entry of a number of new manufacturer/assembler firms, as well as systems architects with existing platform technologies bought into the UK primarily from Australia and New Zealand. This fast-changing environment increased the level of competition in all layers of the value chain. For instance, *“competition is getting stronger even though there are probably 30 platforms now, because the market was growing, there is enough capacity at the moment for them”* (3) and *“what happened was we saw a proliferation of new entrants who are very small which have dragged the existing incumbents to change their businesses and to adopt those models in order to survive”* (25).

The effect on rents is well summed up by one respondent *“the platform space was very profitable thank you very much indeed. In came the fund supermarkets in 2000 – not profitable anymore. So, you basically tried to hang on. And we had more volume than other people so, we could squeeze suppliers harder, so we could still make enough money provided we could keep the overall volumes up but certainly, it wasn't like the game before. But then people kept on coming. Product platforms then came in. All the time chipping away at your underlying margins until now you know, you need £50 billion under management before you can make any money. So they were looking at other ways – how would you make money out of this?”* (24).

The bases of competition began to shift also. Price was important to secure volume and market share, however *“technology reduced the barriers to entry significantly, and it also reduced the basis of competition, or changed it, where it was only once the big life companies and fund managers who could afford a platform to one where you can effectively get a bit of technology that you license, and then you can take it where you like. And, so I think that in this period it has greatly increased the number of competitors and has effectively created a market that didn't previously exist”* (20).

Platform systems architects entered the market around 2005/6 with platform architectures that had *“common components and platforms from Australia to the UK market but day one it was running in ready-to-go. A very obvious example is of the open architecture funds. So that was the expectation in the Australian market”* (3) and *“some influential people and influential firms that came into the market from Australia and New Zealand, who were champions and who led the change as well. So, trailblazers, early adopters and new entrants into the market who wanted to break down this really profitable market; let's knock those barriers down and come in with something world beating to get some of the action”* (22). As one respondent observed, *“it's a natural extension once they've kind of done it in their home market they will look for industries with similar rules, similar types, similar markets to try and play there”* (23).

4.1.3.5. 2012-2014

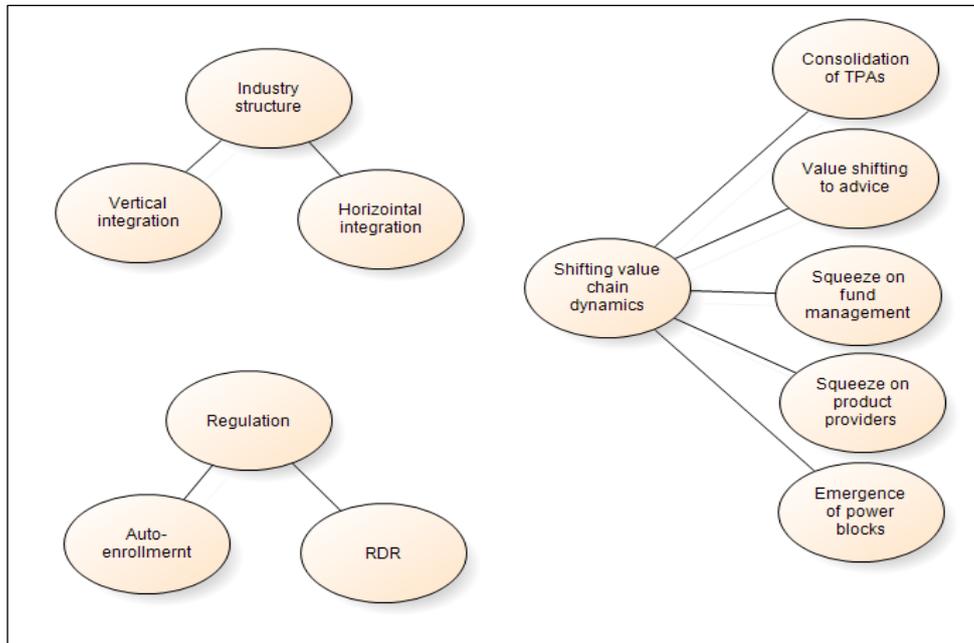


Figure 30: Industry factors: 2012-2014⁶¹

The implementation of RDR in 2012 continued to influence the industry and value chain dynamics such as division of surplus. With a ban now in place on IFA commission, the direction of travel for IFAs was to shift to serving mass affluent and higher net worth consumers who were more able to pay fees, as one respondent points out *“with the end of commission, I think the IFA market will continue its journey to higher wealth”* (3) and *“for IFAs, their margins are being squeezed as well, they can’t take commission, only a fee. And, they’ve got to be very transparent, bring their fees down, and people don’t want to pay as much as they were paying through commission – but didn’t know. So, the amount of time they can spend advising for a small fee is now lower, so it’s a time issue”* (1).

⁶¹ As also shown in tabular form in Table 27 (section 4.3.2.2)

Another factor that is associated with IFAs moving ‘up-market’ comes from the Group Pension product market. Auto-enrolment⁶² came in 2012 and compels an employer to enrol employees in a group pension scheme, subject to certain rules. This significantly limited the size of the personal pensions product market, for example *“we see the demise of the personal pension with auto-enrolment. But you definitely see a blur of the individual market with the corporate market, because if anybody's an employee, you're going to get auto-enrolled into a pension scheme. So that almost makes the individual personal pension redundant now. I mean the volumes, they were starting to fall off a cliff”* (8). As one respondent suggested *“is that not equal to stakeholder with compulsion?”* (13). For manufacturer/assembler firms, auto-enrolment has commercial consequences because *“we're now looking at a much more commoditised market again. A lot more people are going to have pensions because of auto-enrolment. But a lot of them are going to be putting very little into them. So again, it's like stakeholder and the problems of stakeholder because you've got even smaller fund amounts and you've got even more people. So it just drives the financials even further in the direction of get as big a slice of this market as we possibly can, and do it with as small a cost base as we can”* (8) and *“I mean we make a small margin on massive amounts of money under management”* (24).

As a result, *“platform providers are trying to dig out a niche for themselves to make some money because they're being squeezed. I mean everybody is being squeezed. We're going through the pain now. The gain will come to an end when the number of providers has shrunk down”* (24). The open and modular product design is associated with the squeeze on margin *“where you can pick any component you want, from a customer perspective that makes sense, but for the players in the industry it is hard to work out that if we give the consumer this much control, where do we make our money?”* (20) and *“I suppose added value is what it comes down to. You can't add value if all you do is allow access to everything under the sun. You can add value if you can give them something that you've created, which is value to you and the customer”* (15).

⁶² Auto-enrolment came in 2012, and enforces employers to fund a pension for all of their employees which is likely to reduce the size of the UK personal pensions product market

The issue is summed up by one respondent *“there doesn't need to be lots and lots of platforms if all that they're doing is aggregating external funds and components. Because one pension now looks just like another and they're all priced maybe at nil or very low cost. It's a product of convenience. And product provision is no longer a value-add to the customer. Investment advice and the outcome is a value-add for the customer”* (10).

Upstream fund management was also getting squeezed. As one respondent highlights, *“I think the fund management segment is getting massively squeezed and there's a price war going on. They just keep chopping down, step by step by step”* (21). The squeeze is also compounded by the manufacturer/assembler firms squeezing fund management groups exposed for a position within 'selected fund lists', *“we're certainly seeing the fund managers being squeezed very hard by the platform providers. So they're going to the fund managers and saying: “I know we allow you to keep 75 basis points on that lot, if you want to come in this one, you can only have 40”. And the fund managers are saying: “well, it's better to have 40 on this big pot rather than 75 basis points on bugger all”. So they're complying with that but they are feeling the pain”* (24).

The IT platform systems architects have also seen margins under pressure. As one respondent suggests *“the number of IT and admin providers is reducing. One or two of them have had to pull out. [Organisation] now runs the systems for four or five firms. [Organisation] are also running three or four. Still margins are very small and you've got to have enough scale to make it work. So it's a bit like the Hewlett Packard model, that is you sell a printer below cost and then you rip people off on the cartridges. So that's what they're doing, they're selling the system at the price that's dictated by the market because they have no control over the price at all. They sell at the market price and then they try and entice you with proprietary add-ons”* (24).

The compelling need for differentiation and competitive advantage has led to *“firms pulling back particularly as there's been pressure on price, I think some of those bells and whistles have been dispensed with. The drivers, there is one around the commercials of it, particularly as you go to open architecture, the actual operation of the platform gets commoditised and it's where can we add value? And one of those is to actually manage investments on behalf of the customer and charge them a fee for doing that. Even if what you're doing is managing other people's funds. So it is a reversion back, but it's back to a different version of what was there before”* (19).

This has led to manufacturer/assembler firms offering simplified bundles of product components, but within an open and modular systems framework, as one respondent pointed out *“I suppose look at auto-enrolment, you're back to a simple personal pension where people who are looking at what's put in front of them and I suppose relying on the people who either advising the employer or their employer to have made the right decisions for them”* (28). Therefore, following RDR there has been a convergence towards simpler and integrated bundles of product components, but within the context of an open architecture framework.

The value chain also began to reconfigure into ‘power blocks’ to try to mitigate against diminishing rents and gains from trade, as *“firms are wanting to self-restrict themselves, and suppliers are keen to support that happening too”* (6). The respondent continued *“there are power blocks in the market, quasi-vertically integrated through technology and other relationships. So what you end up with is large power blocks. They're not as integrated as the old provider ones, but they're still big power blocks going against each other with the platform at the core of it. The heritage of those power blocks will vary depending on the platform they stand on. So what you are seeing in the market is the platform providers increasingly in concert with the fund groups to build out those power blocks”* (6).

The product market has also started to integrate both horizontally and vertically. For example, *“there were new firms that came up through the late noughties that were, by 2012, consolidating with more mature players”* (11). Horizontal consolidation was occurring at all levels of the value chain. For example, in fund management *“fund-management groups with history were gobbling up newer fund groups to bring their expertise in to widen their own product set”* (11) and *“we see more of a shake out as RDR pushes down the prices and we will see a move towards more super-fund groups. It's certainly created a shift in terms of where the value is and the fact that the fund management is commoditised in some ways”* (19).

In the downstream distribution and advice layer of the value chain (ie, IFAs), consolidation was also evident because *“at the advisor level, there was chaos in the market. Post-2008, you had the big IFA networks going broke. You had the emergence of new networks driven by the same people that sold the last ones just before they went broke. You look back where we are now and look back five years, the market looked very different. So you absolutely had consolidation across all different elements in the market and that all meant huge tightening of margins. So nobody was winning at that point”* (11) and *“you’re seeing IFA consolidators coming in. And, the advice-side becoming less fragmented”* (21).

In complementary product components, horizontal consolidation also took root, *“a classic example would be [organisations] purchased by an Australian company called [organisation], so they were a big software management organisation in Australia, they bought them along with some other sort of peripheral bits of software that are all aimed at an intermediary perspective; but their strategy is very much ‘we want to get a market-leading component in all these different areas’.”*(20)

Interestingly, the manufacturer/assembler layer of the value chain seems currently immune from horizontal consolidation “*you're not seeing it at the moment in terms of existing platforms*” (17), perhaps because “*some of the platforms decided they know exactly what they are, and they're going to stick to being what they are, and some of them have been profitable as a result. I think it's because they stick to their knitting; that has enabled them to focus on key profit drivers and, therefore, have managed to drive and derive some profit. Some of the larger ones have tried to do too much and are still to get to the point of scale, where they're covering fixed costs*” (16). One respondent, however, looked into a crystal ball and suggested that “*but the statistics say that we're unlikely to have 35 or 37 platform providers in three years, because there isn't enough money to make it sustainable. I reckon there could be about ten*” (14)

Forward and backward vertical integration has also begun to be a characteristic of the product market since RDR in 2012. For example, manufacturer/assembler firms forward integrating into advice and distribution again, “*a quasi-arms-length solution where the providers were taking more and more significant financial stakes in the advisory businesses. And were using those stakes to provide additional distribution*” (11) and manufacturer/assembler firms are “*buying IFA's and wealth managers at the moment*” (17). One respondent sums up the logic “*for guidance and advice, we bought an advisor. So we now have our own advice capability in house, which we've now fully integrated, and rebranded. We wanted to do more in retirement for people who need help and advice as a lot of them won't be able to get it from a financial advisor*” (19).

Manufacturer/assembler firms were also backward integrating into complementary product component suppliers to buy-in specialist productive capabilities, such as “*we bought a stockbroking business. I think in terms of development of products now, it's all about actually being very close to how the things behave*” (17), “[*organisation*] *bought a front office portfolio management and reporting solution that sits in front of platforms; can take in standardised data feeds from other areas and give you aggregated reporting and portfolio construction, which is much richer than a platform can do*” (29) and “*one of the big platforms recently has purchased a commercial property supplier and platforms have purchased organisations that do portfolio management construction*” (2).

There is also evidence of downstream IFAs backward integrating into the product provision and fund management. For example, *“an intermediary business that brought a SIPP manufacturing business, and then also brought other intermediary businesses that had SIPP clients, so buy the skill or the capability that you don’t have, and then buy others that then give you the scale, to grow that capability”* (16) and *“[Organisation] has an investment management division that runs funds, they have a master trust, they have the product wrapper. And if you look at the value chain, they own the client relationship, they put them into a product where there’s product manufacturing margin. They invest them in funds where there’s investment manufacturing margin and, for certain types of individual members underneath that, individual wealth management services that they own. You know, they manufacturer a lot of their own product, and build their own investment capability. These firms, you could look at them and say, “Are they actually an intermediary? Or are they a provider? Or are they a bit of both?”* (16).

4.2 RQ2: To what extent are the design characteristics of product architecture associated with the design characteristics of firm and/or industry architecture?

4.2.1 Mid to late-1980s

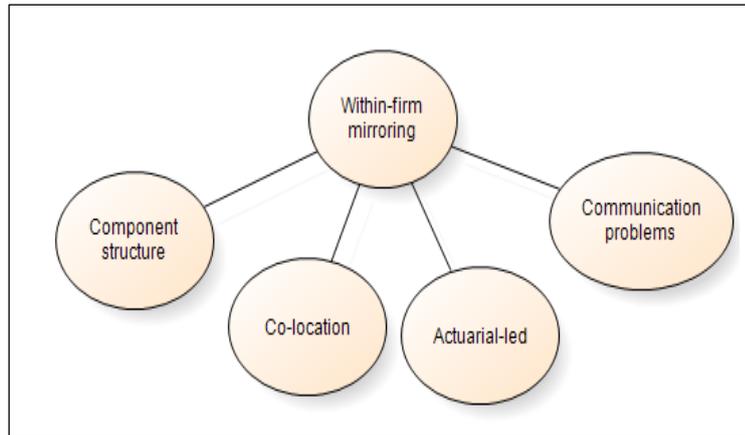


Figure 31: Correspondence: mid to late-1980s⁶³

In the mid-1980s, firm boundaries were characterised by vertical integration. Product development teams were co-located, for instance *“the way the trend was going was to base the product development people with the IT people, so when you did the project you went and worked in the IT area with the project to have everybody in the same room so that when the IT programmer got stuck, if the analyst couldn’t answer the question, someone was there to answer the question that day rather than, “Oh let’s see, I’m seeing him next Wednesday let’s find out the answer then” (4).* Product development was, therefore, a multi-discipline approach, *“development was done by multi-disciplinary teams within the firm so you’d have an actuary; you’d have a marketing person; you’d have an IT person; you’d have a finance person. And they were all sitting in the same building” (3).*

⁶³ As also shown in tabular form in Table 20 (section 4.3.2.2)

The co-location of product development teams may, in part, be associated with the smaller size of firms, *“companies were just smaller, so the idea that everybody knew everything is more likely to be true”* (3). One respondent recalled that in some cases functional specialisms were geographically-separated, but only ever by a short distance, *“mainly they were geographically very close to one another. I can remember the company I worked for had an office in Croydon and then they had their IT team in London. And they used to talk about it as if it was at the other end of the universe. And actually, it was about ten minutes down the road! And even then it was a sort of a “them and us” kind of culture was created”* (8).

Despite multi-disciplinary product development teams, there were very few, if any, project managers, *“I don't recall there being project managers”* (8), or product managers *“I wouldn't say there was anyone who sat across all of it and said “well, actually, I own the proposition, I want pricing to be this, I want investment content to be this”. So, it was quite difficult to work all that together because there was no-one person who bought it together”* (1), and later by the very early 90s *“because so many disasters occurred with that structure, a kind of middle group of product developers seemed to be created”* (8).

In the meantime, the product development process was led by the actuarial product development function, *“product development was very Actuarial led”* (3) and if *“you were looking around to see who was doing sort of job it would have been 99% actuarial, and a very small number of people”* (8). This led to a very linear product development process, for example *“so standard waterfall, do X, once you've done X you do Y, once you've done Y you do Z, if Z doesn't quite work you go back to X or you go back to the point, in the chain, where it broke and go back and fix it. So very functional and linear, in its approach”* (16).

High-levels of information-sharing were common within the actuarial product development team, but information-sharing between them and wider functional teams was characterised as low, “*actuarial would determine the rights and wrongs of product evolution and the features that were offered. Little involvement from marketing and sales. Little involvement from customer services – they were there to do their support function*” (27) and “*it's usually the actuaries that come up with an idea and they usually specify it. Then it goes out to a team with input from admin, marketing, compliance, legal, all commenting on whether that will work. So the admin people will say: “well that charging structure won't work because our systems can't cope with it”* (12). As one respondent recalls “*that's the odd mismatch. You've got a product where every piece affects every other piece. The actuarial team might not talk to the finance team or any team. And it will take them a very long time to suggest something and an even longer time for somebody to realise that it's happened*” (15). As a consequence, “*then what comes out the other end is completely different to what we started with. There was quite a lot of concern about IT developments taking a phenomenal amount of time and having lots of problems at the end anyway*” (4)

As fund components started to be modularised, there is an association with the loose-coupling of the fund development team, for instance “*there was the investment component that would have been in a different site within the same company*” (1). The late-80s also witnesses the early emergence of a ‘product owner’ role, to bring together and coordinate the various component parts of the product architecture, for instance, “*you had now product owners who understood the product, then there were teams of investment people who understood investment content, and finance people who understood the pricing, and they all had accountability for their own bits really*” (14).

4.2.2 Early to mid-1990s

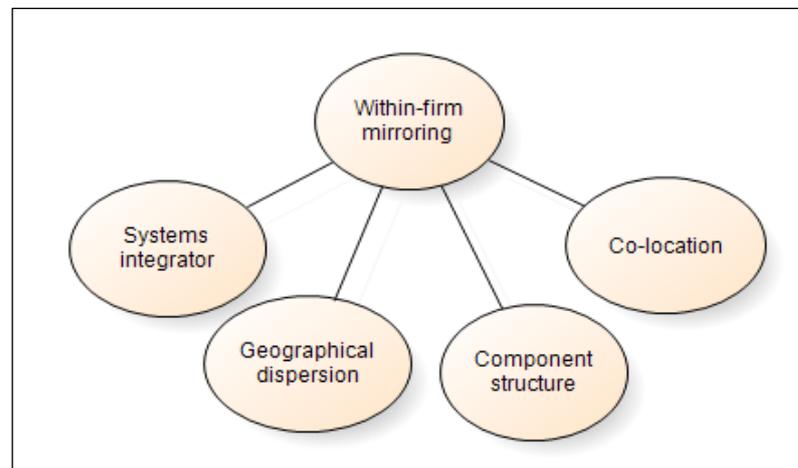


Figure 32: Correspondence: early to mid-1990s⁶⁴

As advice and distribution and fund components were modularised, the respective task and knowledge boundaries are also partitioned. In addition, functional specialisms such as customer services begin to be modularised *“all the customer services was done in particular teams, all the marketing in done in a particular area, actuaries, in one place, it had become quite componentised and not only for pensions, for all products”* (3) and *“it was functionalised and we spent a lot of time on what we call, meta processes, the money out process, a death claim process, and it really didn’t matter which product you had”* (11).

However, loose-coupling of product components, tasks and knowledge was not associated with geographical distance or outsourcing. For example, one respondent recalls that while the fund development team had a separate structure, *“your investments team would have been a lot closer, if not in the same building in some cases”* (23), and where geographical separation occurred it was usually along functional lines and within the same town or city, for instance *“many of the teams had been segmenting by business structure and they typically wouldn’t be co-located. They might be co-located within the same city, same town though”* (27).

⁶⁴ As also shown in tabular form in Table 22 (section 4.3.2.2)

The product owner role, however, began to be more dominant in many firms, and they oversaw *“the ownership of the product development area and the drive of making it happen”* (4). A product owner was needed to coordinate and orchestrate activities because *“it was Chinese whispers. We had a product team who put the product together, and it goes from component to component until it gets out in the market, and if you don’t understand the whole process, and it goes wrong, you need to know which component team dealt with it and who understands it”* (1).

4.2.3 Mid-1990s to mid-noughties

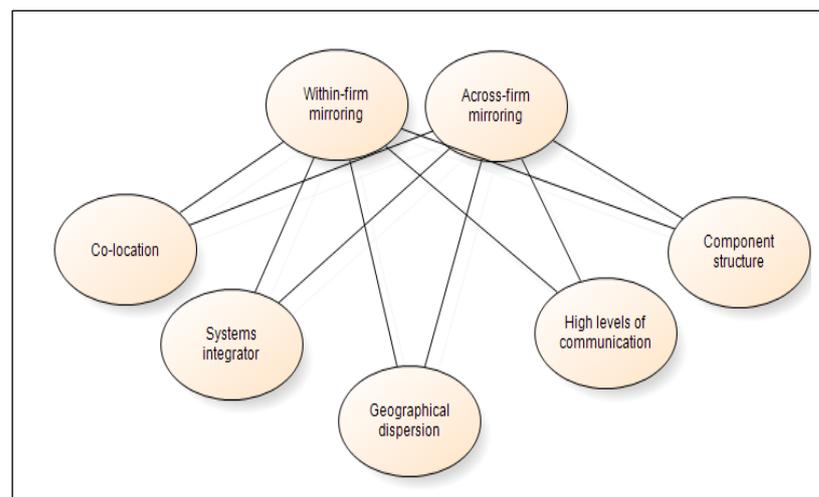


Figure 33: Correspondence: mid-1990s to mid-noughties⁶⁵

By the mid to late-90s as product design evolved towards hybridity, the partitioning of the product development structure – and hence task and knowledge structure - began to shift in a number of ways. For example, many firms began to develop more formalised proposition development teams with a ‘proposition owner’ for each product type, encompassing developers across different functional areas, *“so, you had a product development team responsible for the product. So they would do a lot of the architectural level”* (7).

⁶⁵ As also shown in tabular form in Table 24 (section 4.3.2.2)

Proposition owners (or systems integrators) controlled the architecture, and were “*very unified teams, often in one location where they wouldn’t have broken up*” (18) and “*there would have been probably one team designing the product architecture*” (23), because the nature of the product market was such that “*a lot of the competition was on the features of the product, almost the way the product worked*” (23). The proposition owner had “*big teams that did everything from design the product, associated services, and then you had the teccies, designed the sort of technical aspects of the product*” (21). Co-location remained a dominant feature of the proposition development teams, “*so people that knew each other well, we were all on the same floor. We all sat together, worked together*” (17), “*it was in the same building, same location, always geographically co-located*” (6), and “*if you were working on a new thing, everybody would go and sit in one location*” (28).

The emergence of a proposition owner that co-ordinates product design is associated with the emergence of a componentisation of the firm architecture that corresponds to product architecture, for instance “*if you wanted more funds added, you would outsource that to the existing funds team and they would see what could be done. And then the actual adding those funds would be outsourced internally to the architectural IT people*” (7). In addition, functional component teams were partitioned to offer specialised expertise, “*we pulled out the service component, and that was done as a cross-functional piece, so you could have common service teams across different propositions. They became very independent and very geographically dispersed*” (6).

Furthermore, other component development teams started to become geographically-distant, “*you see things start being broken out, so the investment people moving away from the other bits, centres of excellence around different parts and processes. So an administration centre being up and an actuarial centre being set up*” (18), or “*if you want good quality fund managers, you have to be in London, so you end up having a geographical separation*” (15). Where product components were outsourced, “*fund providers always been a separate team, and they have always worked very closely but are in a totally separate location*” (23) and “*once you were using suppliers, outsourcing components, so you’re bound to get a geographical change*” (1).

Despite the partitioning of the firm architecture into a proposition development team and loosely-coupled product and process component teams, information-sharing between component development teams and the proposition owner was high, for instance, *“a lot of close communication because the product teams and component teams to resolve issues (6).* Where the product component was outsourced, high-levels of information-sharing were also maintained ex-ante and ex-post, such that *“in the early days of using component suppliers it was one of the difficult things, setting them up, and then you also had to set up that communication process and sharing of information throughout” (1).*

In other words *“it moved to a point where we wanted to have a small number of people who had to have a good level of understanding of everything, and within there you would have component teams who specialised. So, more component teams, more wide spread because some are external, but internally you needed a better knowledge about how those components worked and fitted together” (1).*

By the early to mid-noughties, with the evolution to a fund supermarket product architecture, many manufacturer/assembler firms seized upon this as an opportunity to establish a separate business unit structure (SBUs), partitioned into a component structure, and coordinated by a systems integration function, *“so we weren’t building from within the existing infrastructure, if you looked at the proposition, it was divided into the components, so we had a product architecture team, an investment proposition team, and a software team, and I had dotted line report into me from the admin guys” (21).* Moreover, *“the product had a proposition team within the firm who grew to understand all the different components on offer, but then of course, you were also dealing with not just the internal component people, you’re dealing with component people outside, we were dealing with fund management people all over London, we were dealing with IT platform suppliers all over the country. It was like a jigsaw puzzle. You had all these pieces that were all over the place, and all had a part to play, but someone needed how to fit them all together to make a coherent proposition to the market (1).*

Systems integration, however, created challenges around information-sharing and there seem to be a number of different ways this was overcome. In some cases there were “lots of face-to-face meetings, telephone conferences” (7). However, the emergence of internet-enabled technology in the late-90s made information-sharing among geographically-dispersed actors more possible, for instance, “you can share information, you can communicate a lot better than you could in the 80s, for example by using collaboration software. You don’t have to meet somebody” (1). In some cases, external firms were bought together to work on co-development projects, “from time to time we might do a project, where designers and representatives would come in and work with us” (28). However, another model was a ‘relationship interface’, where “there was typically a single point of contact from each firm that communicated together as a team” (25). However, as some product components became more modularised, ex-ante information-sharing replaced the need for high amounts of ex-post information-sharing, “the communication became more co-defined up-front, so it wasn’t required as much thereafter” (6).

4.2.4. Mid-noughties to 2012

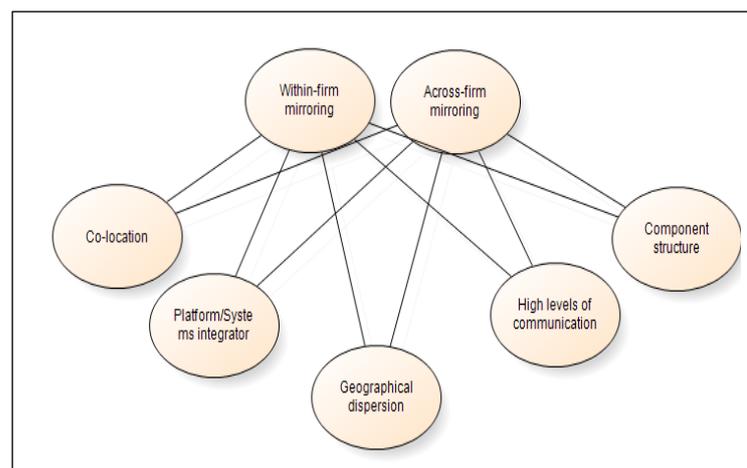


Figure 34: Correspondence: mid-noughties to 2012⁶⁶

⁶⁶ As also shown in tabular form in Table 26 (section 4.3.2.2)

A firm architecture characterised by a systems integration team and loosely-coupled component development was not without issues and problems. As one respondent recalls *“Projects created all sorts of problems. They had a pension team, they had a team that were looking at customer journey, web look and feel and all that, they had a team for pricing and so on, and the trouble was that these teams worked in verticals having agreed a certain standard, the trouble was when you came to end to join back up again, it didn’t quite join back up again very well. So, you found actually, we didn’t think of that, and that doesn’t now work on the customer journey or whatever, and we effectively had to re-work and undo, backwards in those verticals to join back up again”* (21). In other words, ex-ante information-sharing to agree standards failed to eliminate all ex-post interdependencies.

As the product was now effectively a ‘component’ of a platform architecture, a platform systems integration function emerged. Product-level systems integration teams continued to exist, but reduced substantially in size. What emerged was a platform systems integration team that focused on coordinating and integrating the various platform components into a coherent whole, as one respondent explains *“what you saw is the advent of the idea of the platform during this period. So what the platform team was all about was integrating products together, so you retained the product team, typically they got smaller and more specialist in their focus”* (6) and *“so the platform team then effectively brought together the components. So the platform team would have a very close relationship with product teams, but the platform team would agree a common user interface so the platform teams then start to work with other product teams to create a common set of tools. I suppose what’s happened is this platform team structure emerged and that became the actor in the process that brought it all together. So what happens is the platform team built up its own capability, and actually the products were almost quite transactionally handed over to the platform team and they just put them together. And then again, the investment teams handed over their components so they were all kind linked together”* (6). In other words, the systems integration role evolved to the higher hierarchal level of the platform, and the propositions team became a ‘component’ of the platform team.

As the platform architecture was multi-product, few developers had knowledge that ranged across the entire platform architecture and so *“they become 'Jack of all trades'. They know quite a bit about technology. They're prepared to accept that they usually don't know what they need to know about the new thing they're working on because it will always be in a slightly different area to the one that they're working on. And they're very good at building solution designs and sometimes almost architectural solution designs, so that they can work out, usually with help from IT architects, but usually having to drive things, so we've got that component and that component and that component; we're missing something here to do this new job and we need to find something to do that, so their job is to then use expertise that's in the business to try and get to the right third-party supplier. And not always to accept that the one that is in-house is the right one”* (8).

Task partitioning seems to depend upon the coupling characteristics of the product or service component. For example, product component teams had *“ownership and accountability for delivering different strands of that proposition, you know whether it be commercial property team or the drawdown team”* (29) and *“we also used people in the funds team who thought about what new funds we wanted, pricing team who were doing pricing”*(22). Moreover, where a product component could be characterised as loosely-coupled, many firms shared information ex-ante in order to agree standards, and then reverted to a low-level of ex-post information-sharing, for example, *“if it is complex, it is very difficult to specify. You don't know where you're heading until you're on the journey and an agile form of development lends itself well to platforms – it requires you to co-locate or at least collaborate, however if it's a pure waterfall type of development approach, and it's absolutely specified, you've agreed standards, and at a massive level of detail, you can avoid it”* (21). However, as the platform architecture was complex, often high-levels of information-sharing and co-location were maintained throughout an agile product development methodology, *“it's more co-creation and working with colleagues in the business and where you have third parties working in a more collaborative way”* (28).

There also seems to be an association between co-location or geographical distance and the complexity and ambiguity of the development, for instance, *“I think it really depends on the nature and the scale of the project”* (28) and the *“colocation partnering model, where quality and speed are important works really well, particularly in an ambiguous environment, in a market that we think something needs to look like this. You think you know what you want to achieve, and somebody might have come to you with a kernel of an idea, but you actually don’t really know, when you take it to market, exactly what it’s going to look like and that’s, for me, been a critical driver, in what works well. If you try and run that type of build, without doing an element of skill-sharing and knowledge sharing, it won’t work. If you’re building new capability that might break new ground in a market, you would look to collocate”* (16).

Temporary co-development and co-location is a dominant innovation mode, as one respondent recalls *“the team is constructed so it’s more relationship management with third-party providers including fund providers and technology providers. And they are almost like the core team”* (7) and *“it’s co-location and it’s lots of communication, we will have people in our team who have very strong relationships with you know some people on the component side and they’ll have specialist relationship managers for us. But it is co-located. We have people who go and sit in with them and we have people from there come and sit with us. We have people who have very strong relationships between the two”* (23). Advantages of co-location and high-levels of communication were *“good ways to communicate and be co-located for developments so you can knock down issues”* (22), as well as stimulating new innovation *“because what you find is that when you get clever people together just sitting over a coffee having a chat, they fire off new ideas and create brand new innovation out of nothing. Things that none of them as individuals would ever have come up with. I think that there is definitely a $2+2 = 5$ effectively to get people in the same place”* (21).

The decision as to whether to co-locate with external firms or remain geographically-distance was often based on the differentiation potential of the component innovation, such as “*you can collocate or not co-locate for the simple stuff. If it’s a very clearly defined task that you know, for example, the messaging standard for contract enquiry, go and do it, they can be sitting in Xatmandu as far as I’m concerned, it doesn’t matter. But, when you’re not quite sure, and part of the process is creating a differentiated solution, the more you develop things jointly, the more you need to be in the same place*” (21) and “*when standardised, we cut them loose, they go off and do what you’ve asked them to do. We give you a mandate to do this, in this way. This is what we want the outcome to be. There would be much more oversight, ownership and co-development if the component offered differentiation*”(9). As one respondent summed up, “*I think where we want there to be a differentiating factor, we will work very closely with suppliers, the component that we’ve outsourced, given that it’s probably a) important to us and b) the component knowledge is a separate skill to the core product knowledge. We have then build up a team to specifically look after that component, from our side, in terms of thinking about the innovation of that component, even if it’s someone else who actually goes and builds it*” (11).

Conversely, where the component was low-value and standard, “*I think it’s potentially the sort of simplicity of the components and maybe just how much of it was outsourced. So, the illustrations, was just outsourced everything. And we probably don’t particularly want it to be a differentiator. And so it doesn’t get a whole lot of focus from the product side of things. It’s almost something we have to have whereas technology, we do want that to differentiate us, because it’s not one of the big differentiating factors in the market*” (23) and, “*if you can clearly define something, then who you select to build with is quite straightforward, it’s about capability, at the right price, on time*” (16).

4.2.5. 2012 to 2014

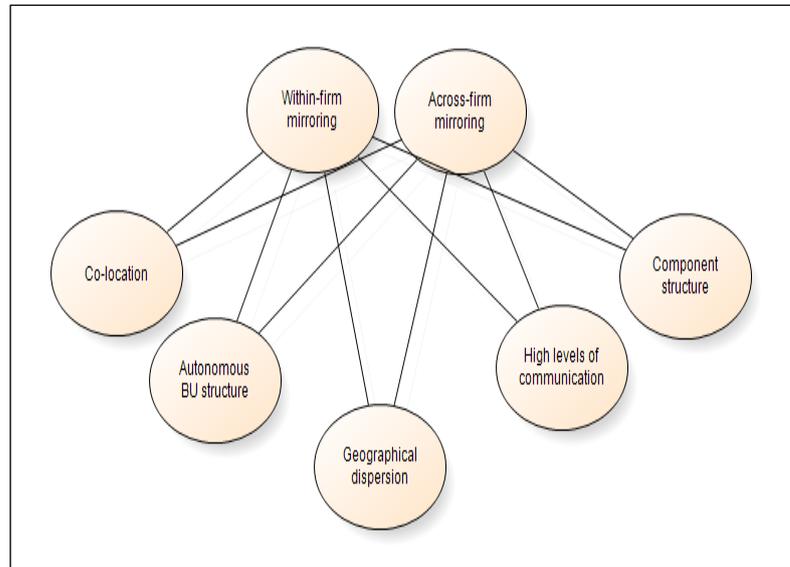


Figure 35: Correspondence: 2012 to 2014⁶⁷

By 2012, many manufacturer/assembler firms have adopted a quasi-autonomous SBU structure that is *“radically smaller than we’ve ever been as we’d lost sight of the customer and the value chain. So when we actually reorganise our business around our value chains, actually we dismantle all the big combined teams that did a little bit of everything so we found that efficiencies by separating out”* (3), and another respondent points out that *“I think now you bring in an SME model”* (14) so that *“we are at a point where you have a team who owned all the pieces of the jigsaw. They were then deciding which of these components do we go externally for, and which internally for. They are probably bringing back in house some of the components so geographically it brings people closer together again”* (1), although there remains a high degree of task and knowledge partitioning *“you’d have somebody that would, let’s say, specialise in the financial planning arm, somebody specialising in pension legislation; somebody that might specialise in investments and build up from there* (14).

⁶⁷ As also shown in tabular form in Table 28 (section 4.3.2.2)

Levels of information-sharing, both within- and across firm boundaries, appears to be associated with setting standards or ex-ante design requirements, such as *“having defined those basics, you can then disperse the team but you still need to regroup regularly to make sure that you continue to understand each other with an ongoing dialogue”* (27), *“so even though there are industry standards, it’s easier to communicate externally anyway and its even easier if you can bring it in house, but in terms of the flow of information, it makes it all easier as its back in house”* (1) and *“fund groups obviously work at a distance but I think the need to manage relationships with them has grown over time. The actual fund managers are locationally quite distant but there is an on-going dialogue between them and across them as well, because you want lots of commonality, you’re dealing with hundreds of them and want to deal with hundreds of them in the same way”* (6)

Geographical distance is a common mode for loosely-coupled and perceived low-value add product components, *“where it just works, say, a new third-party software provider and they want to take valuation data out of our pensions business, we gave them a manual and said ‘just do it, kind of thing’* (20). However, there does, however, appear to be an increase in information-sharing ex-post where the product component offers the potential for differentiation and competitive advantage. For example, internally-designed fund components, *“the internal investment capability offers differentiation, and the information flows have got much stronger* (6) and *“the software suppliers, those relationships began in a transactional way, but we realised there’s actually a lot of intellectual property with those suppliers that we need to tap into. So the relationship has become a lot closer, to the extent that in many cases, it’s not co-location but it’s near as damn it, so they are operating in a building very close to ours and regular meetings are taking place so that relationship has got closer”* (6).

The respondent continued, *“advisor relationships are important because they’re in a really good position to actually define requirements. What the platform provider can bring to that is an understanding what the technology is capable of, and actually that’s where most competitive advantage is, where you marry the two and say, here are the set of requirements, here are the technical capabilities, to what extent can we bring those together to differentiate. So actually it’s a shared intellectual property across the two. And if you separated them, actually neither of them would be that successful without the other. So there’s a lot of time spent back and forth, defining the sort of requirements, getting a proposal by the platform to define that, and going through the project management cycle. If you don’t do that, you have generic components and you probably come to market with something vanilla”* (6).

Another respondent summed up, *“it’s a really involved process, although there’s a lot of standards, but there’s actually so many little nuances there and things, that anything that could go wrong....you need that sort of personal co-location. We’re seeing a lot more of that cross-company working as one, where you’re relationships are so integrated because they’re doing such a major core component of your business, the behaviour is to treat them as part of your business, because how much you are reliant upon them effectively”* (20).

5. Discussion

This chapter presents a discussion of the key findings from the primary research presented in Chapter 4 and synthesises them with the relevant extant literature.

The discussion presented is structured into three parts: section 5.1 confirms the key findings derived from the product architecture typology between 1984 and 2014 and positions the design evolutions within the innovation framework presented by Henderson & Clark (1990). Next, by then incorporating an efficiency and productive capabilities perspective⁶⁸, section 5.2 examines how architectures may co-evolve across time. Finally, section 5.3 highlights how and if architectures correspond as they co-evolve.

By way of a reminder, this chapter aims to illuminate:

- RQ1: What supply-side processes are associated with the co-evolution of architectures? (sections 5.1 and 5.2)
- RQ2: To what extent are the design characteristics of product architecture associated with the design characteristics of firm and/or industry architecture? (section 5.3)

5.1 Product architecture evolution and innovation modes

By way of a summary of the findings presented in chapter 4, I highlight in Figure 36 how the product architecture evolved from one design type to another in the UK personal pensions product market between 1984 and 2014. In further detail, appendices six to eight show the evolution of product architectures at the product component level and the associated governance mode.

⁶⁸ Section 5.2 develops its arguments from an efficiency and productive capabilities perspective as the templates illustrated in chapter 4 in response to RQ1 clearly identify them as ‘integrative themes’ (King, 1998, 2014) throughout the 5 time periods. In other words, while the findings presented in chapter 4 outline a number of different possible explanations, efficiency and productive capabilities considerations seem to represent the ‘glue’ across time that can help illuminate the processes underpinning architectural co-evolution in the UK personal pensions product market. The Northumbria University PhD regulations on word count also provide a secondary logic for narrowing the scope of this discussion to such integrative themes. It is acknowledged that further research opportunities exist to explore other theoretical perspectives and themes arising from the templates.

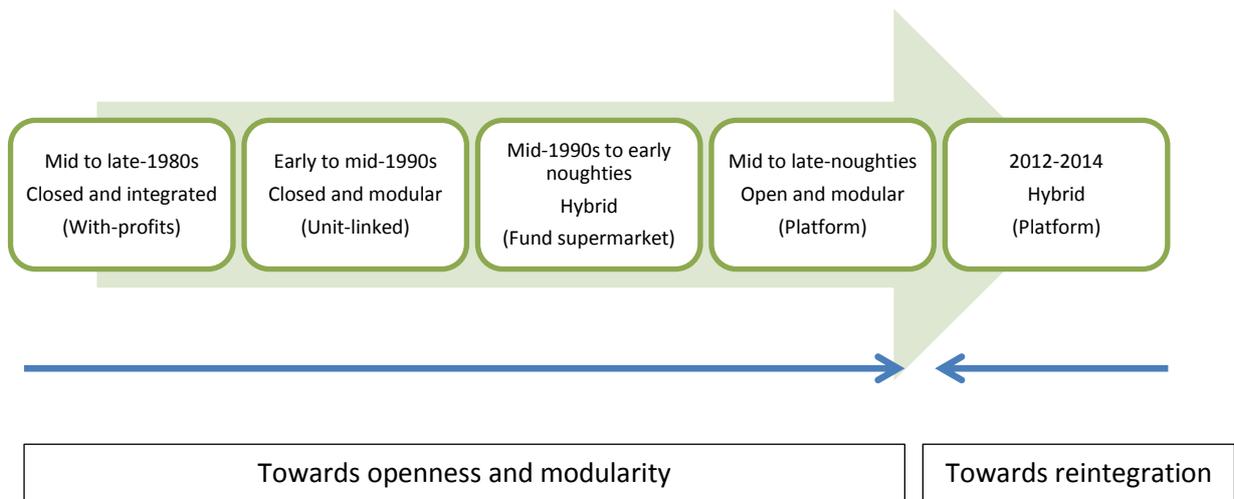


Figure 36: Product architecture evolution

Wolter and Veloso (2008) argue that one of the challenges to examining and understanding the combined effects of efficiency considerations and productive capabilities on innovation is that the mode of innovation being analysed often lacks precise definition. Hence, like Walter and Veloso (2008), it seems reasonable to follow the seminal work of Henderson and Clark (1990) and develop the discussion primarily from the perspective of the manufacturer/assembler layer of the value chain responding to incremental, modular, architectural or radical innovation modes. However, as Fixson and Park (2008) note, the typology proposed by Henderson and Clark (1990) does not envisage a type of innovation mode associated with reintegration, and hence I also follow Fixson and Park by proposing ‘integrative innovation’ as a further stylised innovation type.

As I shall go on to highlight, the findings in chapter 4 indicate that there are four distinct innovation phases as shown in Figure 37:

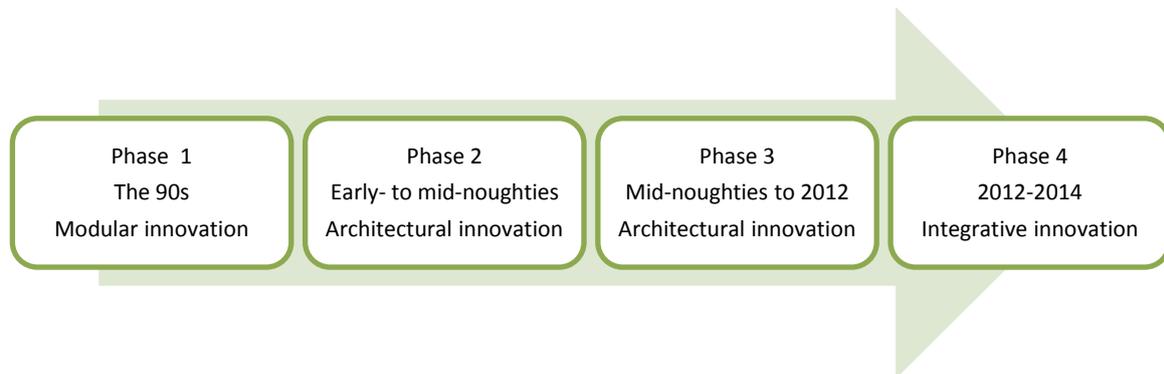


Figure 37: Innovation modes

Having proposed the stylised innovation modes that permeate the findings, I now turn to discuss how efficiency considerations and productive capabilities may combine in different ways across time to co-evolve architectures in the UK personal pensions sector.

5.2. Architectural co-evolution

Cacciatori & Jacobides (2005:1851) posit the question “how do value chains evolve?” and argue that in fact very little is known about how the institutional and vertical architectures of product markets change over time. To provide analytical focus, this section will relate the findings presented in chapter 4 to the extant efficiency and productive capabilities literatures. Although using interchangeable terminology, the gap in the extant literature is succinctly summarised by Walter & Veloso (2008:586) who argue that “*most studies of vertical structure incorporating TCE and firm competencies have neglected innovation*” and further posit that “*technological change reasonably complicates the task of identifying transaction costs, relevant competencies, and how the interplay between them shapes firm boundaries*”.

Efficiency considerations in determining the ‘vertical architecture’ (Cacciatori & Jacobides, 2005:1852) of a product market are often associated with production costs (Stigler, 1951), bureaucratic costs (ie, Jacobides, 2008; Jacobides & Hitt, 2005), and the presence or absence of transaction costs (Klein, *et al.*, 1978; Williamson, 1975; 1985). Therefore, in general terms, the balance of bureaucratic costs (BC) + internal production costs (IPC) versus external production costs (EPC) + transaction costs (TC) may determine firm boundary decisions (ie, Williamson, 1975). As a generic term for external governance costs, transaction costs economics (TCE) has typically examined firm boundaries at the level of the transaction (Jacobides, 2005), presenting a static picture of the menu of alternative governance modes. Many scholars, such Baldwin (2008), Jacobides & Winter (2005), Langlois (2002, 2003, 2006), and Sanchez, *et al.*, (2013), in contrast, argues that transaction costs are often transient and dynamic, hence they can decline or increase across time, especially under conditions of technological change (ie, Furlan, *et al.*, 2014; Wolter & Veloso, 2008). Indeed, Williamson (1985) acknowledged that TCE would be less persuasive in dynamic environments as compared to a static analysis.

Jacobides and Winter (2005) highlight that transaction costs are often mediated by the productive capability distribution in a given product market. Hence, the notion of ‘productive capabilities’ embraces an ‘efficiency’ perspective in so far as it relates to how efficiently or effectively a given firm carries out its productive activities. According to Jacobides and Winter (2005), productive capabilities along a value chain are often heterogeneous and path-dependent and differences in productive capabilities are often amplified and evolve to reveal comparative advantages and disadvantages between value chain layers that can be the source of potential gains from trade between two or more parties (Jacobides, 2005; Jacobides & Hitt, 2005).

Therefore, governance modes are often likely to depend upon both efficiency considerations and heterogeneity in productive capabilities. Following, there is growing empirical support for the notion that transaction costs trade off against productive capability considerations (ie, Jacobides & Hitt, 2005; Jacobides & Winter, 2005 Poppo & Zenger, 1998; Schilling & Steensma, 2001; Walter & Veloso, 2008). However, as noted by Jacobides and Winter (2005), often such contributions focus on static, firm-level considerations, rather than focus attention on dynamic, systemic perspectives. Therefore, there remains a gap in our understanding of how transaction costs and productive capabilities combine dynamically across time.

In proposing a longer run, dynamic perspective, Jacobides & Winter (2005:400) argue that evolutionary forces can temporally change the respective trade-off between transaction costs and productive capabilities as (1) selection forces such as competition and imitation initially reinforce and amplify the static combination, however (2) productive capability differences are amplified, often reducing transaction costs to reveal gains from trade, which in turn (3) changes the scope and trajectory of productive capability development, which (4) allows new entrants to participate in the product market thereby changing the capability pool and its subsequent distribution. The implication of Jacobides and Winter's (2005) work is that in the long run firms can and do shape their own institutional environment to fit their own advantage (ie, Penrose, 1959). In sum, there is a continued need to move away from a static analysis of efficiency considerations and productive capabilities and shift to a more dynamic and longer run analysis of how they combine in context-specific settings to underpin architectural co-evolution.

I structure this section of the discussion as follows: section 5.2.1 examines the progression towards product modularity and industry specialization (1990-2012) and section 5.2.2 examines the progression towards product and industry integration (2012-2014), as highlighted previously in Figure 36.

5.2.1 Progression towards product modularity and industry specialization (1990-2012)

5.2.1.1. Phase one: the 90s – modular innovation

Table 29 provides a summary of the mechanisms underpinning modular innovation in the 90s.

Co-evolutionary cycle 1 – modular innovation	Process	Findings
1. Exogenous shocks	Exogenous shocks provide a discontinuity to the vertical architecture of the industry.	Exogenous shocks such as deregulation (1988), provide industry standards (ie, exogenous reductions in transaction costs) that promote product modularity in peripheral components and industry fragmentation
2. Expanded roster of industry participants	Exogenous shocks provide incentives for new specialised firms to enter the product market in all value chain activities, and establish a new mix of industry participants	Sets in motion a heterogeneous productive capability development process
3. Endogenous investments in mundane transaction costs	Exogenous shocks provide incentives to invest in modular product partitioning and standard-setting	Unit-linked insurance companies invest ex-ante in specialised standards to take advantage of heterogeneous productive capabilities in certain product component intermediate markets
4. Selection forces	Selection forces amplify differences in profitability or market share	With-profit insurance companies suffer decline in profitability and are prompted to imitate the unit-linked design
5. Heterogeneous productive capabilities	Specialisation amplifies the heterogeneity in productive capabilities, revealing gains from trade	Heterogeneity in productive capabilities and ex-ante investments in mundane transaction costs create a stable unit-linked product architecture and specialised vertical architecture

Table 29: Phase one: the 90s – modular innovation

In the mid to late-80s, insurance companies in the product market were vertically integrated, owing to the absence of efficient intermediate markets and, thus, reliance upon internal suppliers. However, exogenous shocks associated with deregulation in 1988 provided standards that attracted new unit-linked insurance companies to enter the product market in the manufacturer/assembler layer, as well as new entrants in upstream and downstream intermediate markets, causing industry fragmentation. Exogenous shocks also provided the impetus for product partitioning and standards to take root in the ‘peripheral’ fund and advice product components. I now turn to discuss the modularisation processes of these two product components.

5.2.1.1.1 Modularisation of fund components

With-profits insurance companies were vertically integrated. Developing Williamson’s (Williamson, 1983), three parameters of transaction costs⁶⁹, modularisation can be characterised in terms of a technological uncertainty that presented transaction hazards, especially in the presence of architectural and knowledge specificity associated with a closed and integrated product architecture. Furthermore, with-profits insurance companies had scale-efficient bureaucratic and production costs and no significant prior contracting relationships with external fund management groups which would seem likely to increase the perception of ex-post opportunism. With limited incentives to conduct market exchange, the sum of bureaucratic costs (BC) + internal production costs (IPC) can be characterised as less than external production costs (EPC) + transaction costs (TC), which provided an efficiency-based force in favour of vertical integration.

On the other hand, upstream fund management groups had significant productive capabilities owing to their specialisation and global reach, and hence greater incentives to expand their knowledge and initiate new technological developments. Interviews suggest, however, that with-profits insurance companies held a perception that the correlation between productive capabilities was quite strong owing to, in some cases, 40 year+ knowledge and experience in managing with-profits funds.

⁶⁹ Williamson (1983) cites uncertainty, frequency of contract upgrades and asset specificity as the primary drivers of transaction costs and, hence, vertical integration

As a consequence, no significant gains from trade were perceived as appropriate. Moreover, in the presence of a transaction environment where $BC + IPC < EPC + TC$, the product architecture and vertical scope remained closed and integrated. In other words, any meaningful differences in productive capabilities, net of transaction costs, were not sufficient enough to reveal gains from trade – at least for a short while.

For *unit-linked insurance companies*, there is a different story. New unit-linked insurance companies entered the product market with few existing productive capabilities in fund management. As such, there was a weak correlation in the productive capabilities of unit-linked insurance companies and specialised upstream fund management groups, resulting in a strong upstream comparative advantage revealing gains from trade. Possessing a near-modular IT architecture at the fund component level, ‘unit-linking’ across firm boundaries can be characterised as less of a technological change, and perhaps more as an ‘incremental innovation’, rather than a modular innovation.

Characterising the innovation mode as incremental has a number of potential implications, (1) incremental innovation is typically characterised as an innovation mode within firm boundaries, however I find, in contrast, that the extent of comparative advantage held by fund management groups is significant enough to push the task activity across firm boundaries, (2) the prior product partitioning and specialised ‘mirror fund’ interface standards for connecting fund components can be characterised as an ex-ante mundane transaction cost (ie, Baldwin, 2008) that also served to reduce threats of ex-post transaction costs, (3) investments in buyer-supplier relationships further served to mitigate transaction hazards, and (4) as gains from specialisation attracted more upstream fund management groups to the product market, small numbers bargaining problems reduced over time.

By the mid-90s, however, most *with-profit insurance companies* were sourcing modular fund components from upstream suppliers. Initially, with-profits insurance companies relied upon their own internal suppliers to play catch-up and replicate the modular innovation in order to mitigate against selection forces and a loss of market share. However, once internal suppliers had modularised fund components, moving the task activity across firm boundaries represented a lower level of technological change and hence lower transaction costs.

The process can be characterised as follows, (1) the codification process involved in modularising the product component interface subsequently permitted the task activity to be transferred across firm boundaries as a basis for efficient market exchange, (2) the intermediate market now offered a much wider range of specialised fund component suppliers, thereby reducing hazards associated with small numbers bargaining problems, and (3) the production costs of upstream fund management groups had reduced owing to scale efficiencies associated with supplying a number of manufacturer/assembler firms in the product market. In other words, the transaction cost environment was now pushing in the direction of modularisation and specialisation, as $BC + IPC > EPC + TC$.

With external fund management groups offering access to an almost limitless range of UK and global stock markets and indices, the distance between the new technological alternatives and the manufacturer/assemblers capability base grew substantially, revealing latent gains from trade. Therefore, in the case of *with-profits insurance companies*, I find that superior productive capabilities in the intermediate market and efficiency considerations eventually reinforce each other to promote modularisation at the fund component level and a disintegrated vertical architecture. In sum, new unit-linked insurance companies ‘going modular’ served to create selection forces for all firms in the product market so that there was eventually no choice but to eventually follow (ie, Sanchez & Mahoney, 2013).

5.2.1.1.2 Modularisation of advice components

Advice and distribution components moved across firm boundaries in the late-80s and early-90s for both *with-profits and unit-linked insurance companies* simultaneously. Prior to the emergence of Government-sponsored ‘industry standards’ for advice and distribution in 1988, interviews suggest that manufacturer/assembler firms held strong productive capabilities for the provision of advice and distribution.

However, 'standards' had two main impacts, (1) regulation embodied in the Financial Services Act 1986 significantly increased the bureaucratic costs associated with internal management of the activity, and (2) characterised as an ex-ante mundane transaction cost, industry standards codified the nature of the relationship between manufacturer/assembler firms and external IFA firms. As such, the 'modularisation' of advice changed the nature of the relationship between internal costs and external costs, such that $BC + IPC > EPC + TC$, providing an efficiency-based force for specialisation across the entire roster of product market participants.

Productive capabilities, however, seem to play only a minor or even non-existent role. The correlation of productive capabilities between manufacturer/assembler firms and downstream IFA firms can be characterised as strong, and under such circumstances it would be usual logic to suggest that negligible differences in productive capabilities would result in vertical integration, transaction costs permitting. However, in contrast, despite high external production costs (ie, the amount of commission demanded) and transaction costs (ie, ex-post opportunism), interviews suggest that the increases in internal bureaucratic costs arising from regulation were so significant as to reveal gains from trade from market-based exchanges.

5.2.1.2 Phase two: early to mid-noughties – architectural innovation

Table 30 provides a summary of the mechanisms underpinning the architectural innovation between 2000 and 2005.

Co-evolutionary cycle 2 – architectural innovation (“fund supermarkets”)	Process	Result
1. Exogenous shocks	Exogenous shocks provide a discontinuity to the vertical architecture of the product market	Exogenous shocks such as tax changes to CIS funds and the launch of ISAs provide further industry standards (ie, exogenous reductions in mundane transaction costs) that promote product modularity and industry fragmentation. Such shocks cannot be easily accommodated within existing stable unit-linked product architecture
2. New vertical scope	Exogenous shocks and limits to specialisation gains (ie, through imitation) provide incentives for existing participants to re-evaluate vertical scope	Upstream fund management groups forward integrate into product provision to leverage existing productive capabilities and limit volatility of specialisation gains
3. Expanded roster of industry participants	Exogenous shocks and new product architecture attract expanded roster of providers of complementary product components and sets in motion a new productive capability development process	Wide range of new upstream firms enter the market providing complementary product components creating network effects
4. Endogenous investments in mundane transaction costs	Exogenous shocks provide incentives to invest in further product partitioning and standard-setting	Standards agreed for connecting with numerous complementary product components
5. Selection forces	Selection forces amplify differences in profitability or market share	Insurance companies suffer decline in profitability/market share and are prompted to imitate by quickly licensing-in new technologies
6. Heterogeneous productive capabilities	Heterogeneous productive capabilities promote further specialisation	Specialisation gains begin to erode as a result of modular effects of commoditisation and imitation. Capability-destroying characteristics of new architecture lead insurance companies to develop systems integration capabilities

Table 30: Phase two: 2000 – 2005 – architectural innovation

The institutional structure of the product market presented in 5.2.1.1 was punctuated by an architectural innovation (“fund supermarkets”) in c2000, which resulted in further product modularity, but also forward vertical integration at the industry level, supporting the idea that architectures at different hierarchal levels may sometimes (temporarily) co-evolve in different directions.

Throughout the 90s, despite fund and advice and distribution components having moved across firm boundaries, interviews suggest that manufacturer/assembler firms maintained superior productive capabilities in unit-linked ‘pensions’ IT systems. Furthermore, at the industry level, gains from specialisation (ie, Jacobides, 2005) dominated a ‘stable’ vertical architecture as firms focused on task activities that yielded a comparative advantage and appropriable gains. For example, the institutional structure was characterised by (1) insurance companies with productive capabilities in IT and actuarial science, (2) fund management groups with productive capabilities in fund design and development, and (3) IFA firms who specialised in distribution and advice.

However, the stability of the unit-linked product architecture and specialised vertical architecture was disrupted not by manufacturer/assembler firms with existing productive capabilities in ‘pensions’ IT systems, but rather by upstream fund management groups who forward integrated into product provision. Many upstream fund management groups had productive capabilities in modular ‘fund management’ IT systems, owing to the presence of existing industry standards for fund design and development⁷⁰. 1999, however, saw the government-sponsored launch of Individual Savings Accounts (ISAs), a tax-efficient ‘modular’ product for medium-term savings and investments. More importantly, the Government issued ‘ISA rules’, a set of quasi industry-standards that, in the stroke of a pen, gave fund management groups a capability-based advantage⁷¹ enshrined in the ISA rules.

The launch of ISAs also occurred almost simultaneously with exogenous changes to the regulation and taxation of CIS funds designed by fund management groups that enabled them to also be held for the first-time directly in personal pensions (SIPPs), undermining the existing ‘mirror fund’ specialised standard propagated by insurance companies. As a consequence, fund management groups spotted an opportunity to not only forward integrate into ISAs but also SIPPs, and create, what the industry initially described as a multi-product ‘fund supermarket’.

⁷⁰ Fund components were designed according to UCITS standards, a European standard

⁷¹ ISA rules allowed investments into CIS funds offered by fund management groups, but not unit-linked mirror funds offered by insurance companies

While the architectural innovation was competence-destroying for incumbent insurance companies, the innovation mode can be characterised as ‘modular’ for fund management groups, as it leveraged existing productive capabilities. As a modular innovation from the perspective of fund management groups, forward integration into product provision occurred for a number of reasons (1) downstream firms (ie, insurance companies) possessed few (if any) productive capabilities in ISA product provision, fund management or modular IT systems, (2) despite the presence of standards, as a brand new product category the ISA product is associated with a degree of technological uncertainty, and (3) within the fund management segment of the value chain, the trajectories of productive capabilities had become so super-specialised such that, according to interviews, they had reached a maximum yield. In a product market where classes of investment move in and out of fashion in line with the so-called investment cycle, there was a strong need to develop productive capabilities at the product system level in order to capture additional rents, or at least minimise the volatility of rents in their existing product portfolio.

Few, if any, incumbent insurance companies held any substantive productive capabilities in the new multi-product systems architecture or its modular complementary product components. As a consequence, most insurance companies chose to quickly license-in fund supermarket architectures from fund management groups or from other external partners in the fund management supply chain, a type of ‘opening’ strategy that waives control of the commodity layer (West, 2003). However, despite these strategic risks, licensing the new technology allowed many insurance companies to take advantage of a number of opportunities: (1) it enabled them to reduce the sunk cost risks associated with technological uncertainty and to access new product markets (ie SIPP and ISA), (2) due to its greater modular properties, a number of providers of complementary product components had already begun to converge around the fund supermarket architecture, such as decision-based software tools and back-office suppliers, promoting supply-side network externalities and bandwagon effects (ie, Schilling, 1999; Sanchez, 2008), (3) it was becoming clear that fund supermarkets were quickly becoming established as a ‘supply-chain platform’ (Gawer, 2010; 2014) that would replace the existing unit-linked product architectures as a dominant design.

From the perspective of manufacturer/assembler firms, architectural innovations are often associated with vertical integration (Walter & Veloso, 2008), the opposite of the case here. This is because architectural innovations are often associated with (1) high levels of technological uncertainty, (2) unforeseen product component interdependencies, (3) high levels of ex-ante and ex-post transaction costs, and (4) the need for close coordination, better accomplished within firm boundaries. However, despite technological uncertainty in the product market, explanations for why insurance companies turned to market-based exchanges despite these risks appear to be related to (1) the very weak correlation in productive capabilities at both the architectural and complementary product component level, (2) the sheer necessity of ensuring speed to market to mitigate against strong selection forces, (3) if the architectural innovation is characterised as participation in an *existing* (largely) modular systems architecture, the presence of *a priori* standards across many complementary product components may have mitigated ex-ante transaction costs, (4) the existing long-term relationships with fund management groups may have eased ex-post transaction costs, and hence resulting in more efficient market exchanges, and, (5) interestingly, incumbent insurance companies remained sceptical that the new architecture would become embedded as a dominant design, and continued to vociferously defend existing technologies (ie, Schilling, 1999) at the same time. In other words, market-based exchanges can be characterised as an effort to minimise the risks of technological obsolescence and uncertainty (Walter & Veloso, 2008) and maximise the ‘option value’ associated with the new technology (Baldwin & Clark, 2000).

In other words, while insurance companies had previously held productive capabilities in unit-linked ‘pensions’ IT systems, the architectural innovation at the ‘system-level’ had reinforced the direction of co-evolutionary forces towards product modularity and industry specialisation. Following, efficiency considerations where $BC + IPC > EPC + TC$, and the weak correlation across productive capabilities both pushed in the direction of increased product modularity and industry specialisation.

For many insurance companies, the cycle of modular and architectural innovations throughout the 90s and the noughties had ‘hollowed-out’ organisational capabilities (ie, Parmigiani & Mitchell, 2009), with a loss of both core architectural and complementary product component knowledge. As a consequence, and as I discuss further in section 5.3, many insurance companies began to focus on developing systems integration capabilities (ie, Brusoni, *et al.*, 2001, Ceci, *et al.*, 2014). Perhaps through strategic necessity, insurance companies began to maintain a broader knowledge than task boundary across the value chain in order to acquire knowledge and productive capabilities in the new product architecture and ‘critical’ complementary product components by maintaining intensive inter-firm collaboration devices such as co-location with fund management groups and other complementary product component suppliers. In other words, manufacturer/assembler firms began to try to reverse the weak correlation of productive capabilities as gains from specialisation and trade began to erode.

5.2.1.3. Phase three: mid to late-noughties – architectural innovation

Table 31 provides a summary of the mechanisms underpinning the architectural innovation between 2005 and 2012.

Co-evolutionary cycle 3 – architectural innovation (“product platforms”)	Process	Result
1. Exogenous shocks	Exogenous shocks provide a discontinuity to the vertical architecture of the industry.	Exogenous shocks such as A-day provide further industry standards (ie, exogenous reductions in mundane transaction costs) that cannot be easily accommodated within the existing fund supermarket product architecture
2. Expanded roster of industry participants	Exogenous shocks provide incentives for further new firms to enter the product market	Upstream IT systems providers enter the product market that reinforces the existing specialised vertical architecture
3. Endogenous investments in mundane transaction costs	Exogenous shocks provide incentives to invest in further product partitioning and standard-setting	Further standards agreed for numerous complementary product components
4. Selection forces	Extreme modularity effects promote selection forces that begin to erode heterogeneity	Insurance companies leapfrog the fund supermarket architecture and license-in new platform technologies
5. Heterogeneous productive capabilities	Specialisation amplifies heterogeneity in productive capabilities in many upstream value chain segments, but limits specialisation gains in others	Insurance companies continue to focus on developing systems integration capabilities

Table 31: Phase three: 2005 -2012 – architectural innovation

It soon became apparent, however, that the fund supermarket architecture was not well-suited to operating as a ‘multi-product platform’, as it was based on a modular ‘fund administration hub’, and it remained difficult to design and develop a ‘product layer hub’ without designing a significant number of interdependencies and workarounds into the architecture.

By the mid-noughties a further systems-based architectural innovation disrupted the product market. Rather than the architectural innovation being instigated by firms in the manufacturer/assembler layer, or even by fund management groups who had earlier forward integrated into product provision, this time it was initiated by a number of new upstream platform systems IT architects who entered the market with open and modular IT architectures, with origins in Australia and New Zealand, where financial services product platforms had already become dominant in financial services product markets. Many incumbent insurance companies and a wide array of new entrants saw this as an opportunity to ‘technologically-leapfrog’ early-movers who had developed the fund supermarket architecture by licensing-in the new platform architecture technologies.

From a productive capabilities perspective, the correlation between the productive capabilities of the platform systems architects and manufacturer/assembler firms was weak, and as a consequence the IT systems architects held a significant comparative advantage and hence gains from trade were appropriable. In efficiency terms, the platform systems architects had committed significant ex-ante investments in standards that served to reduce mundane transaction costs for buyers. Furthermore, although the platform systems architecture was a new innovation in the UK product market, the degree of technological uncertainty had been minimised through its prior embeddedness and dominance in other international product markets. As such, with low asset specificity and low technological uncertainty, the presence of transaction costs can be characterised as low, despite the hazards associated with coordination and exchange between unfamiliar market counterparties. New, productive inter-firm relationships were required, and, in part, this was developed through an extended period of intense inter-firm collaboration between manufacturer/assembler firms, platform systems architects and providers of upstream complementary product components.

Interviews suggest that despite the known threat of commoditisation and imitation posed by extreme product and systems modularity, speed to market was, again, critical in the decision to license-in the new technology. The pensions simplification regime (“A-day”) was scheduled for 2006 which, at the stroke of a pen, unified all existing occupational and personal pensions rules into a coherent single set of rules – in other words, an ‘industry standard’ - that ultimately meant that consumers could consolidate all of their existing personal pensions, perhaps amassed from various pension schemes with previous employers, and personal schemes, into one product architecture – a SIPP. Furthermore, the ‘industry standard’ SIPP rules provided by government created a modular product standard, much like the ISA rules had done in 1999. As many in the industry stated, this was a ‘do or die’ or ‘winner takes all’ moment (ie, Schilling, 1999), and speed to market was essential.

In sum, from the perspective of manufacturer/assembler firms, there was a weak correlation between productive capabilities revealing gains from trade. In an environment where $BC + IPC > EPC + TC$, a continued push towards product modularity and industry specialisation ensued. However, manufacturer/assembler firms continued to invest in developing systems integration capabilities as a mechanism to re-acquire knowledge and productive capabilities across the product system and in critical complementary product components.

5.2.2. Progression towards product and industry reintegration (2012-2014)

5.2.2.1 Phase 4: 2012-2014 – integrative innovation

Table 32 provides a summary of the mechanisms underpinning integrative innovation between 2012 and 2014.

Co-evolutionary cycle 4 – integrative innovation	Process	Result
1. Exogenous shocks	Exogenous shocks provide a discontinuity to the vertical architecture of the industry.	Exogenous shocks such RDR provide further industry standards (ie, exogenous reductions in mundane transaction costs)
2. Selection forces	Product modularity effects limit gains from specialisation and trade	Insurance companies leverage existing systems integration capabilities to bundle product components to secure rents
3. Endogenous investments in integrative innovation	Selection forces promote investments throughout the value chain in integrative innovation	Emergence of guided modularity and vertical specificity as integrative innovation modes
4. Heterogeneous productive capabilities	Integrative innovation amplifies heterogeneity in productive capabilities	Acquisition of selected product component suppliers amplifies heterogeneity and reduces roster of participants

Table 32: Phase four: 2012 – 2014 – integrative innovation

Despite the commonality of core product components in the open and modular platform architecture, licensees could request co-development to establish proprietary product components, so long as the systems’ interface standards were maintained. However, licensing agreements only permitted exclusivity periods of c12-18 months, from which time the product component would be offered to other licensees as part of the standard licensing package. This created a significant challenge for manufacturer/assembler firms to sustain gains from trade and resist commoditisation and imitation threats. As West (2003) reminds us, without innovation, differentiation or some form of lock-in, it will be very difficult to appropriate gains. Furthermore, with c35-40 firms now competing in the manufacturer/assembler layer of the value chain, the product development process was slow and protracted, owing to limited capacity and responsiveness from platform systems’ suppliers.

In response, many manufacturer/assembler firms decided to try to appropriate rents through foregrounding the marketing of their own proprietary fund components, in line with West's (2003) suggestion that marketing and branding may offer the only opportunity to secure rents. However, this approach had limited success owing to the superior and specialised productive capabilities of external fund management groups, and the multi-vendor configurations demanded by consumers (ie, Schilling, 1999; 2000). By 2012, coupled with further exogenous shocks, the need to innovate for differentiation and competitive advantage was pressing in order to deter continued commoditisation and imitation. As a consequence, 'integrative innovation' came to the fore (ie, Fixson & Park, 2008), leveraged from the systems integration capabilities developed and enhanced throughout the noughties.

Types of integrative innovation			
	Guided modularity	Vertical specificity	Insourcing
Product component specificity	Low	High	High
Adherence to standards	Yes	Yes	Yes
Use	Open	Restricted	Both
Governance mode	Market	Hybrid	Vertical integration

Table 33: Integrative innovation modes

According to interviews, integrative innovation took three different shapes, as shown in Table 33. First, manufacturer/assembler firms bundled⁷² combinations of modular product and process components for particular market segments. The modular bundles consisted of peripheral product and process components, governed via weak organisational ties with upstream and downstream suppliers, and supported by marketing and branding. For instance, the bundles would be heavily marketed to the downstream IFA channel as a 'governed' proposition that would save them time and reduce compliance risk. To participate in the bundles, fund management groups and providers of complementary product components were squeezed ex-post so as to create rents for the manufacturer/assembler firm.

⁷² I shall call this approach 'guided modularity'

This approach to ‘integrative innovation’ is similar to West’s (2003) description of bundling in the IT platform and computer games industries as a method to secure differentiation and competitive advantage in the face of open source software. However, the extent of product component specificity within the bundles was often marginal owing to weak incentives to participate and invest in integrative innovative activities.

Second, modular product and process components were bundled for the specific and exclusive use of particular value chain combinations or, using different terminology, various heterogeneous ‘power blocks’ in the product market⁷³. Each so-called ‘power block’ consisted of (1) the lead manufacturer/assembler firm, (2) the platform systems architect, (3) a number of participating upstream providers of complementary product components, and (4) a downstream IFA firm or network. Coordinated by the lead manufacturer/assembler firm, the modular product components would often be significantly reengineered within boundaries to suit the specific performance requirements of the power block, but maintaining adherence to the standards-based architecture. Interviews suggest that the key differences between guided modularity bundles and ‘vertical specificity’ bundles was (1) the extent of the reengineering or specificity of the inner-workings of the product and process components, (2) its exclusivity of use within the power block, and (3) the strength of the organisational tie, as often power blocks shifted to hybrid forms of governance, suggesting that as modularity reaches its architectural limits, firms may replace weak organisational ties across a diverse range of external firms and instead establish stronger ties with fewer players (ie, Chesbrough & Prencipe, 2008).

⁷³ An approach I shall call vertical specificity

Seen through a productive capabilities lens, interviews suggest that vertical specificity can be characterised as a leveraging of productive capabilities in systems integration by manufacturer/assembler firms in order to appropriate gains from trade. Moreover, high-levels of information-sharing and co-development were enacted within the arrangement, including co-location for the design and development of reengineered product components. Furthermore, the closer organisational ties between participants in the power block served to establish parameters for how ex-ante investments and ex-post surplus would be divided and can also perhaps be characterised as an attempt to further increase efficiencies in a product market characterised by commoditisation and imitation.

A third integrative innovation mechanism was insourcing the design and development of previously outsourced ‘critical’ product components. According to interviews, precisely which product components were insourced depended upon the specific particulars of firm-level strategy and how different firms saw opportunities for differentiation and competitive advantage. In all cases, the modular properties of the insourced product components were maintained in order to ensure interoperability with the systems-based platform architecture, however it may be the case that integrative innovation eventually leads to a decrease in the overall modularity of the product system.

Insourcing occurred via acquisition of the product component supplier in order to acquire superior productive capabilities at the product component level, and the findings in chapter 4 include many examples of acquisition including upstream providers of commercial property, discretionary fund managers, stockbrokers, and downstream IFA firms. Moreover, insourced product components were also often reengineered and incorporated into power block configurations, so as to positively influence the ex-post division of surplus within the arrangement. Despite the low transaction cost environment associated with the platform architecture, acquisition of product component suppliers can also be characterised as an investment in mundane transaction costs to establish the product component within firm boundaries (Sanchez, *et al.*, 2013).

Section summary and contribution

This section summarises the discussion in sections 5.2.1 and 5.2.2 and identifies the research study's main contribution to the extant literature on architectural co-evolution.

First, I highlight the importance of a dynamic, co-evolutionary perspective towards the interaction between efficiency considerations and productive capabilities, and the findings support the work by Jacobides and Hitt (2005) and Jacobides and Winter (2005) in the US mortgage banking sector where vertical disintegration took root in the value chain where the correlation across productive capabilities was weak and prior ex-ante investments were made in standards, reducing transaction costs. However, I extend prior work by arguing that architectural co-evolution may be examined not only at the firm and systemic level, but also at the level of the individual product component, as firm-level decisions concerning the degree of modularisation and hence firm boundaries occur at the product component level. In other words, assuming the product component is modular with limited co-complementarity, I demonstrate that the boundaries of product components are also the boundary in which firms consider efficiency and productive capabilities trade-offs such that 'thin crossing points' (Baldwin, 2008) in the component boundary enabled by ex-ante standards are often the points at which task and firm boundaries break apart. Moreover, I also show how product modularity and industry specialisations can reverse as the commoditisation and imitation threats posed by ever-increasing modularity limit gains from specialisation and trade, and erode productive capabilities. I also show how the trade-offs at the individual product component level can impact the overall aggregate degree of product architecture modularity and vertical scope of the product market.

Figure 38 summarises how efficiency considerations and productive capabilities may impact upon product component innovation and vertical scope. The figure is not intended to suggest cause-and-effect but rather the association between processes observed in the study.

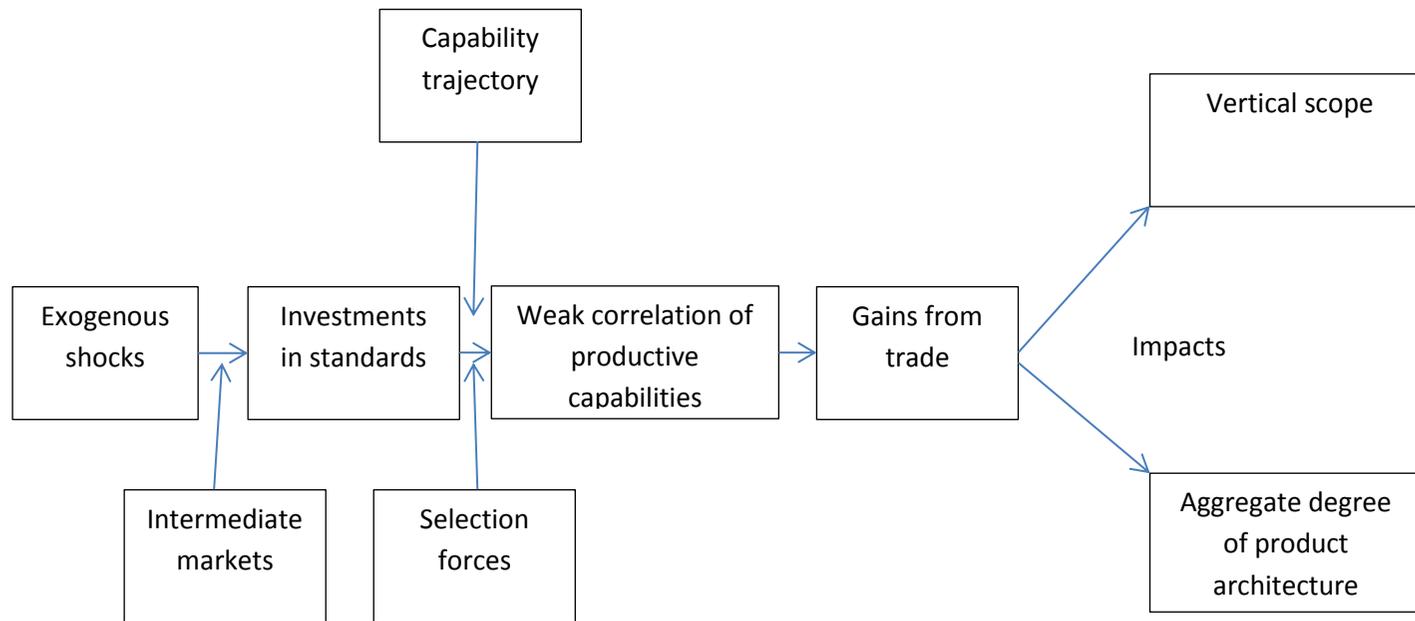


Figure 38: Dynamic effects of productive capabilities and efficiency considerations on modular product component evolution

The findings suggest that firms are likely to remain integrated in the product component and the associated task activity where a firm possesses superior productive capabilities, regardless of prior ex-ante investments in standards. This scenario is indicative of product components in closed and modular and hybrid product architectures. On the other hand, selection forces and specialised productive capability trajectories in time weaken the correlation in productive capabilities and hence net gains from trade eventually ensue. In other words, subject to an efficient intermediate market, ex-ante investments in standards combined with superior productive capabilities in the intermediate market may propel a closed and modular product component in one time period across firm boundaries in a future time period.

As product architectures are made up of a number of multiple product components, the different productive capabilities and efficiency trade-offs at the product component level may help illuminate the under-explored phenomena of hybridity at the aggregate level of the product architecture. It is probable that most product architectures exhibit different degrees of hybridity depending upon the precise productive capabilities and efficiency trade-offs occurring at the individual product component level at any one temporal cross-section. In other words, subject to prior investments in standards, product architecture hybridity may evolve towards or away from greater modularity dependent upon the level of ex-ante investments in standards and the degree of heterogeneity in productive capabilities at the individual product component level.

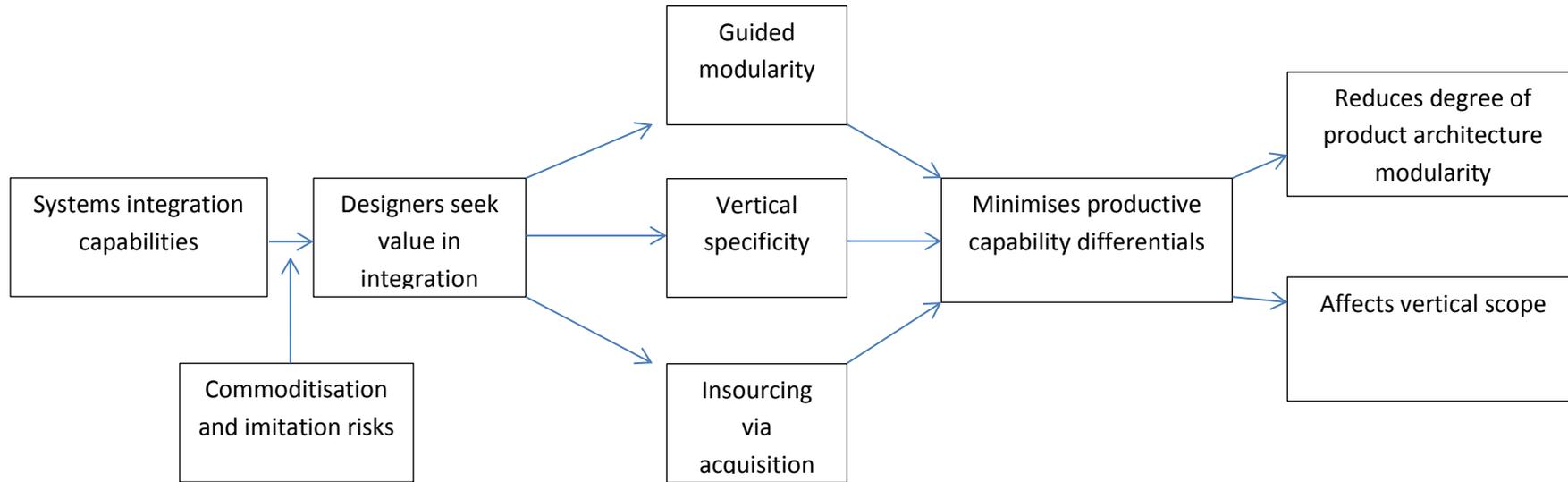


Figure 39: Dynamic effects of productive capabilities and efficiency considerations on integrative innovation

Second, I contribute to our understanding on the types of integrative innovation pursued by manufacture/assembler firms faced by the increasing modularity effects of commoditisation and imitation. Figure 39, again, is not intended to suggest cause-and-effect but rather the association between processes observed in the study. As shown, in response to the erosion of architectural and product component knowledge, I find that manufacturer/assembler firms turned to the development of systems integration capabilities in one time period (c2000) that led to three types of integrative innovation in another time period (2012-2014). Interviews suggest that, first, guided modularity is an integrative innovation mechanism characterised by (1) modular product and process component bundling, (2) limited investments in product component specificity, and (3) weak market-based organisational ties. Second, vertical specificity is characterised as comprising of (1) modular product and process component bundling, (2) significant investments in specificity within product component boundaries while maintaining adherence to standards, (3) exclusivity of use, (4) high-levels of knowledge exchange and knowledge integration, and (5) hybrid governance modes. Finally, product component insourcing via acquisition is characterised as (1) significant investments in specificity within product component boundaries, while maintaining adherence to standards, (2) a form of forward or backward vertical integration to acquire product component suppliers and associated productive capabilities.

In other words, rather than reintegration taking root in the core platform architecture, control remained waived, and integrative innovation focused on ‘critical’ product components that offered manufacturer/assembler firms some form of differentiation or competitive advantage. For instance, coordinated by the manufacturer/assembler firms, the power blocks created high integrative mechanisms, switching costs, and shifted competition towards a ‘between power blocks’ mode, a quasi-integrated product market structure. Moreover, insourcing critical product components via acquisition had the benefit of acquiring critical productive capabilities, positively influencing the division of surplus, and also reduced the roster of firms available in the intermediate market.

Supporting the findings of Fixson and Park (2008) in the bicycle industry, the integrative innovation mechanisms are also endogenously-determined. Designers seeking value in integrative innovation can and do affect the degree of co-specialisation among product components which, in turn, may enhance product performance outcomes and the availability of gains. However, whereas Fixson and Park (2008) highlight the role of Shimano designing non-modular product components, in contrast, I find that integrative innovation occurs largely within existing product component boundaries, increasing specificity, but maintaining adherence to the standards-based platform architecture. In other words, product components remain interoperable with the standards-based architecture.

However, reintegration occurred more tightly at the product market level via quasi-integrated power blocks governed via hybrid mechanisms such as alliances and, in some cases, joint venture arrangements, as well as forward and backward vertical integration through acquisition of product component suppliers, supporting the notion that reintegration at the product level can have profound implications for product market structure.

Third, I contribute to our understanding of systems integrator firms. I show how the process of developing and maintaining systems integration capabilities in one time period may be associated with an 'option value' of reintegration in a subsequent time period. For example, (1), around 2000, interviews suggest that following the erosion of productive capabilities in the core systems architecture and peripheral product components, insurance companies turned to the development of systems integration capabilities and maintained close collaborative relationships with upstream and downstream suppliers as a mechanism to survive. By 2005, (2) as the architectural and product component knowledge of insurance companies grew, and new platform systems architects entered the product market, insurance companies initiated deeper and more intensive knowledge integration mechanisms across the value chain encompassing devices such as product component co-development, temporary co-location, and employee-rotation. By 2012, (3) insurance companies leveraged its system integration capabilities to form quasi-integrated power blocks in the product market, based on hybrid governance forms and insourced, via acquisition, productive capabilities in component design where the product component offered a form of differentiation or competitive advantage.

As shown in Table 34, the narrative suggests that the development of systems integration capabilities was (1) instrumental in coping with discontinuous, competence-destroying, architectural innovation, something which Henderson and Clark (1990) argue is difficult for incumbent firms to do, (2) remaining integrated in the knowledge domain for product components with the potential for differentiation or competitive advantage allowed insurance companies to strengthen the previously weak correlation in productive capabilities across time (as noted by Zirpoli and Becker, 2007; Zirpoli & Camuffo, 2009), and (3), in many cases, it eventually led to hybrid forms of governance within quasi-integrated power blocks and insourcing of product component via forward and backward vertical integration.

Although I stop short of hypothesising a direct cause-and-effect relationship across time periods, it seems highly probable that the development of systems integration capabilities in one time period may provide an ‘option value’ of integrative innovation in a subsequent time period. Without such systems integration capabilities, it seems likely that firms would lose the productive capabilities required to innovate at the product system level, and fall into the ‘modularity trap’ as envisaged by Chesbrough and Kusunoki (2001).

Time period	Time period	Time period
Early to mid-noughties	Mid-noughties to late-noughties	2012-2014
Systems integrator (1)	Systems integrator (2)	Systems integrator (3)
Develops SIC to survive in response to architectural innovation and prior modular innovations. Maintains broader knowledge than task boundary to strengthen weak correlation in productive capabilities	Extends SIC via component co-development, co-location and employee exchanges across the value chain	Leverages SIC through creation of power blocks and hybrid governance Insources superior productive capabilities in component design via acquisition, where component offers differentiation or competitive advantage
Primarily market-based transactions	Primarily market-based exchanges	Hybrid and some vertical reintegration

Table 34: Evolution of systems integration capabilities

5.3 Architectural correspondence

Table 35 is a summary of the correspondence v. non-correspondence of product component, task, knowledge and firm boundaries between 1984 and 2014. I show the degree to which correspondence is evident by drawing upon Colfer and Baldwin's (2010) classification of 'support', 'no support', 'partial support' and 'mixed support' (although for simplicity, I collapse 'partial' and 'mixed' into one overall classification of 'partial support').

	Mid to late-1980s	Early to mid-1990s	Mid-90s to 2000	2000-2005	2005-2014
	Closed and integrated	Closed and modular	Hybrid (unit-linked)	Hybrid (fund supermarket)	Platform – open and modular (2005-2012) and Hybrid (2012-2014)
Product component type	Integrated product component	Some product components partitioned and loosely-coupled within firm boundaries	Some product components partitioned and loosely-coupled, within and across firm boundaries	Some product components partitioned and loosely-coupled, within and across firm boundaries	Most product components partitioned and loosely-coupled, within and across firm boundaries, some reintegrated
Task and firm boundaries	Task boundaries partitioned based on functional specialism (partial)	Where product/process components are loosely-coupled, product component and task boundaries are partitioned (support) but task remains within firm boundaries (support)	Where product/process components are loosely-coupled, product component, task and firm boundaries are partitioned (support)	Where product/process components are loosely-coupled, product component, task and firm boundaries are partitioned (support)	Where product/process components are loosely-coupled, product component, task and firm boundaries are partitioned (support)
Knowledge boundaries	Knowledge boundaries partitioned across functional domains (partial)	Architectural knowledge decoupled from product component knowledge (support)	Architectural and product component knowledge held by systems integrator (no support). Product component knowledge held by component teams (support)	Architectural and product component knowledge held by systems integrator (no support). Product component knowledge held by component teams (support)	Architectural and product component knowledge held by platform systems integrator (no support). Product component knowledge held by component teams (support)
Information-sharing (IS)	High within functions, but low across and between functions (partial)	No data	Low between component development teams (support), but high with systems integration team (no support)	Low between component development teams (support) and high with systems integration team (no support). In some cases, ex-ante IS replaces ex-post IS (partial)	High between component development teams (no support), and high with platform systems integration team (no support). In some cases, ex-ante IS replaces ex-post IS (partial)
Firm membership	Firm co-membership (support)	Firm co-membership (support)	Integrated components = firm co-membership (support), modular components = no firm co-membership (support)	Integrated components = firm co-membership (support), modular components = no firm co-membership (support)	Integrated components = firm co-membership (support), modular components = no firm co-membership (support)
Location	Co-location (support)	Co-location (no support)	Loosely-coupled component teams geographically-dispersed (support)	Geographical dispersion only where product component is loosely-coupled and low value add (partial)	Geographical dispersion only where product component is loosely-coupled and low value add (partial)

Table 35: Summary of correspondence across time

5.3.1 Stage 1: Mid to late-1980s (closed and integrated product architecture)

In the mid to late-1980s, interviews suggest that the with-profits personal pensions product architecture is characterised as closed and integrated. In the case of integrated product designs, there is unlikely to be a clear mapping of product component, task and knowledge boundaries and, as such, high-levels of information-sharing and co-location is often required to progress its design and development, and hence product development tasks are likely to be encompassed more efficiently within a single firm boundary (ie, Campagnolo & Camuffo, 2010; Colfer, 2007, Colfer & Baldwin, 2010).

In the period between 1984 and 1988, all firms in the product market exhibited high-levels of vertical integration, owing primarily to the absence of an intermediate market (with the exception of fund components) and high levels of architectural and knowledge specificity of the actuarial product development team who led the product development process. The actuarial product development function encompassed integration of architectural and product component knowledge of the core product design tasks such as product architecture development, pricing, and with-profits fund design. Business process tasks such as marketing, customer service, and distribution were partitioned on the basis of functional specialism, and all functions shared firm co-membership, supporting the view of Sanchez and Mahoney (1996) that integrated product development is enacted within a single firm.

The extant literature assumes that high-levels of information-sharing are necessary between all product development actors throughout an integrated product development process (ie, Colfer & Baldwin, 2010; Galbraith, 1973), in order to resolve product interdependencies, and reduce information asymmetries. However, actuarial science is well-known in the product market for its function-specific dialect and language, and this may illuminate why information-sharing within the actuarial product development team was high, and co-location necessary. However, information-sharing between the actuarial product development team and the other supporting business functions (such as customer services, marketing, distribution) was comparatively low, an unusual finding.

For instance, many respondents referred to the presence of ‘Chinese walls’ that existed between the actuarial product development team and other business functions (perhaps owing to information asymmetries, routines, etc), often resulting in a number of iterative product redesigns and reworks. A close reading of the interviews suggests that it is probable that the transmission of rich, tacit and technical knowledge across all product development actors represented a high information cost (Galbraith, 1973), even within-firm boundaries.

Co-location across all teams, however, was a dominant product development mode, perhaps owing to the smaller-size of firms in this period. In terms of correspondence, then, product development actors shared the organisational tie of firm co-membership and co-location (ie, Colfer & Baldwin, 2010), but the amount of information-sharing between actors offers only partial support.

In general terms, there is support for correspondence based on product component integration, and the organisational ties of firm co-membership and co-location. However, the picture is not straight-forward. Task and knowledge boundaries were not integrated, and instead partitioned across functional boundaries, with low-levels of information-sharing between functional teams, as shown in Table 36.

Closed and integrated (with-profits) mid to late-80s	
Product component integration	Correspondence
Task integration	Partial – correspondence within actuarial product development, but non-correspondence across separate functionally specialist teams
Knowledge integration	Partial – correspondence within actuarial product development, but non-correspondence across separate functionally specialist teams
Firm co-membership	Correspondence
Co-location	Correspondence
Information-Sharing (IS)	Partial – correspondence within actuarial product development, but non-correspondence across separate functionally specialist teams

Table 36: Correspondence summary: closed and integrated product architecture

5.3.2 Stage 2: early to mid-90s (closed and modular product architecture)

By the early to-mid-1990s, the product architecture evolved towards a closed and modular design. Although the core unit-linked ‘pensions’ IT systems architecture remained tightly-coupled and integrated throughout the period, the fund, advice and customer service components modularised, underpinned by specialised standards. Via the creation of fund and customer service teams, task and knowledge partitioning ensued within firm boundaries corresponding to the modular characteristics of the product/process components.

In other words, the product component partitioning acted as a map for task and knowledge partitioning, but is not (as yet) associated with significant outsourcing. The findings have a strong resonance with the study of Black and Decker by Lehnerd (1987), where Black and Decker were able to design a family of modular products via collaboration across different intra-firm functional units, and the study of computer notebooks by Hoetker (2006) where product component modularisation need not be associated with the movement of task activities across firm boundaries. This finding highlights that product component modularity *per se* is not enough to shift tasks across firm boundaries. As discussed in section 5.2, a weak correlation in productive capabilities and gains from trade must also ensue (ie, Jacobides & Winter, 2005) and hence it is possible that temporal delays in reconfiguring firm boundaries to reflect co-evolutionary changes in productive capabilities and efficiency considerations may be one possible explanation.

In terms of co-location v. geographical distance, co-location remained the dominant product development mode. This is possibly due to the continued small-size of firms at this point in time, and a temporal delay in shifting organisational structures to reflect modular product component and task boundaries. Temporal delay was highlighted as a potential explanation of non-correspondence by Zirpoli and Camuffo (2009) in their study of the motor vehicle industry.

In terms of information-sharing, I do not have sufficient meaningful data to make any observations, although it seems reasonable to assume that co-location may have resulted in high-levels of information-sharing ex-ante (to agree specialised standards) and ex-post (to coordinate activities across numerous product or process component teams).

As discussed in 5.3.1, in the mid to late-80s the existing organisational structure produced a number of coordination problems owing to substantial design iterations and reworks. By the early to mid-90s, however, this was addressed through the emergence of a ‘product owner’ role. In this context, a ‘product owner’ is similar to the notion posited by Henderson and Clark (1990:11) of someone with ‘architectural knowledge’ – *“knowledge about the ways in which components are integrated and linked together into a coherent whole”*.

According to interviews, architectural and product component knowledge was more or less partitioned, whereby the product owner focused on product architecture design, whereas product and process development teams focused on component level design. Product and process component development teams tended to work independently of each other and exchanged information via the product owner. It seems probable that there was minimal overlap between the component knowledge of product and process component development teams and the architectural knowledge held by the product owner.

Closed and modular (unit-linked) early to mid-90s	
Product component partitioning	Funds and customer service (within-firm) and advice (across-firm)
Task partitioning	Correspondence
Knowledge partitioning	Correspondence
Firm co-membership	Correspondence (with the exception of advice)
Co-location v. geographic dispersion	Correspondence (with the exception of advice)
Information-Sharing (IS)	No data

Table 37: Correspondence summary: closed and modular product architecture

5.3.3 Stage 3: mid-90s to 2000 (hybrid unit-linked product architecture)

By the mid-90s, there are a number of changes in organisational structure. Product development is now coordinated and orchestrated by a systems integrator who holds architectural knowledge and orchestrates and coordinates the design and development of product and process components. In this context, the systems integrator concept is similar to the distinction made by Ceci, *et al.*, (2014) between a static systems integrator and a dynamic systems integrator. Here, the systems integrator role can perhaps be conceptualised as a static systems integrator, where architectural knowledge is used ex-ante to define product component partitioning schemes and to resolve system-wide issues ex-post.

Co-location becomes less of a dominant product development mode. In many firms, the systems integrator and business functions remain co-located. Whereas the modularisation of fund, advice and customer service components is associated with so-called geographically-distant ‘centres of excellence’.

Despite a product component and task correspondence, interviews suggest that ex-post information-sharing is channelled through the systems integrator who coordinates and orchestrates the various task activities both within and across firm boundaries, and resolves system-wide issues. The high-levels of information-sharing between the systems integrator and the component development teams is associated with the broader knowledge than task boundary maintained by the systems integrator. In contrast, ex-post information-sharing between the de-coupled product and process component development teams was low, owing to a clear partitioning of product component, task and knowledge boundaries facilitated by the presence of standards. This is shown in Figure 40.

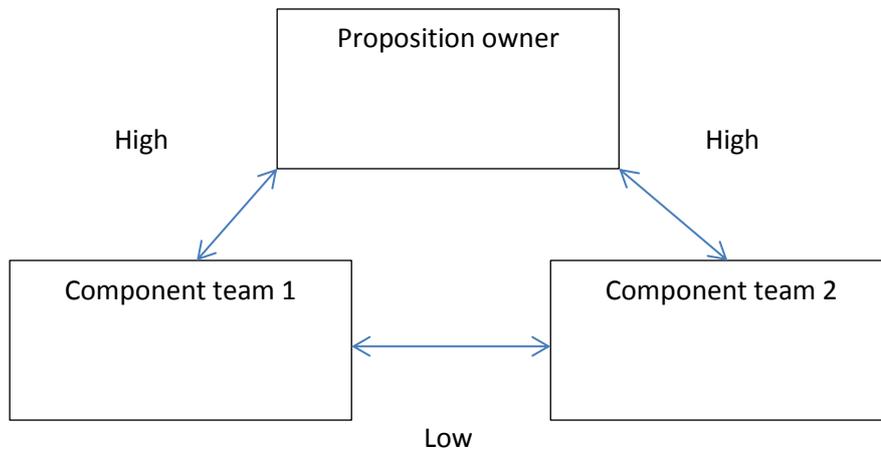


Figure 40: Information-sharing mid to late-90s

Throughout the 90s, the unit-linked product architecture can be characterised as ‘stable’ (relative to post-2000) and, as such, this research study builds upon the work of Cabigiosu and Camuffo (2012) in the air-conditioning industry who found that depending upon whether the product architecture can be characterised as ‘stable’ or ‘complex’ impacted ex-post levels of information-sharing such that low-levels of ex-post information-sharing were observed across firm boundaries under conditions of a stable product architecture, and high-levels of ex-post information-sharing were observed in the context of complex product development.

In general terms, I find resonance with their findings. However, I refine their observation further by highlighting that high-levels of ex-post information-sharing are maintained by the systems integrator, even under conditions of product architecture stability in order to resolve system-wide issues as they arise. In other words, modular product architecture stability may not be a sufficient condition to minimise information-sharing needs.

Hybrid (unit-linked) mid-90s to 2000	
Product component partitioning	Customer service (within-firm), funds and advice (across-firm)
Task partitioning	Correspondence
Knowledge partitioning	Correspondence (with the exception of the 'proposition owner')
Firm co-membership	Correspondence
Co-location v. geographic dispersion	Correspondence
Information-Sharing (IS)	Correspondence between component development teams, Non-correspondence with proposition owner

Table 38: Correspondence summary: hybrid (unit-linked) product architecture

5.3.4. Phase 4: 2000-2005 (hybrid fund supermarket architecture)

From c2000, the systems integration teams evolved in a quasi-SBU structure, with accountability for product profit and loss metrics. The evolution is accompanied by two mechanisms, (1) a competence-destroying architectural innovation in 2000 and, (2) the strategic need to develop and maintain outsourced architectural and product component knowledge.

The need to develop and maintain external product component knowledge is critical in this period. The modular innovations throughout the 90s eroded product component knowledge, and the architectural innovation in 2000 served to erode any remaining architectural capabilities, leaving incumbents with 'old' and out-dated architectural and product component knowledge. Hence, insurance companies simply did not hold any meaningful productive capabilities and so turned to a strategy of developing systems integration capabilities in order to survive. This observation is associated with Cabigiuso, *et al.*, (2013) study of air-conditioning systems in the motor vehicle industry, where the authors find that co-development practices with external suppliers were foregrounded as a result of the manufacturer/assembler firms lack of product component level knowledge.

As a consequence, information-sharing requirements also evolve. Interviews suggest that (1) ex-post information-sharing requirements continue to be low between modular product and process component development teams and, (2) high between component development teams and the systems integration team, both within- and across firm boundaries. However, the architectural innovation in c2000 triggers a need to vary ex-post information-sharing needs, depending upon the value characteristics of the product component type. For instance, (1) where a product component is perceived as low-value, ex-ante information-sharing is high in order to develop standards, followed by low levels of ex-post information-sharing. In such cases, geographical dispersion is considered an appropriate governance mode. In contrast, (2) where a product component is perceived as high-value (ie, it offers some form of differentiation or competitive advantage), high-levels of both ex-ante and ex-post information-sharing are maintained throughout. In such cases, temporary co-location is considered an appropriate governance mode. This is summarised in Table 39.

	Low value modular component	High value modular component
Knowledge integration	Low	High
Ex-ante IS	High	High
Ex-post IS	Low	High
Location	Geographic distance	Co-location

Table 39: Value characteristics of product components on knowledge integration, information-sharing and location

This observation has some similarities with Furlan *et al.*, (2014) who argue that product modularity only works as a substitute for inter-firm coordination when product components are stable. I find congruence with this observation but moreover, the product component must also be perceived as low-value. I also find resonance with Cabigosu, *et al.*, (2013:673) who remark that product modularity has limited traction in easing inter-firm coordination, but I find that this observation only holds to the extent that the modular product component offers some form of differentiation and/or competitive advantage. In order words, inter-firm coordination is mediated by the firm's strategy and value proposition, supporting the observation by Zirpoli and Camuffo (2009).

In sum, firms rely on knowledge integration, extensive information-sharing throughout, and co-location in order to co-develop outsourced high-value modular product components. On the other hand, where the modular product component is low-value, product modularity does seem to act as an embedded coordination mechanism a la Sanchez and Mahoney (1996) as firms choose to invest ex-ante in defining standards, and then revert to a low-level of ex-post information-sharing and geographical distance. This finding may help illuminate further the suggestion by Cabigiosu and Camoffo's (2012) that many firms may rely on different strategies for product and organisation correspondence. For instance, I find that firms rely on a 'trade-off hypothesis' where the product component is low value, and the 'complementarity hypothesis' where the product component is high-value, such that "one strategy does not exclude the other" (Cabigiosu & Camuffo, 2012:699).

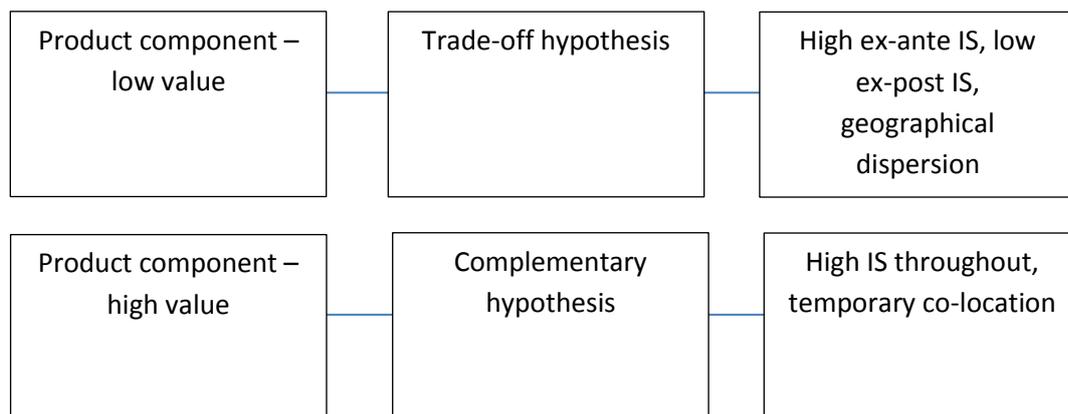


Figure 41: Varying information-sharing needs based on value characteristics of the product component

A correspondence summary between 2000 and 2005 is shown in Table 30.

Hybrid (fund supermarket) 2000-2005	
Product component partitioning	Customer service (within-firm), funds and advice (across-firm) Explosion of loosely-coupled complementary products (across-firm)
Task partitioning	Correspondence
Knowledge partitioning	Correspondence between component development teams, Non-correspondence with systems integrator. Knowledge integration for high value components
Firm co-membership	Correspondence
Co-location	Partial correspondence
Information-Sharing (IS)	Correspondence between component development teams, Non-correspondence with systems integrator

Table 40: Correspondence summary: hybrid (fund supermarket) product architecture

5.3.4. Phase 5: 2005-2014 (open and modular and hybrid platform architecture)

The emergence of a technological platform architecture in 2005 can be characterised as a further architectural innovation (Henderson & Clark, 1990), but one that did not render existing architectural or systems integration knowledge redundant. The evolution to a multi-product platform is accompanied by the evolution of systems integration teams into platform integration teams. In other words, the product moves down a level in the overall systems hierarchy to become, in effect, a component of the platform architecture. Product systems integration teams subsequently become much smaller, and platform integration teams much larger as a result.

Knowledge integration, high levels of information-sharing and co-location for high value components is maintained throughout the period between 2005 and 2014, perhaps owing to the fact that once a firm decides to act as a systems integrator and ‘know more than they do’ (Brusoni, *et al.*, 2001), further changes in product architecture may not lead to subsequent changes in firm architecture.

However, the emergence of multi-product platforms is associated with changes to ex-post information-sharing between product and process component development teams, as shown in Figure 42.

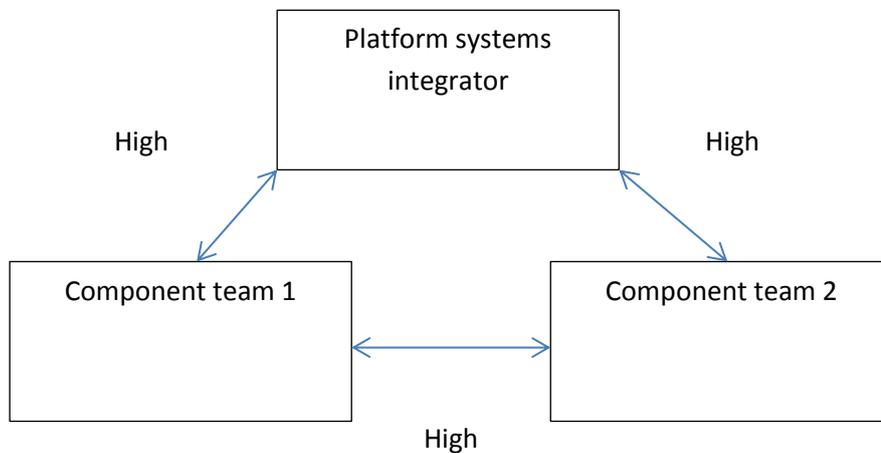


Figure 42: Information-sharing needs 2005 onwards

Whereas between the mid-90s and mid-noughties, product and process component development teams shared minimal information between them, but high amounts of information with the system integration team, the emergence of a multi-product platform architecture changes this somewhat. Now, ex-post information-sharing increases not only between component development teams and the platform systems integration team, but also between the loosely-coupled component teams themselves, perhaps owing to increasing product complexity and unforeseen interdependencies at the platform system level. For example, fund component teams needed to share information not only with the platform integration team, but also the product team, and providers of complementary product components in order to ensure effective integration, despite the presence of standards. Therefore, it is probable that as product complexity increases, the intensity of information-sharing across the entire system increases, perhaps owing to technological newness (Zirpoli and Becker, 2007), uneven rates of product component change and to resolve unforeseen system design issues (Busoni, *et al.*, 2001; MacDuffie, 2013).

Platform architecture 2005-2014	
Product component partitioning	Customer service (within-firm), funds and advice (across-firm) Explosion of loosely-coupled complementary products (across-firm)
Task partitioning	Correspondence
Knowledge partitioning	Correspondence between component development teams, Non-correspondence with platform systems integrator
Firm co-membership	Correspondence
Co-location	Partial correspondence
Information-Sharing (IS)	Non-correspondence

Table 41: Correspondence summary: open and modular and hybrid platform architecture

Section summary and contribution

This section summarises the discussion in section 5.3 and identifies the research study's main contribution to the extant literature on architectural correspondence.

First, I show the extent to which architectural correspondence is sustained in the face of co-evolutionary forces across time. By charting the correspondence across a 30-year period, I have shown how product markets may evolve through periods of correspondence and non-correspondence. Although I have been unable to isolate the effect of temporal delays, the findings show that there is no simple relationship between the design characteristics of product, firm and industry architecture. Whether correspondence is evident at any temporal cross-section seems to depend upon a number of factors at both the product architecture *and* product component level, such as (1) the stylised productive architecture type, (2) the degree of product complexity, (3) whether the product component is integrated or modular, and (4) the degree to which the product component has differentiation or competitive advantage properties. Furthermore, it may also depend upon the target for correspondence, for instance (1) product component boundary, (2) task activity boundary, (3) knowledge boundary, or (4) firm boundary. Moreover, how correspondence is measured plays an important role and different results may be obtained depending upon whether an empirical study examines (1) information-sharing needs, (2) location, or (3) the presence or absence of firm co-membership. There is clearly a need for uniformity in further research.

In general terms, the findings show some support for the broad proposition that architectures correspond. But the support is not unqualified. Based upon the extant literature (ie, Colfer, 2007; Colfer and Baldwin, 2010), I propose two different ways in which correspondence can be assessed: (1) a 'standard test' and, (3) a 'strong test'. A standard test of correspondence may seek to examine the relationship between product component, task and firm boundaries, mapped against an organisational tie of either co-location/geographical distance or firm co-membership (or both).

In such a test, the findings provide strong support for correspondence across all time periods. For instance, in many time periods (1) modular product components are associated with task partitioning and outsourcing and, (2) integrated product components are associated with task integration and insourcing. Furthermore, (1) modular product components are associated with geographical distance and an absence of firm co-membership and, (2) integrated product components are associated with co-location and the presence of firm co-membership. These observations highlight congruence with the simple observation that integrated products are developed by integrated firms, and modular products are developed by specialised firms (ie, Colfer, 2007; Fine, 1998).

However, I find a clear exception to the rule. First, closed modularisation clearly breaches the assumption that product component modularity is associated with outsourcing and an absence of firm co-membership (ie, Chesbrough, 2003; Hoetker, 2006). In this research study, I have proposed in section 5.2 that closed modularisation occurs when ex-ante investments in standards precede the emergence of superior productive capabilities in intermediate markets. In other words, assuming a path-dependant and specialised productive capability trajectory, temporal delays in shifting firm boundaries to correspond to product component and task boundaries may be a probable explanation.

A strong test of correspondence seeks to further examine the correspondence of knowledge boundaries and the degree of information-sharing as a measurement proxy. Much of the extant literature assumes that modular product architectures permit vertical⁷⁴ knowledge partitioning and low-levels of ex-post information-sharing (ie Sanchez & Mahoney, 1996). I find the evidence mixed. In general terms, I find support for the notion that modular product architectures permit knowledge to be partitioned at the product component level, ie, product component, task and knowledge boundaries correspond. However, I find little evidence to support the notion of Sanchez & Mahoney (1996) and Sanchez (2003) that architectural knowledge and product component knowledge can be vertically decoupled. Interviews suggest that cycles of modular and architectural innovation necessitate the need for firms to develop systems integration capabilities, resulting in a ‘vertical’ integration of architectural and product component knowledge, and hence high levels of information-sharing throughout the product development process.

⁷⁴ ie architectural and product component knowledge

As discussed in section 5.2, in the UK personal pensions product market, manufacturer/assembler firms, in response to competence-destroying innovations, turned to the development and maintenance of systems integration capabilities to ensure (1) survival, and (2) to attempt to reverse the weak correlation in productive capabilities for outsourced modular product components that offered a form of differentiation or competitive advantage. In other words, I can formalise a further contribution. While I find that the reasoning behind knowledge integration is similar to technological newness (Zirpoli and Becker, 2007), fast rates of product component change (Furlan, *et al.*, 2014), and uneven rates of product component change (Brusoni, *et al.*, 2001), I can extend prior work arguing that knowledge integration, high-levels of information-sharing and co-location are associated with (1) only high value outsourced modular product components and (2) as an option value for future integrative innovation modes. In other words, I can extend the work of Furlan, *et al.*, (2014) on contingencies that ‘mist the mirror’ and argue that the value characteristics of the product component are important in determining knowledge boundaries and hence whether a strong test of correspondence is supported or unsupported.

Many scholars have also highlighted the role of product complexity in determining the degree of architectural correspondence (ie, MacDuffie, 2013), as complexity is often associated with unforeseen product component interdependencies (ie, Pil & Cohen, 2006; Zirpoli and Becker, 2007). Interviews suggest that the degree of product complexity in the UK personal pensions product market increases across time. As a consequence, the findings suggest the degree of product complexity may reach a ‘tipping point’ that necessitates the emergence of systems integration capabilities and knowledge integration. For example, in the early to mid-90s the product architecture was characterised as ‘simple’ and insurance companies relied upon a vertical division of knowledge. However, the architectural innovations in 2000 and 2005 are associated with a step-change increase in product complexity, owing to the multi-technology explosion of complementary product components that could plug and play into the architecture, and then a widening-out of the product layer to a multi-product and multi-technology platform architecture.

In 2000, increases in product complexity are associated with manufacturer-assembler firms maintaining a broader knowledge than task boundary across firm boundaries, and high-levels of information-sharing throughout the product development process. Between 2005 and 2014, the multi-product, multi-technology characteristics of the platform architecture further increased product complexity which, in turn, necessitated even higher levels of knowledge integration such as co-location and employee rotation and higher levels of information-sharing across the entire product system. As such, I can make a third contribution to prior work by arguing that increasing product complexity is associated with higher levels of knowledge integration, information-sharing, and co-location across the entire system. In other words, at least under conditions of a multi-product, multi-technology platform architecture, product complexity ‘mists the mirror’. However, while extant literature assumes that controlling product component interdependencies is the primary logic, in this research setting, incentives for increasing knowledge integration also appear to be associated with maintaining an option value to future integrative innovation in respect of high-value product components.

6. Conclusions and contribution

6.1 Contribution to knowledge

This sub-section provides a summary of the contributions outlined in chapters 2 and 5.

6.1.1 Architectural co-evolution

In respect of architectural co-evolution, I contribute to the extant literature in the following ways:

1. Existing product architecture typologies (ie, Sanchez, 2008; Shibata, *et al.*, 2005) fail to capture the notion of hybridity in product architecture. The stylised product architecture typology developed in chapter 2 encompasses the notion of a hybrid product architecture and embeds a co-evolutionary perspective. While the typology formed the basis of the research methodology, as outlined in chapter 3, I hope future empirical work in the research community may provide opportunities to develop its construct validity and applicability in scholarly work.
2. I show that the co-evolutionary dynamics of productive capabilities and efficiency considerations trade-off at the product component level, revealing either integration gains or gains from specialisation and trade. In other words, on an evolutionary path towards increasing product modularity and industry specialization, I find that exogenous shocks in the industry create the necessary selection forces to compel firms to (1) invest ex-ante in standards, however product component modularity is not a sufficient condition for outsourcing across firm boundaries. Rather, (2) superior productive capabilities must also be available in the intermediate market, (3) net of ex-post transaction costs. Subject to all three conditions being met, a modular product component may subsequently migrate across firm boundaries, providing a correspondence of product component, task and firm boundaries (as discussed in 6.1.2 to follow). Therefore, it is highly probable that the degree of product modularity in a product architecture at any one time is likely to be the aggregate sum of the respective productive capabilities and efficiency trade-offs at the underlying individual product component level. As product architectures are made up of multiple product components, it is probable that product modularity may emerge for some product components and not others, underlining the importance of embedding the notion of hybridity into future product architecture research.

3. In this research context, ‘extreme’ product modularity led to commoditization and imitation risks that ultimately posed limits to gains from specialization and trade which led to integrative innovation modes and industry consolidation. This finding supports the work of Fixson and Park (2008) foregrounding the role of ‘designers seeking value through integrative innovation’, but provides a different perspective to studies by Cacciatori and Jacobides (2005), Christensen, *et al.*, (2002), and Jacobides and Winter (2005). On a path to increasing product and industry reintegration, I extend the work of Fixson and Park (2008) and show how manufacturer/assembler firms initiated three types of integrative innovation modes, (1) under guided modularity, manufacturer/assembler firms bundled combinations of modular product and process components, (2), power blocks emerged consisting of modular product and process components bundled for the specific and exclusive use of particular value chain combinations. Coordinated by the lead manufacturer/assembler firm, the modular product components would often be reengineered within its boundaries to suit the specific performance requirements of the power block, but maintaining adherence to the standards-based platform architecture, and (3) insourcing, via acquisition, the design and development of previously outsourced ‘critical’ product components. I also show how the governance mode differed for each integrative innovation mode. In other words, (1) guided modularity was governed by bi-lateral market contracts, (2) power blocks were often configured into hybrid governance modes such as alliances or joint venture initiatives, and (3) insourcing product component was governed by forward or backward vertical integration.

4. In terms of integrative innovation, Fixson and Park (2008) ask ‘where do the capabilities come from?’ I contribute to our understanding of systems integrator firms. I show how the process of developing and maintaining systems integration capabilities in one time period may be associated with an option value for integrative innovation modes in future time periods. For example, (1), around 2000, interviews suggest that in response to the erosion of productive capabilities in both the core systems architecture and peripheral product components, insurance companies turned to the development of systems integration capabilities and maintained close collaborative relationships with upstream and downstream suppliers as a mechanism to survive. By 2005, (2) as the architectural and product component knowledge of insurance companies grew, and new platform systems architects entered the product market upstream, manufacturer/assembler firms initiated deeper and more intensive knowledge integration mechanisms across the entire value chain encompassing devices such component co-development, temporary co-location, and employee-rotation. By 2012, (3) manufacturer/assembler firms leveraged existing system integration capabilities developed across the time periods to exercise the options to form quasi-integrated power blocks based on hybrid forms of governance and insourced, via acquisition, productive capabilities in component design where the product component offered routes to differentiation or competitive advantage.

6.1.2 Architectural correspondence

I offer the following contributions:

1. I show the extent to which architectural correspondence is sustained in the face of co-evolutionary forces across time. By charting architectural correspondence across a 30-year period, I have shown how product markets may evolve through periods of correspondence and non-correspondence across time. Moreover, I show how correspondence and non-correspondence may occur *at the same time*. This occurs depending upon which boundaries are being examined (ie, product component, task, knowledge and firm boundaries) and which measurement proxy is being used (ie, firm co-membership, geographical location and levels of information-sharing).

2. I develop further contingencies that may ‘mist the mirror’. At the *product architecture level*, I find that closed modularisations provide a contingency that has been previously highlighted in the extant literature (ie, Hoetker, 2006), calling into question the assumption that modular product architectures are always associated with outsourcing across firm boundaries. However, I extend prior work by proposing that closed and modular product architectures emerge under specific conditions when ex-ante investments in standards precede the emergence of superior productive capabilities in intermediate markets. In other words, assuming a path-dependent productive capability trajectory, temporal delays in shifting firm boundaries to correspond to product component and task boundaries may be a probable explanation.

3. *At the product architecture level*, I also highlight the role of product complexity in determining the degree of architectural correspondence. I find congruence with the prior work of MacDuffie (2013) and Zirpoli and Becker (2007) that increasing product complexity increases the need for a non-correspondence of knowledge boundaries, high-levels of information-sharing and co-location, even where the product component is modular. I find that the incentives for knowledge integration in one time period are associated with maintaining an option value to future integrative innovation in respect of high-value product components in a future time period.

4. *At the product component level*, I highlight the importance of the value characteristics of the outsourced modular product component in determining knowledge boundary correspondence or non-correspondence. While I find that the reasoning behind knowledge integration is similar to the ideas of technological newness (Zirpoli and Becker, 2007), fast rates of product component change (Furlan, *et al.*, 2014), and uneven rates of product component change (Brusoni, *et al.*, 2001), I extend prior work by arguing that maintaining a broader knowledge than task boundary, high-levels of information-sharing throughout and co-location are associated with high-value product components. In other words, where an outsourced modular product component is high-value, firms use the ‘complementarity hypothesis’ (Cabigiosu & Camuffo, 2012) to organize its innovative activities. In contrast, where an outsourced modular product component is low value, firms will rely upon the ‘trade-off hypothesis’ (Cabigiosu & Camuffo, 2012), invest ex-ante in standards and then revert to a low-level of ex-post information-sharing and geographic distance to manage product development.

6.2 Contribution to practice and public policy

The ability for managers to understand and predict how a product architecture may evolve towards or away from product modularity should be of great strategic concern, both to existing incumbents and new entrants.

For instance, this research study highlights that where a firm that has chosen to sponsor a closed and integrated product design it may be at risk of losing market share and ultimately risks technological lock-out if it does not adapt to firms 'going modular'. In a similar vein, open and modular product architectures often result in commoditisation and imitation, and therefore the findings of this research study may help strategizing managers decide which aspects of a product architecture to retain control of and which to open-up. It is probable that retaining an intermediate degree of modularity may be the best strategy even in the face of forces pushing towards the extremes (Ethiraj, *et al.*, 2008).

The question arises as to how managers should decide between architectural choices. This is clearly a multi-dimensional issue and almost impossible to answer without detailed costs and benefits compared across a number of alternative architectural choices. However, this research study highlights the need for managers to take a dynamic view and ultimately it may depend perhaps on the prevailing stylised characteristics of the dominant product architecture design (Ethiraj, *et al.*, 2008) at a particular point in time.

For instance, where the prevailing architectural design is closed and integrated, and evolutionary processes are pushing towards a modular design, firms may be better off either retaining a closed and modular or hybrid product architecture in order to minimise the threats from imitation and commoditisation associated with extreme open and modular product architectures. On the other hand, in product markets where the dominant architecture is open and modular, firms may be better off developing or maintaining systems integration capabilities in order to develop superior productive capabilities and an option value for future integrative innovation in a subsequent time period.

As Fixson and Park (2008) acknowledge, changing the rules of the game requires *a priori* investments in maintaining a broad knowledge base, even when the product architecture is modular. Maintaining systems integration capabilities seems to be an important factor that managers should strive to develop. Systems integration capabilities may be essential if firms are to avoid a loss of both architectural control and access to product component level knowledge that may be strategically critical in determining which product components to insource and which to outsource, in other words, how best to partition product component, task, knowledge and firm boundaries. In fact, systems integration capabilities may confer firms a number of strategic advantages, (1) it may help firms avoid limits to gains from trade and/or specialisation in the face of increasing product modularity, (2) systems integration capabilities may allow firms to avoid knowledge decay at both the architectural and product component level, and, (3) it may create an ‘option value’ associated with future integrative innovation and the ability to vary the degree of product architecture modularity.

MacDuffie (2013:37) suggests that managers also often tend to make the mistake of assuming that once product component boundaries have been partitioned, task, knowledge and firm boundaries can be subsequently partitioned so that firms can rely on external suppliers to design modular product components independently. In terms of architectural correspondence, the degree to which a product and task structure is modularised is an important strategic decision, and I have shown that firms should pay careful attention to not only the modular characteristics of each product component, but also the opportunity to create differentiation and competitive advantage. It may then be possible to pursue both a ‘trade-off’ and ‘complementarity’ hypothesis (Cabigiosu & Camuffo, 2012) at the same time.

6.3 Limitations and recommendations for future research

As a single industry case study, I acknowledge that such studies often run the risk of being non-generalisable (ie, Eisenhardt, 1989) to other product market contexts. I also acknowledge that it is also risky to attempt to generalise when co-evolutionary processes are still emergent (especially in the period following 2012). However, like any other contextualist process research, my aim has been to understand the dynamics of the specific product market setting rather than the concerns of generalisability of the research findings to other settings.

As discussed in Chapter 3, I also acknowledge that my positionality may, despite my best efforts, have exerted an undue influence on the research design, methodology and methods. I have outlined what steps I have taken to ensure the credibility and plausibility of the research findings, including the use of an expert panel.

It is also worth noting that the research design was created as a primarily deductive study. Based upon the valuable feedback from examiners, the literature review has been structured to focus on the main contributions of the study in order to improve readability. The author highlights that the main contributions of the study were only known following the initial drafting of the literature review.

6.3.1 Architectural co-evolution

I observe that further empirical research would be beneficial in the following areas:

1. The stylised product architecture presented in chapter 2 would benefit from empirical application in other product market contexts. Moreover, although I have contributed to the extant literature in proposing three different integrative innovation mechanisms, future empirical work may seek to examine different product architecture hybridity types in different settings

2. I have focused on supply-side co-evolution processes. A future synthesis of supply-side and demand-side processes may offer further opportunities to progress the direction of research. It is possible that demand-side factors, such as a buyer's willingness to mix and match product components, may also exert a strong influence on the co-evolutionary path of products and industries
3. Technological platform architectures in service-orientated environments remain a largely underexplored topic (Gawer, 2014). I have taken a small step in this direction, and further empirical research may wish to further develop a theory of how product architectures co-evolve into technological platform architectures. The literatures on product architecture and platforms remain largely separate and distinct and a synthesis of work across the two domains is necessary.

6.3.2 Architectural correspondence

In terms of architectural correspondence, this research study has responded to numerous calls in the extant literature for a longitudinal and retrospective perspective. However, there remain a number of additional gaps in the extant literature.

1. I highlight the role of the value characteristics of product components in determining correspondence, and further research may be directed towards the economics of product modularity (ie, Campagnolo & Camuffo, 2010), specifically examining the precise value characteristics of product components in order to better understand how firms organize and manage different product development tasks.
2. This research study is a single industry case study. Further empirical work may wish to examine cross-industry cases in order to uncover generalizable product market contingencies
3. I have not attended to the performance implications of architectural correspondence and non-correspondence. Under what circumstance is a firm better-off corresponding or non-corresponding architectures? Affuah (2001) initiated work in this area, and further empirical studies remain necessary and relevant

4. The emergent perspective on the reverse mirroring hypothesis (ie, Sanchez, *et al.*, 2013) suggests that products may not design organisations, and a reverse direction of causation may be possible. Although Sako (2003) and Fixson and Park (2008) examine the direction of causality in their work, further empirical is required to unpack the precise paths of architectural correspondence or non-correspondence. This necessarily requires a longitudinal or retrospective perspective.

6.4 Final thoughts and reflections

I finish with a little speculation.

It is too early to say whether architectures will continue to co-evolve towards increasing integration, or evolve back towards further openness and modularity. For example, Langlois (2003) has suggested that temporary shifts towards integration are short punctuations in the broader, macro shift towards modularity and disintegration. Fine (1998) similarly highlights the unstable nature of product architecture.

However, in the interview process many respondents shared a commonly-held view of the direction they believed the product market would take. Many expected significant horizontal consolidation in all value chain segments, especially the advice and manufacturer/assembler segments as smaller firms are either forced to exit or are acquired by larger firms in the pursuit of scale, as predicted by Fixson and Park (2008). Some respondents believed that restricted access to IT product development resources would also force many manufacturer/assembler firms to exit. Commonly, many interviewees felt the product market could only sustain c5-10 industry platforms in the medium-term, and many were already plotting how to maximise value in the context of an exit strategy. Horizontal integration was expected to occur within the product market segment, as well as across product market segments. For example, ‘industry platforms’ also exist in the investment and retirement market such as ‘workplace platforms’, as well as share dealing platforms. Many respondents believed that these different product segment platforms would converge into an industry standard platform architecture across different segments.

Many respondents also believed that further vertical integration would follow, as manufacturers/assemblers acquired smaller specialist product component suppliers which may, of course, reduce the number of product component suppliers. In line with Gawer (2010, 2014), it also appears manufacturers/assembler firms have lost control of the value chain to some degree. For instance, many larger IFA firms are reportedly in the process of bypassing the manufacturer/assembler firm to contract directly with the platform systems providers in order to establish their own product platforms, following the lead of fund management groups in the late-90s and early-noughties. Moreover, back-office suppliers are creating ‘platforms of platforms’ in order to competitively attack and forward integrate into the product platform space. In general terms, it would appear that the product market is in a period of ‘ferment’ as described by Anderson and Tushman (1990).

In a way, it is a shame that this research study was conducted soon after the introduction of RDR and auto-enrolment in 2012. As many industry commentators are predicting ‘the death of platforms as we know them’, it would be interesting and insightful to revisit the product market in a few years to uncover how the dynamics I describe in this research study have continued to evolve.

Looking back at the PhD journey I feel it is appropriate at its closure to reflect on the research journey and its data collection, analysis and writing approach. After leaving the UK financial services sector in 2011 to pursue a career in academia it was very interesting to re-enter the financial world again after a 3-year absence. I had a suspicion that major changes may have taken place, but was nonetheless surprised by how fast the pace of change had progressed, and continues to progress. When I started the PhD, I had a loose set of ideas in mind, but found that researching the relevant literature extremely helpful in refining and revising my research question. In fact, one of the hardest parts of the journey was in deciding on the research question and de-marking the boundary of this research study; there were so many options, it was difficult to focus on one specific domain. Thank-you to the supervision team for assertively providing guidance.

However, the data collection process proved an extremely valuable learning experience having moved on from the literature review to data collection. I was fortunate to secure the time and funding to interview respondents during semester-time, and managed to conduct the thirty-one interviews over an approximate three-month period. I had to learn new skills in respect of interviewing, recording and listening. As a new qualitative researcher, I was also very conscious of my positionality and so I read numerous journal papers and talked to experienced researchers before the interviews in order to try to understand the issues. On the whole, however, I found the data collection process less difficult than anticipated and, in fact, rather enjoyable.

The same cannot be said about data analysis. The data analysis proved very different than anticipated. I had no previous experience in using NVivo and so attended two day-long workshops in order to get a handle on the basics, and how to structure the longitudinal nature of the methodology within the software package. However, the qualitative data collected from the interviews proved a much longer and complex process than expected, producing over half-a-million words. Each interview took at least two days to analyse, consequently delaying the overall progress of the writing up phase of the PhD. In hindsight, it would have been impossible to analyse the data without the use of software (NVivo) to assist in the coding of the data and extracting the relevant themes. As a techno-phobe, my initial plan to use Microsoft Excel would have been a very bad decision. Although I achieved some level of saturation after about twenty-five interviews, I had thirty-one interviews in the diary and so progressed on. However, qualitative data analysis requires substantial time even with the assistance of software (NVivo) and so keeping the number of interviews to a manageable size is highly recommended. Thirty-one interviews took about sixty consecutive days to code and a similar amount to analyse into a hierarchal thematic structure.

With so much data, reporting the study findings became difficult as I would have to ‘leave out’ so much of interest. However, I followed a structured approach by focusing on aspects of the interview data that most closely related to the research questions. The writing-up process was quite difficult. I was fortunate to be awarded with a 14-week sabbatical, without which I would be some time away from completing this research study. As someone who needs large blocks of time to write, I found the sabbatical very helpful in making sustained progress.

To break-up the process, I also developed two conceptual papers based upon the literature review. One paper has been submitted to *Academy of Management Review*, and the other submitted to *Industry and Innovation*. At the time of writing, I await responses. I hope more papers will follow as a direct result of this research study.

Overall, I strongly believe that this research study offers a valuable insight into architectural co-evolution and correspondence in the UK personal pensions product market. With data that would have been almost impossible for other academic researchers to access, this research study provides a comprehensive account of the product market and how its product and industry architecture has evolved. As a consequence, I hope that the findings presented are of benefit for both fellow researchers and management practitioners. As such, my medium-term objective as a new researcher is to bring these findings into the public domain, with further papers to be submitted for publication.

REFERENCES

- Abermathy, W. & Utterback, J. (1978). Patterns of industrial innovation. *Technology Review*, 50 (7), 40-47.
- Acker, S. (2000). In/out/side: positioning I in feminist qualitative research. *Resources for Feminist Research* 28 (1/2), 189-208.
- Adner, R. & Levinthal, D. (2001). Demand heterogeneity and technology evolution: implications for product and process innovation. *Management science*, 47(5), 611-628.
- Afuah, A. (2001). Dynamic boundaries of the firm: are firms better off being vertically integrated in the face of a technological change? *Academy of Management Journal*, 44(8), 1211-1228.
- Alexander, C. (1964). *Notes on the synthesis of form*. Cambridge: Harvard University Press.
- Argyres, N. (1999). The impact of information technology on coordination: evidence from the B-2 stealth bomber. *Organisation Science*, 10(2), 162-180.
- Argyres, N., & Bigelow, L. (2007). Does transaction misalignment matter for firm survival at all stages of the industry life cycle?. *Management Science*, 53(8), 1332-1344.
- Argyres, N. & Bigelow, L. (2010). Innovation, modularity and vertical disintegration: evidence from the early US motor industry. *Organisation Science*, 21(4), 842-853.
- Argyres, N. & Liebeskind, J. (1999). Contractual commitments, bargaining power and governance inseparability: incorporating history into transaction cost theory. *Academy of Management Review*, 24(1), 49-63.
- Anderson, P. & Tushman, M. (1990). Technological discontinuities and dominant designs: a cyclical model of technological change. *Administrative science quarterly*, 35(4), 604-633.

- Arrow, K. (1974). *The limits of organisation*. New York: Norton.
- Asselin, M. (2003). Insider research: issues to consider when doing qualitative research in your own setting. *Journal for Nurses in Staff Development*, 19(2), 99-103.
- Bach, N. & Galvin, P. (2011). *Why do firms set up permeable firm boundaries? A theoretical analysis integrating efficiency and capability perspectives*. Paper presented at EURAM 2011: 11th European Academy of Management Conference, June 2011, Tallin, Estonia.
- Baldwin, C. (2008). Where do transactions come from? modularity, transactions and the boundaries of firms. *Industrial and Corporate Change*, 17(1), 155-195.
- Baldwin, C. (2015). *Bottlenecks, Modules and Dynamic Architectural Capabilities*. Harvard Business School Working Paper, No. 15-028, October 2014. (Revised May 2015).
- Baldwin, C. & Clark, K. (2000). *Design rules: the power of modularity*. Cambridge, MA: MIT Press
- Baldwin, C. & Clark, K. (2003). Managing in an age of modularity. In R. Garud, A. Kumarswamy, R. Langlois (Eds), *Managing in the modular age: architectures, networks, and organizations*. (pp149-171) Oxford: Blackwell Publishing.
- Baldwin, C. & Woodard, C. (2008). *The architecture of platforms: a unified view*. Harvard Business School Finance Working Paper, (09-034). September 2008.
- Banfield, G. (2004). What's really wrong with ethnography? *International Education Journal*, 4 (4), 53-63.
- Barney, J. (1991). Firm resources and sustained competitive advantage. *Journal of Management*, 17(1), 99-120.
- Blakie, N. (1993). *Approaches to Social Enquiry*. Cambridge: Polity Press.

- Boisot, M. & Sanchez, R. (2010). Organisation as a nexus of rules: emergence in the evolution of systems and change. *The International Review of Management Studies*, 21(4), 378-405.
- Boudreau, K. (2010). Open platform strategies and innovation: granting access vs. devolving control. *Management Science*, 56(10), 1849-1872.
- Boudreau, K. (2012). Let a thousand flowers bloom? An early look at large numbers of software app developers and patterns of innovation. *Organization Science*, 23(5), 1409-1427.
- Boudreau, K. & Hagiu, A. (2009). Platform rules: multi-sided platforms as regulators. In A, Gawer (Ed), *Platforms, markets and innovation*. (pp163-191). London: Edward Elgar.
- Brannen, J. & Nilsen, A. (2011). Comparative biographies in case-based cross-national research: methodological considerations, *Sociology*, 45 (4), 603-618.
- Braun, V. & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77-101.
- Breen, L. (2007). I in the middle: negotiating the insider/outsider dichotomy. *The Australian Community Psychologist*, 19(1), 163-174.
- Brooks, J. & King, N (2012). *Qualitative psychology in the real world: the utility of template analysis*. Paper presented at the 2012 British Psychological Society Annual Conference, 18-20 April 2012, London, UK (Unpublished).
- Brooks, J. & King, N. (2014). *Doing template analysis: evaluating an end of life care service*. Sage Research Methods Cases. This version is available at <http://eprints.hud.ac.uk/19707/>.
- Brusoni, S. & Prencipe, A. (2001). Managing knowledge in loosely coupled networks: exploring the links between product and knowledge dynamics. *Journal of Management Studies*, 38(7), 1019-1035.

- Brusoni, S. & Prencipe, A. (2006). Making design rules: A multi-domain perspective. *Organization science*, 17(2), 179-189.
- Brusoni, S. & Prencipe, A. (2011). Patterns of modularisation: the dynamics of product architecture in complex systems. *European Management Review*, Vol8, 67-80.
- Brusoni, S., Prencipe, A. & Pavitt, K. (2001). Knowledge specialisation, organisational coupling, and the boundaries of the firm: why do firms know more than they make? *Administrative Science Quarterly*, 46(4), 597-621.
- Cabigiosu, A. & Camuffo, A. (2012). Beyond the mirroring hypothesis: product modularity and inter-organisational relations in the air conditioning industry. *Organisation Science*, 23(3), 686-703.
- Cabigioso, A., Zirpoli, F. & Camuffo, A. (2013). Modularity, interfaces definition and the integration of external sources of innovation in the automotive industry. *Research Policy*, 42(3), 662-675.
- Cacciatori, E. & Jacobides, M. (2005). The dynamic limits of specialisation: vertical integration reconsidered. *Organisation Studies*, 26(12), 1851-1883.
- Campagnolo, D. & Camuffo, A. (2010). The concept of modularity in management studies: a literature review. *International Journal of Management Reviews*, 12(3), 259-283.
- Campbell, D. (1969). Variation and selective retention in socio-cultural evolution. *General Systems*, 14, 69-85.
- Carter, R. & Hodgson, G. (2006). The impact of empirical tests of transaction cost economics on the debate on the nature of the firm. *Strategic Management Journal*, 27(5), 461-476.
- Cebon, P., Hauptman, O., & Shekhar, C. (2008). Product modularity and the product life cycle: new dynamics in the interactions of product and process technologies. *International Journal of Technology Management*, 42(4), 365-386.

Ceci, F., Masini, A., & Prencipe, A. (2014). *The limits of systems integration: complementarity, contingencies, and solution design choices*. HEC Paris Research Paper No. MOSI-2014-1069. This paper can be downloaded from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2535356

Chandler, A. (1962). *Strategy and structure: chapters in the history of American industrial enterprise*. Cambridge, MA: MIT Press.

Chavez, C. (2008). Conceptualising from the inside: advantages, complications and demands on insider personality. *The Qualitative Report*, 13(3), 474-494.

Chesbrough, H. (2003). Towards a dynamic of modularity: a cyclical model of technical advance. In A. Prencipe, A. Davies, & M. Hobday (Eds), *The business of systems integration*. (pp174-200). Oxford: Oxford University Press.

Chesbrough, H., & Kusunoki, K. (2001). The modularity trap: innovation, technology phase shifts, and the resulting limits of virtual organizations. In I. Nonaka and D. Teece (Eds), *Managing industrial knowledge: creation, transfer, and utilization*. (pp220-230). London: Sage.

Chesbrough, H. & Prencipe, A. (2008). Networks of innovation and modularity: a dynamic perspective. *International Journal of Technology Management*, 42(4), 414-425.

Chesbrough, H. & Teece, D. (1996). When is virtual virtuous. *Harvard Business Review*, 74(1), 65-73.

Christensen, C. & Rosenbloom, R. (1995). Explaining the attacker's advantage: technological paradigms, organisational dynamics, and the value network. *Research Policy*, 24(2), 233-257.

- Christensen, C, Verlinden, M., & Westerman, G. (2002). Disruption, disintegration, and the dissipation of differentiability. *Industrial and Corporate Change*, 11(5), 955-993
- Chuma, H. (2006). Increasing complexity and limits of organization in the microlithography industry: implications for science-based industries. *Research Policy* 35(3), 394–411
- Ciarli, T., Leoncini, R., Montresor, S., & Valente, M. (2007). Innovation and competition in complex environments. *Innovation*, 9(3-4), 292-310.
- Clark, K. (1985). The interaction of design hierarchies and market concepts in technological evolution. *Research Policy*, 14(5), 235–251.
- Coase, R. (1937). The nature of the firm. *Economica*, 4(16), 386-405.
- Colfer, L. (2007). *The mirroring hypothesis: theory and evidence on the correspondence between the structure of products and organizations*. Unpublished manuscript, Harvard Business School, Boston, MA.
- Colfer, L. & Baldwin, C. (2010). *The mirroring hypothesis: theory, evidence and exceptions*. Working Paper, 10-058, Harvard Business School, Boston, MA.
- Conner, K., & Prahalad, C. (1996). A resource-based theory of the firm: knowledge versus opportunism. *Organization science*, 7(5), 477-501.
- Conway, M. (1968). How do committees invent. *Datamation*, 14(4), 28-31.
- Corbin, J. & Strauss, A. (2008). *Basics of qualitative research: techniques and procedures for developing grounded theory*. California: Sage.
- Cox, J. & Hassard, J. (2007). Ties to the past in organisation research: a comparative analysis of retrospective methods. *Organization*, 14(4), 475-497.

- Crabtree, B. & Miller W. (1992). A template approach to text analysis: developing and using codebooks. In B. Crabtree. & W. Miller (Eds), *Doing qualitative research*. (pp163-178). London: Sage.
- Crotty, M. (1998). *The foundations of social research*. London: Sage.
- D'Adderio, L. & Pollock, N. (2014). Performing modularity: competing rules, performative struggles and the effect of organizational theories on the organization. *Organization Studies*, 35(12), 1813-1843.
- DeLyser, D. (2001). Do you really live here? thoughts on insider research. *Geographical Review*, 91(1/2), 441-453.
- De Meuse, K., Bergmann, P., & Lester, S. (2001). An investigation of the relational component of the psychological contract across time, generation and employment status. *Journal of Managerial Issues*, 13, 102-118.
- Denzin, N, & Lincoln, Y. (2000). Introduction: the discipline and practice of qualitative research. In N. Denzin & Y. Lincoln (Eds). *Handbook of qualitative research*. (pp1-28). Thousand Oaks, CA: Sage.
- Dosi, G. (1982). Technological paradigms and technological trajectories, *Research Policy*, 22(2), 147-162.
- Duncan, E. & Nicol, M. (2004). Subtle realism and occupational therapy: an alternative approach to knowledge generation and evaluation. *British Journal of Occupational Therapy*, 67(10), 453-456.
- Dwyer, S. & Buckle, J. (2009). The space between: on being an insider-outsider in qualitative research. *International Journal of Qualitative Methods*, 8(1), 54-63.
- Easterby-Smith, M., Thorpe, R., & Jackson, P. (2012). *Management research*, Fourth Edition, London: Sage.

- Easton, G. (2010). Critical realism in case study research. *Industrial Marketing Management*, 39(1), 118-128.
- Eisenhardt, K. (1989). Building theories from case study research. *Academy of Management Review*, 14(4), 532-550.
- Eisenmann, T, Parker, P, & Van Alstyne, M. (2008). *Opening platforms: how, when and why?* Harvard Business School Working Paper 09-030. September 2008
- Emerson, R. (1983). *Contemporary field research: a collection of readings*. Prospect Heights, Ill: Waveland Press.
- Ethiraj, S. & Levinthal, D. (2004). Modularity and innovation in complex systems. *Management Science*, 50(2), 159-173.
- Ethiraj, S., Levinthal, D., & Roy, R. (2008). The dual role of modularity: innovation and imitation. *Management Science*, 54(5), 939-955.
- Fereday, J. & Muir-Cochrane, E. (2006). Demonstrating rigour using thematic analysis: a hybrid approach of inductive and deductive coding and theme development. *International Journal of Qualitative Methods*, 5 (1), 80-92.
- Fine, C. (1998). Clockspeed-based strategies for supply chain design. *Production and Operations Management*, 9(3), 213-221.
- Fine, C., Golany, B. & Naseraldin, H. (2005). Modelling trade-offs in three-dimensional concurrent engineering: a goal programming approach. *Journal of Operations Management*, 23(3), 389-403.
- Fine, C. & Whitney, D. (1996). *Is the Make versus Buy Decision a Core Competence?*, MIT International Motor Vehicle Program Working Paper. Cambridge, MA.

- Fixson, S. (2003). The multiple faces of modularity – a literature analysis of a product concept for assembled hardware products, Technical Report, 03-05. University of Michigan. June 2003.
- Fixson, S. (2005). Product architecture assessment: a tool to link product, process and supply chain decisions. *Journal of Operations Management*, 23(3), 345-369.
- Fixson, S. & Park, J. (2008). The power of integrality: linkages between product architecture, innovation, and industry structure. *Research Policy*, 37(8), 1296-1316.
- Fleming, L. & Sorenson, O. (2001). The dangers of modularity. *Harvard Business Review*, 79(8), 20-21.
- Funk, J. (2008). Systems, components and modular design: the case of the US semiconductor industry. *International Journal of Technology Management*, 42(4), 387-413.
- Furlan, A., Cabigiosu, A. & Camuffo, A. (2014). When the mirror get misted up: modularity and technological change. *Strategic Management Journal*, 35(6), 789-807.
- Galbraith, J. (1973). *Designing Complex Organizations*. Reading: Addison-Wesley.
- Galbraith, J. (1977). *Organization design: an information processing view*. Reading: Addison-Wesley.
- Galvin, P. & Morkel, A. (2001). The effect of product modularity on industry structure: the case of the world bicycle industry. *Industry and Innovation*, 8(1), 31-48.
- Galvin, P. & Rice, J. (2008). Managing knowledge in the mobile telephone industry: a case study of knowledge protection and diffusion for innovation. *International Journal of Technology Management*, 42(4), 426-438.
- Gardner, G. (2001). Unreliable memories and other contingencies: problems with biographical knowledge. *Qualitative Research*, 1(2), 185-204.

Garud, R. & Kumaraswamy, A. (1995). Technological and organizational designs for realizing economies of substitution. *Strategic Management Journal*, 16(S1), 93-109.

Garud, R., Jain, S., & Kumaraswamy, A. (2002). Institutional entrepreneurship in the sponsorship of common technological standards: the case of Sun Microsystems and Java. *Academy of management journal*, 45(1), 196-214.

Glaser, B. & Strauss, A. (1967). *The discovery of grounded theory: strategies for qualitative research*. New York: Aldine.

Gomes, P. & Joglekar, N. (2008). Linking modularity with problem solving and coordination efforts. *Managerial and Decision Economics*, 29(5), 443-457.

Guba, E. & Lincoln, Y. (1982). Epistemological and methodological bases of naturalistic inquiry. *Educational Communication and Technology Journal*, 30(4), 233-252.

Johnson, P., Buehring, A., Cassell, C., & Symon, G. (2006). Evaluating qualitative management research: towards a contingent criteriology. *International Journal of Management Reviews*, 8 (3), 131-156.

Hammersley, M. (1992). *What's wrong with ethnography?* New York: Routledge.

Hammersley, M. (1993). *What is social research?* Milton Keynes: Open University Press.

Hammersley, M. (2007). The issue of quality in qualitative research. *International Journal of Management Research and Method in Education*, 30 (3), 287-305.

Hammersley, M. (2008). Troubles with triangulation. In M. Bergman (Ed), *Advances in Mixed Methods Research*. (pp22-36). London: Sage.

Hammersley, M. & Atkinson, P. (1995). *Ethnography: principles in practice*. London: Routledge.

- Hao, B., Feng, Y., & Frigant, V. (2015). Rethinking the 'mirroring' hypothesis: implications for technological modularity, tacit coordination, and radical innovation. *R&D Management*. doi: 10.1111/radm.12106
- Henderson, R. & Clark, K. (1990). Architectural innovation: the reconfiguration of existing product technologies and the failure of established firms. *Administrative Science Quarterly*, 35(1), 9-30.
- Herod, A. (1999). Reflections on interviewing foreign elites: praxis, positionality, validity and the cult of the insider. *Geoforum*, 30(4), 313-327.
- Hobday, M., Davies, A., & Prencipe, A. (2005). Systems integration: a core capability of the modern corporation. *Industrial and corporate change*, 14(6), 1109-1143.
- Hoetker, G. (2006). Do modular products lead to modular organisations? *Strategic Management Journal*, 27(6), 501-518.
- Jacobides, M. (2005). Industry change through vertical disintegration: how and why markets emerged in mortgage banking. *Academy of Management Journal*, 48(3), 465-498.
- Jacobides, M. (2008). How capability differences, transaction costs, and learning curves interact to shape vertical scope. *Organisation Science*, 19(2), 306-326.
- Jacobides, M. & Billinger, S. (2006). Designing the boundaries of the firm: from make, buy or ally to the dynamic benefits of vertical architecture. *Organisation Science*, 17(2), 249-261.
- Jacobides, M. & Hitt, L. (2005). Losing sight of the forest for the trees: productive capabilities and gains from trade as drivers of vertical scope. *Strategic Management Journal*, 26(13), 1209-1227.

Jacobides, M, Knudesen, T & Augier, M. (2006). Benefitting from innovation: value creation, value appropriation, and the role of industry architectures. *Research Policy*, 35(8), 1200-1221.

Jacobides, M. & Winter, S. (2005). The co-evolution of capabilities and transaction costs: explaining the institutional structure of production. *Strategic Management Journal*, 26(5), 395-413.

Kanuha, V. (2000). Being native versus going native: conducting social work research as an insider. *Social Work*, 45(5), 439-447.

Katz, M, & Shapiro, C. (1994). Systems competition and network effects. *The Journal of Economic Perspectives*, 8(2), 93-115.

Kerswetter, K. (2012). Insider, outsider, or somewhere in-between: the impact of researchers' identities on the community-based research process. *Journal of Rural Social Sciences*, 27(2), 99-117.

King, N. (1998). Template analysis. In G. Symon. & C. Cassell, (Eds), *Qualitative methods and analysis in organisational research*. London: Sage.

King, N. (2004). Using templates in the thematic analysis of texts. In G. Symon, & C. Cassell. (Eds), *Qualitative methods and analysis in organisational research*. London: Sage.

King, N. (2012). Doing Template Analysis. In G. Symon. & C. Cassell. (Eds). *Qualitative organisational research: core methods and current challenges*. London: Sage.

King, N. & Horrocks, C. (2010). *Interviews in Qualitative Research*. London: Sage.

King, N., Mirza, M., & Bates, E. (2007). *Paddock pathways to health allotment project: qualitative evaluation report*. Centre for Applied Psychological Research, School of Human and Health Sciences, University of Huddersfield.

- Kirk, J., & Miller, M. (1986). *Reliability and validity in qualitative research*. Newbury Park, CA: Sage.
- Klein, B., Crawford, R., & Alchian, A. (1978). Vertical integration: appropriable rents and the competitive contracting process. *Journal of Law and Economics*, 21(2), 297-326.
- Kogut, B. & Zander, U. (1992). Knowledge of the firm, combinative capabilities, and the replication of technology. *Organization science*, 3(3), 383-397.
- Koriat, K. & Goldsmith, M. (1994). Memory in naturalistic and laboratory contexts: distinguishing the accuracy-orientated and quantity-orientated approaches to memory assessment. *Journal of Experimental Psychology: General*, 123(3), 297-315.
- Langlois, R. (2002). Modularity in technology and organization. *Journal of economic behavior & organization*, 49(1), 19-37.
- Langlois, R. (2003). The vanishing hand: the changing dynamics of industrial capitalism. *Industrial and Corporate Change*, 12(2), 351-385.
- Langlois, R. (2006). The secret life of mundane transaction costs. *Organization Studies*, 27(9), 1389-1410.
- Langlois, R. (2012). Design, institutions and the evolution of platforms, *Journal of Law, Economics and Policy*, 9(1), 1-14.
- Langlois, R. & Robertson, P. (1989). Explaining vertical integration: lessons from the American automobile industry. *The Journal of Economic History*, 49(02), 361-375.
- Langlois, R. & Robertson, P. (1992). Networks and innovation in a modular system: lessons from the microcomputer and stereo component industries. *Research Policy*, 21(4), 297-313.
- Langlois, R. & Robertson, P. (1995). *Firms, markets and economic change: a dynamic theory of business institutions*. London: Routledge Press.

- Langdrige, D. (2004). *Introduction to research methods and data analysis in psychology*. Harlow: Pearson.
- Lawrence, P. & Lorsch, J. (1967). Differentiation and integration in complex organisations, *Administrative Science Quarterly*, 12(1), 1-47.
- Lehnerd, A. (1987). Revitalising the manufacture and design of mature global products. In B. Guile and H. Brooks (Eds). *Technology and global industry: companies and nations in the world economy*. (pp49-64). Washington DC: National Academy Press.
- Leiblein, M. (2003). The choice of organisational governance form and performance: predictions from transaction cost, resource-based and real options theories. *Journal of Management*, 29(6), 937-961.
- Leiblein M. & Miller D. (2003). An empirical examination of transaction and firm level influences on the vertical boundaries of the firm. *Strategic Management Journal*, 24(9), 839-860.
- Lewis, K. (2015). *Engaging children and young people in physical activity*. Unpublished Doctoral Thesis, University of Huddersfield.
- Lincoln, Y. & Guba, E. (1985). *Naturalistic Enquiry*. California. Sage.
- Lippman, S. & Rumelt, R. (1982). Uncertain imitability: an analysis of interfirm differences in efficiency under competition. *The Bell Journal of Economics*, 13(2), 418-438.
- Luo, J., Whitney, D., Baldwin, C., & Magee, C. (2011). *How firm strategies influence the architecture of transaction networks*. Harvard Business School Working Paper 11-076. January 2011.
- MacCormack, A., Baldwin, C., & Rusnak, J. (2012). Exploring the duality between product and organizational architectures: a test of the “mirroring” hypothesis. *Research Policy*, 41(8), 1309-1324.

MacDuffie, J. (2006). *Modularity and the automobile: what happened when the concept hit the road*. IMVP Working Paper, June 2006.

MacDuffie, J. (2008). *Technological and organizational barriers to modularity: persistent integrality in the global automotive industry*. The Wharton School, University of Pennsylvania.

MacDuffie, J. (2013). Modularity-as-property, modularization-as-process, and modularity-as-frame: lessons from product architecture initiatives in the global automotive industry. *Global Strategy Journal*, 3(1), 8-40.

Madill, A., Jordan, A., & Shirley, C. (2000). Objectivity and reliability in qualitative analysis: realist, contextualist and radical constructionist epistemologies. *British Journal of Psychology*, 91(1), 1-20.

March, J. & Simon, H. (1958). *Organizations*. Oxford: Wiley.

Marshall, C. & Rossman, G. (2006). *Designing qualitative research*. London: Sage.

Masten, D. (1991). A legal basis for the firm. In O. Williamson & S. Winter (Eds). *The nature of the firm: origins, evolution and development*. (pp196-212). Oxford: Oxford University Press.

Merton, R. (1946). The focused interview. *American Journal of Sociology*, 51, 541-557.

Merton, R. (1972). Insiders and outsiders: a chapter in the sociology of knowledge. *American Journal of Sociology*, 78, 9-47.

Mikkola, J. (2003). Managing modularity of product architectures: toward an integrated theory. *Engineering Management, IEEE Transactions*, 50(2), 204-218.

Mikkola, J. (2006). Capturing the degree of modularity embedded in product architectures. *Journal of Product Innovation Management*, 23(2), 128-146.

- Miles, M. & Huberman, M. (1994). *Qualitative data analysis: an expanded sourcebook*. London: Sage.
- Miller, C., Cardinal, L., & Glick, W. (1997). Retrospective reports in organisational research: a re-examination of recent evidence. *Academy of Management Journal*, 40(1), 189-204.
- Mitchell, W. & Parmigiani, A. (2009). Complementarity, capabilities, and the boundaries of the firm: the impact of within-firm and interfirm expertise on concurrent sourcing of complementary components. *Strategic Management Journal*, 30(10), 1065-1091
- Monteverde, K. (1995). Technical dialog as an incentive for vertical integration in the semiconductor industry. *Management Science*, 41(10), 1624-1638.
- Moss, P. (1995). Reflections on the gap as part of the politics of research design. *Antipode*, 27(1), 82-90.
- Murmann, J. & Frenken, K. (2006). Toward a systematic framework for research on dominant designs, technological innovations, and industrial change. *Research Policy*, 35(7), 925-952.
- Murphy, E., Dingwall, R., Greatbach, C., Parker, S., & Watson, P. (1998). Qualitative research methods in health technology assessment: a review of the literature. *Health Technology Assessment*, 2 (16), 1-276.
- Nelson, R. & Winter, S. (1977). In search of useful theory of innovation. *Research Policy*, 6(1), 36-76.
- Nelson, R. & Winter, S. (1982). The Schumpeterian trade-off revisited. *The American Economic Review*, 72(1), 114-132.
- Nickerson, J., & Zenger, T. (2004). A knowledge-based theory of the firm — the problem-solving perspective. *Organization science*, 15(6), 617-632.

- Nisbett, R. & Ross, L. (1980). *Human interference: strategies and shortcomings of social judgement*. New Jersey: Prentice Hall.
- North, D. (1990). A transaction cost theory of politics. *Journal of theoretical politics*, 2(4), 355-367.
- Orton, J. & Weick, K. (1990). Loosely coupled systems: a reconceptualization. *Academy of Management Review*, 15(2), 203-223.
- Office of National Statistics (2012). UK Pensions Landscape, URL available www.ons.gov.uk/ons/dcp171766_269503.pdf
- Park, J. & Ro, Y. (2013). Product architectures and sourcing decisions: their impact on performance. *Journal of Management*, 39(3), 814-846.
- Parmigiani, A. & Mitchell, W. (2010). The hollow corporation revisited: can governance mechanisms substitute for technical expertise in managing buyer-supplier relationships? *European Management Review*, 7, 46-70.
- Parnas, D. (1971). Information distribution aspects of design methodology. *Proceedings of IFIP Congress*, Ljubljana, Yugoslavia, Volume 1, 339-344
- Parnas, D. (1972). On the criteria to be used in decomposing systems into modules, *Communications of the ACM*, 15(12), 1053-1058.
- Paton, M. (2002). *Qualitative research methods and evaluation methods*, Third Edition. California: Sage.
- Penrose, E. (1959). *The Theory of the Growth of the Firm*, Oxford: Oxford University Press.
- Pil, F. & Cohen, S. (2006). Modularity: implications for imitation, innovation, and sustained advantage. *Academy of Management Review*, 31(4), 995-1011.
- Pisano, G. & Teece, D. (2007). How to capture value from innovation: shaping intellectual property and industry architecture. *California Management Review*, 50(1), 278-296.

- Poppo, L. & Zenger, T. (1998). Testing alternative theories of the firm: transaction cost, knowledge-based, and measurement explanations for make-or-buy decisions in information services. *Strategic management journal*, 19(9), 853-877.
- Puranam, P. & Jacobides, M. (2006). *The dynamics of coordination regimes: implications for organizational design*. London Business School discussion paper (April).
- Reinstaller, A. (2011). The modularity of technology and organisations: implications for the theory of the firm, WIFO working papers, 398/2011.
- Rivkin, J. & Siggelkow, N. (2006). Organizing to strategize in the face of interactions: preventing premature lock-in. *Long Range Planning*, 39(6), 591-614.
- Rochet, J. & Tirole, J. (2003). Platform competition in two-sided markets. *Journal of the European Economic Association*, Vol1, 990-1029.
- Sako, M. (2003). Modularity and outsourcing: the nature of co-evolution of product architecture and organisation architecture in the global automotive industry. In A. Prencipe, A. Davies, & M. Hobday (Eds), *The business of systems integration*. (pp229-253). Oxford: Oxford University Press.
- Sanchez, R. (1995). Strategic flexibility in product creation. *Strategic Management Journal*, 16 (summer), 135-159.
- Sanchez, R. (2003). *Knowledge management and organizational competence*. Oxford: Oxford University Press.
- Sanchez, R. (2008). Modularity in the mediation of market and technology change. *International Journal of Technology Management*, 42 (4), 331-364.
- Sanchez, R. & Collins, R. (2001). Competing – and learning – in modular markets, *Long Range Planning*, 34(6), 645-667.

- Sanchez, R., Galvin, P., & Bach, N. (2013). *'Closing the Loop' in an architectural perspective on strategic organizing: towards a reverse mirroring mypothesis*, Frederiksberg: Department of Innovation and Organizational Economics, Copenhagen Business School.
- Sanchez, R. & Mahoney, J. (1996). Modularity, flexibility and knowledge management in product and organisation design. *Strategic Management Journal*, 17, Special Issue, 63-76.
- Sanchez, R. & Mahoney, J. (2013). Modularity and economic organization: concepts, theory, observations, and predictions. In A. Grandori (Ed). *Handbook of economic organization: integrating economic and organization theory*. (pp383-399), Cheltenham, England: Edward Elgar Publishing.
- Schilling, M. (1998). Technological lockout: an integrative model of the economic and strategic factors driving technology success and failure. *Academy of Management Review*, 23(2), 267-284.
- Schilling, M. (1999). Winning the standards race: building installed base and the availability of complementary goods. *European Management Journal*, 17(3), 265-274.
- Schilling, M. (2000). Towards a general modular systems theory and its application to inter-firm product modularity. *Academy of Management Review*, 25(2), 312-334.
- Schilling, M. & Steensma, K. (2001). The use of modular organisational forms: an industry-level analysis. *Academy of Management Journal*, 44(6), 1149-1168.
- Schilling, M. & Steensma, H., (2002). Disentangling the theories of firm boundaries: a path model and empirical test. *Organization Science*, 13(4), 387-401.
- Seale, C. (1999). Quality in qualitative research. *Qualitative Enquiry*, 5, 465-478
- Seale, C. (2002). Quality issues in qualitative inquiry. *Qualitative Social Work*, 1(1), 97-110.
- Shibata, T., Yano, M., & Kodama, F. (2005). Empirical analysis of evolution of product architecture: Fanuc numerical controllers from 1962 to 1997. *Research Policy*, 34(1), 13-31.

- Siggelkow, N. & Levinthal, D. (2003). Temporarily divide to conquer: centralized, decentralized, and reintegrated organizational approaches to exploration and adaptation. *Organization Science*, 14(6), 650-669.
- Siggelkow, N. & Levinthal, D. (2005). Escaping real (non-benign) competency traps: linking the dynamics of organizational structure to the dynamics of search. *Strategic organization*, 3(1), 85-115.
- Simon, H. (1962). The architecture of complexity. *Proceedings of the American Philosophical Society*, 106(6), 468-482.
- Simon, H. (1982). *Models of bounded rationality: behavioural economics and business organisation*. Cambridge, MA: MIT Press.
- Snelgrove, R. & Havitz, M. (2010). Looking back in time: the pitfalls and potential of retrospective methods in leisure studies. *Leisure Studies*, 32(4), 337-351.
- Sosa, M., Eppinger, S., & Rowles, C. (2004). The misalignment of product architecture and organisational structure in complex product development. *Management Science*, 50(12), 1674-1689.
- Srikanth, K. & Puranam, P. (2007). *Coordination in distributed organizations*. Working paper, available at SSRN: <http://ssrn.com/abstract=939786>
- Srikanth, K. & Puranam, P. (2008). *Co-presence not communication: how ICT's really enable software services offshoring*. Working paper, available at <http://ssrn.com/abstract=928072>
- Srikanth, K. & Puranam, P. (2011). Integrating distributed work: comparing task design, communication, and tacit coordination mechanisms. *Strategic Management Journal*, 32(8), 849-875.
- Steward, D. (1981). *Systems analysis and management: structure, strategy and design*. New York: Petrocelli Books.

- Stigler, G. (1951). The division of labor is limited by the extent of the market. *The journal of political economy*, 59(3), 185-193.
- Staudenmayer, N., Tripsas, M., & Tucci, C. (2005). Interfirm modularity and its implications for product development. *Journal of Product Innovation Management*, 22(4), 303-321.
- Suarez, F. (2004). Battles for technological dominance: an integrative framework. *Research Policy*, 33(2), 271-286.
- Takeishi, A. (2002). Knowledge partitioning in the interfirm division of labor: the case of automotive product development. *Organization Science*, 13(3), 321-338.
- Takeishi, A., & Fujimoto, T. (2003). Modularization in the car industry: interlinked multiple hierarchies of product, production, and supplier systems. In A. Prencipe, A. Davies, & M. Hoibday (Eds). *The business of systems integration*. (pp254-278). Oxford: Oxford University Press.
- Taylor, J. (2011). The intimate insider: negotiating the ethics of friendship when doing insider research. *Qualitative Research*, 11(1), 3-22.
- Tee, R. (2011). *Modularity, architecture and the role of knowledge: a multi-level exploration*. Paper presented at the DRUID conference at Copenhagen Business School, June 2011. Copenhagen, Denmark.
- Tee, R. & Gawer, A. (2009). Industry architecture as a determinant of successful platform strategies: a case study of the i-mode mobile Internet service. *European Management Review*, 6(4), 217-232.
- Teece, D. (1986). Profiting from technological innovation: implications for integration, collaboration, licensing and public policy. *Research policy*, 15(6), 285-305.
- Teece, D. (1996). Firm organization, industrial structure, and technological innovation. *Journal of Economic Behavior & Organization*, 31(2), 193-224.

- Ulrich, K. (1995). The role of product architecture in the manufacturing firm. *Research Policy*, 24(3), 419-444.
- Heutgen, van, K. (2004). Managing insider research: learning from experience. *Qualitative Social Work*, 3(2), 203-219.
- Von Hippel, E. (1990). Task partitioning: an innovation process variable. *Research Policy*, 19(5), 407-418.
- Waring, T. & Wainwright, D. (2008). Issues and challenges in the use of template analysis: two comparative case studies from the field. *The Electronic Journal of Business Research Methods*, 6 (1), 85-94.
- Weick, K. (1976). Educational organizations as loosely coupled systems. *Administrative Science Quarterly*, Vol21, 1-19.
- Wernerfelt, B. (1984). A resource-based view of the firm. *Strategic management journal*, 5(2), 171-180.
- West, J. (2003). How open is open enough?: melding proprietary and open source platform strategies, *Research Policy*, 32(7), 1259-1285 (downloaded from San Jose State University sjsu.edu)
- West, J. (2007). The economic realities of open standards: black, white and many shades of grey. In S. Greenstein. & V. Stango. (Eds), *Standards and Public Policy*. (pp87-122). Cambridge: Cambridge University Press.
- Wheelwright, S. & Clark, K. (1992). *Revolutionizing product development: quantum leaps in speed, efficiency, and quality*. New York: Free Press.
- Williamson, O. (1971). The vertical integration of production: market failure considerations. *American Economic Review*, 61(2), 112-123.
- Williamson O. (1975). *Markets and Hierarchies: analysis and antitrust implications*. New York: Free Press.

Williamson, O. (1979). Transaction cost economics: the governance of contractual relations, *Journal of Law and Economics*, 22(2), 233-261.

Williamson, O. (1985). *The economic institutions of capitalism*. New York: Free Press.

Williamson, O. (1991). Strategizing, economizing, and economic organization. *Strategic management journal*, 12(S2), 75-94.

Williamson, O. (1996). *The mechanisms of governance*. Oxford. Oxford University Press.

Willig, C. (2001&2013). *Introducing qualitative research in psychology*. Buckingham: Open University Press.

Wolter, C. & Veloso, F. (2008). The effects of innovation on vertical structure: perspectives on transaction costs and competences. *Academy of Management Review*, 33(3), 586-605.

Woodward, C. & West, J. (2009). *Architectural dualities in complex systems: a synthesis of research on technology and organisations*, Paper presented at DRUID 2009 summer conference, Copenhagen Business School, Copenhagen, Denmark.

Yacoub, G. (2015). Heads or tails? The openness-appropriability duality and its implications for innovative performance among UK manufacturing and services, Paper presented at the DRUID Academy conference January 2015, Aalborg, Denmark.

Yin, R. (2014). *Case study research: design and methods*. 5th edition. London: Sage.

Zirpoli, F. & Becker, M. (2007). *Beyond product architecture: division of labour and competence accumulation in complex product development*. SSRN Working Paper Series.

Zirpoli, F. & Camuffo, A. (2009). Product architecture, inter-firm vertical coordination and knowledge partitioning in the auto industry. *European Management Review*, 6(4), 250-264.

Zirpoli, F. & Becker, M. (2011). The limits of design and engineering outsourcing: performance integration and the unfulfilled promises of modularity. *R&D Management*, 41(1): 21-43.

APPENDICES

Interview Date	Reference	Organisation reference	Job title
02/10/2014	Interviewee_1	Firm_1	Head of Product Development
03/10/2014	Interviewee_2	Firm_2	Head of Product Development
10/10/2014	Interviewee_3	Firm_3	Director of Product Development
15/10/2014	Interviewee_4	Firm_1	Director of Product Development
15/10/2014	Interviewee_5	Firm_2	Director of Product Development
18/10/2014	Interviewee_6	Firm_2	Head of Product Development
20/10/2014	Interviewee_7	Firm_1	Senior Product Development Manager
20/10/2014	Interviewee_8	Firm_1	Head of Product Development
21/10/2014	Interviewee_9	Firm_6	Head of Product Development
21/10/2014	Interviewee_10	Firm_6	Head of Product Development
22/10/2014	Interviewee_11	Firm_1	Head of Product Development
23/10/2014	Interviewee_12	Firm_4	Head of Product Compliance
24/10/2014	Interviewee_13	Firm_4	Director of Product Development
25/10/2014	Interviewee_14	Firm_4	Head of Product Development
30/10/2014	Interviewee_15	Firm_1	Director of Product Development
31/10/2014	Interviewee_16	Firm_1	Director of Product Development
04/11/2014	Interviewee_17	Firm_4	Director of Product Development
07/11/2014	Interviewee_18	Firm_2	Director of Product Development
08/11/2014	Interviewee_19	Firm_6	Director of Product Development
11/11/2014	Interviewee_20	Firm_1	Senior Product Development Manager
13/11/2014	Interviewee_21	Firm_2	Director of Product Development
13/11/2014	Interviewee_22	Firm_1	Director of Product Development
20/11/2014	Interviewee_23	Firm_1	Senior Product Development Manager
20/11/2014	Interviewee_24	Firm_4	Director of Product Development
21/11/2014	Interviewee_25	Firm_4	Director of Product Development
24/11/2014	Interviewee_26	Firm_4	Head of Product Development
24/11/2014	Interviewee_27	Firm_4	Director of Product Development
12/12/2014	Interviewee_28	Firm_5	Head of Product Development
14/12/2014	Interviewee_29	Firm_5	Director of Product Development
16/12/2014	Interviewee_30	Firm_5	Senior Product Development Manager
16/12/2014	Interviewee_31	Firm_5	Head of Product Development

Appendix 1: Interview date, firm background and job title of interview respondents

INTERVIEW SCHEDULE

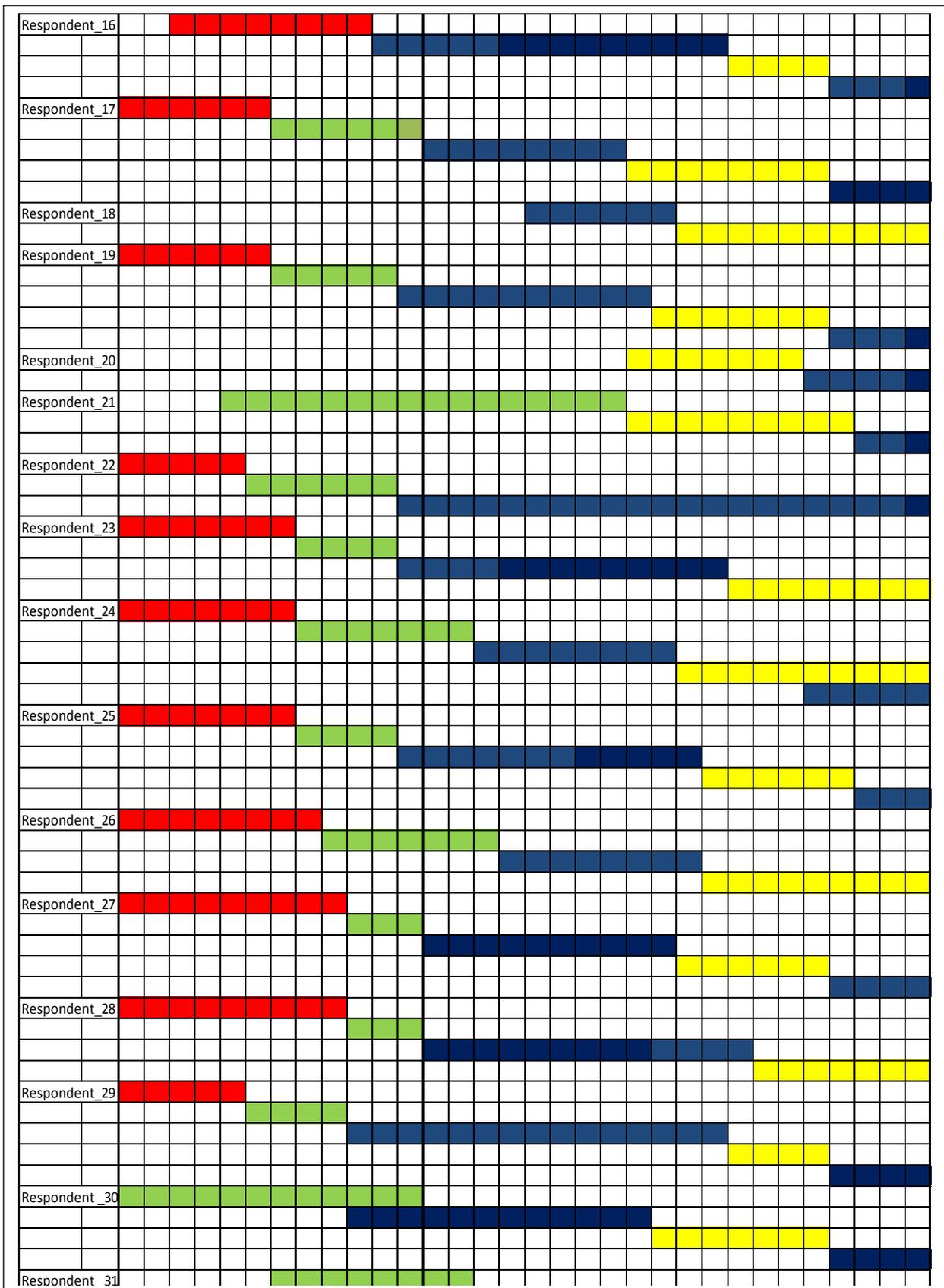
PART 1	
<p>Introduction</p> <p>Background to research study aims and objectives</p> <p>Recorder</p> <p>Consider positionality</p> <p>Allow respondents time to read the product architecture typologies – estimate 5-8 min</p> <p>Respondents to create a product timeline</p> <p>Record results and leave out as aide memoir</p>	
PART 2	
Product questions	<p>Examples</p> <p>What was going on?</p> <p>What led to that?</p> <p>What was the impact?</p> <p>What preceded that?</p> <p>What happened next?</p> <p>What happened first?</p> <p>What influenced it?</p> <p>Who started it?</p> <p>Who did it impact?</p> <p>Why did that happen?</p> <p>When did it start?</p> <p>When did it finish?</p>
Firm questions	
Industry questions	
Probes	<p>Tell me more about that</p> <p>Is there anything else?</p> <p>Why wasn't the impact bigger/smaller?</p> <p>Why did it matter?</p> <p>How does that relate to...?</p>
Close	

Appendix 2: Interview schedule

PRODUCT ARCHITECTURE TYPES	
CLOSED AND MODULAR (CM)	<p>Proprietary</p> <p>Protected by means of IP, secrecy or other means</p> <p><i>Low</i> no. of product components are tightly-coupled and integrated</p> <p><i>High</i> no. of product components are loosely-coupled and independent of each other</p> <p><i>High</i> no. of interfaces between product components are specified but firm-specific</p> <p><i>Low</i> no. of interfaces between product components are idiosyncratic or non-existent</p>
CLOSED AND INTEGRATED (CI)	<p>Proprietary</p> <p>Protected by means of IP, secrecy or other means</p> <p><i>High</i> no. of product components are tightly-coupled and integrated</p> <p><i>Low</i> no. of product components are loosely-coupled and independent of each other</p> <p>Interfaces between product components are idiosyncratic or non-existent</p>
HYBRID (H)	<p>Some elements are proprietary and some are not</p> <p><i>Mix</i> of product components that tightly-coupled and loosely-coupled, integrated and independent</p> <p><i>Mix</i> of interfaces between product components, some are firm-specific and some are industry standards</p>
OPEN AND MODULAR (OM)	<p>Non-protected, controlled by two or more firms</p> <p><i>High</i> no. of product components are loosely-coupled and independent of each other</p> <p><i>Low</i> no. of product components are tightly-coupled and integrated</p> <p><i>High</i> no. of interfaces between product components are industry standards</p> <p><i>Low</i> no. of interfaces between product components are firm-specific</p>

Appendix 3: Product architecture descriptions

Appendix 4 cont'd



Appendix 4: Synthesised product timelines

Faculty of Business and Law

Informed Consent Form for research participants

Title of Study:	Multi-level architectural oscillation
Person(s) conducting the research:	Nicholas Burton
Programme of study:	PhD
Address of the researcher for correspondence:	c/o Eagle Cottage, Chapel Street, Nunnington, YORK, North Yorkshire, YO62 5UP
Telephone:	01439 748510/07827 442914
E-mail:	Nicholas.burton@yahoo.co.uk
Description of the broad nature of the research:	To investigate the pre-conditions, motivations and effects of product architecture change over time, and whether such change is mirrored by architectural changes in firm and industry structure
Description of the involvement expected of participants including the broad nature of questions to be answered or events to be observed or activities to be undertaken, and the expected time commitment:	<p>The expected involvement of the research participants is as follows:</p> <ul style="list-style-type: none"> • Initial exercise assigning generic product firm and industry architecture descriptions to discreet time periods ((approximately 15 minutes) • Follow up telephone call (20 minutes) • Follow up interview (approximately 1 hour). • Any other meetings deemed necessary for the research upon negotiation with the research participant. <p>The interviews will be semi structured and based upon individual's experiences of product and firm architecture within a given industry context</p> <p>The initial exercise will involve research participants being emailed generic architecture ideal types. The ideal types are to be assigned by the research participant to discreet time periods. The research participant will also be given a list of architectural characteristics and need to tick present / not present.</p> <p>The follow up telephone call will be informed by the responses arising from the initial exercise and is aimed at clarifying any potential areas.</p> <p>The follow up interview will explore the pre-conditions, motivations and effects of any changes in product, firm and industry architecture over time.</p> <p>All interviews will be recorded with a digital voice recorder and transcribed.</p> <p>Anonymity will be assured by changing the names of the participants, the organizations and people that they name during the interview in the transcripts.</p>
Description of how the data you provide will be securely stored and/or destroyed upon completion of the project.	All data will be stored securely either electronically on computer or in hard copy version in a locked cupboard. The recordings on the digital recorder will be deleted once they have been saved on the researcher's password-protected laptop. This will be done within 2 weeks of the initial recording. As part of the data analysis process, hard copies of the anonymised transcripts (raw data) may be given to the doctoral supervision team and a small number of other research participants to review. Hard copies will be returned to the researcher.

	The saved audio recordings and text transcriptions will be deleted from the laptop at the end of the project's 'life'. However, it is envisaged that they will be held for a period of up to 5 years, to allow time to maximize the data for publication purposes
--	---

Information obtained in this study, including this consent form, will be kept strictly confidential (i.e. will not be passed to others) and anonymous (i.e. individuals and organisations will not be identified *unless this is expressly excluded in the details given above*).

Data obtained through this research may be reproduced and published in a variety of forms and for a variety of audiences related to the broad nature of the research detailed above. It will not be used for purposes other than those outlined above without your permission.

Participation is entirely voluntary and participants may withdraw at any time.

By signing this consent form, you are indicating that you fully understand the above information and agree to participate in this study on the basis of the above information.

Participant's signature:

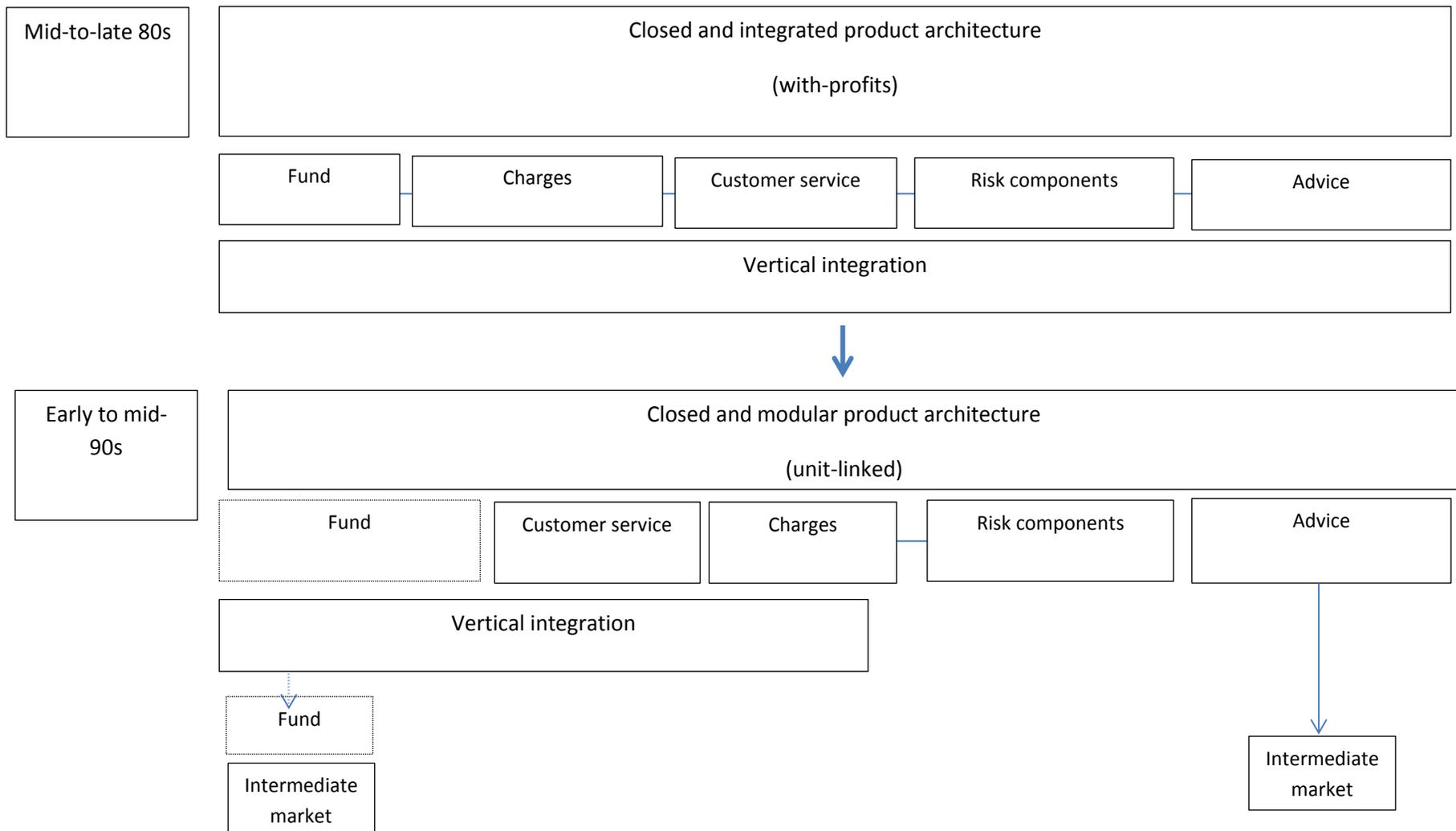
Date:

Student's signature:

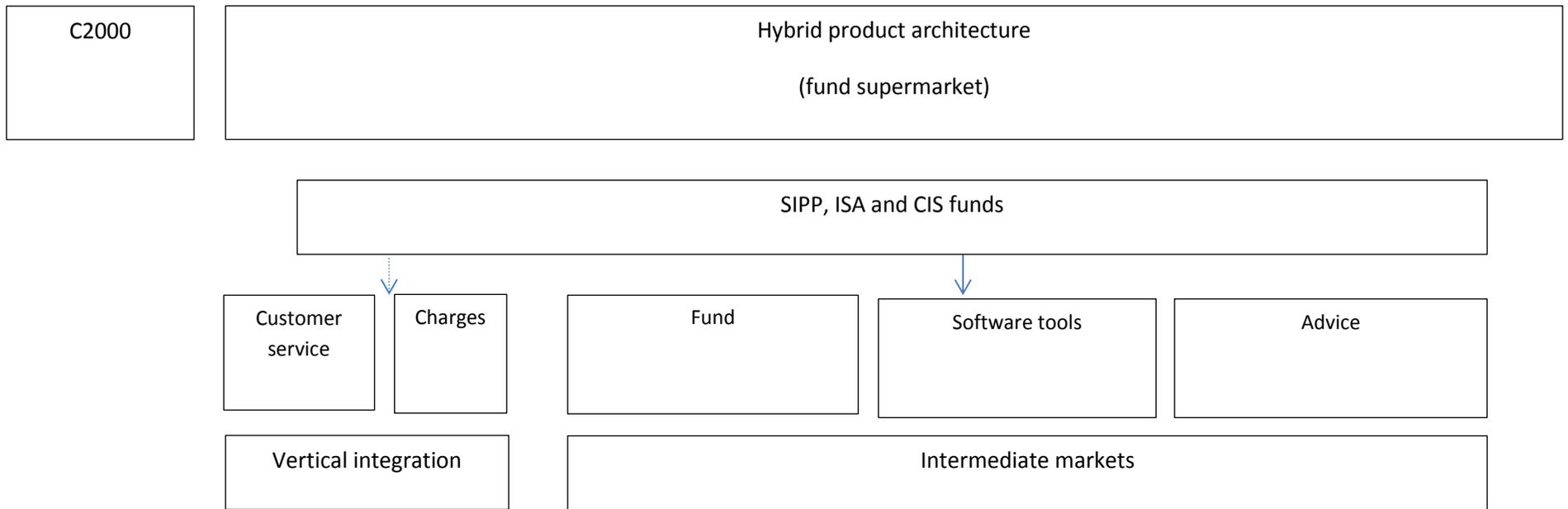
Date:

Please keep one copy of this form for your own records

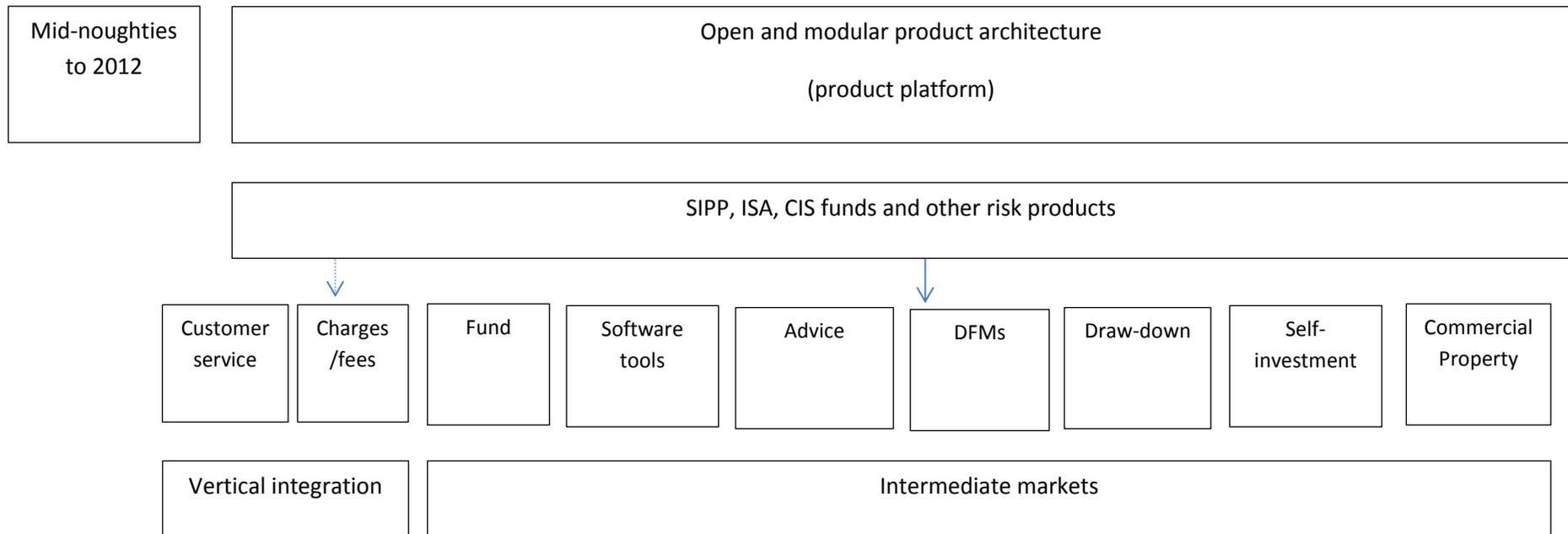
Appendix 5: Informed consent form for research participants



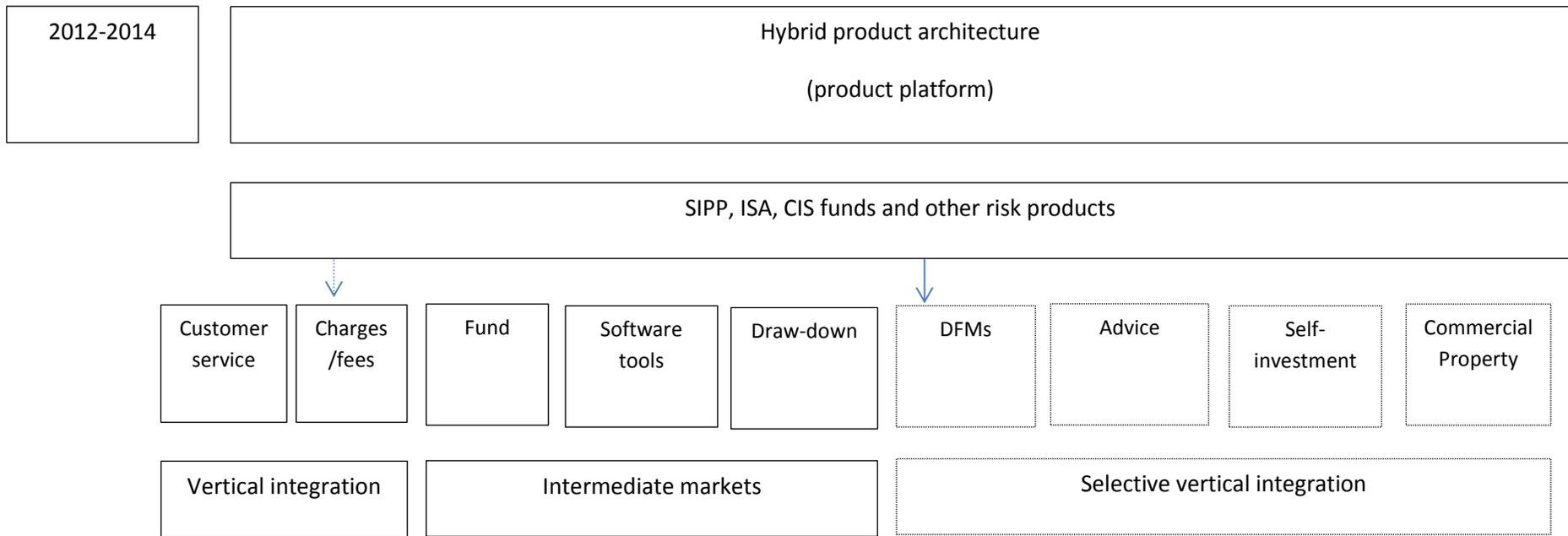
Appendix 6: Summary of evolution towards a closed and modular product architecture



Appendix 7: Summary of evolution towards a hybrid product architecture



Appendix 8: Summary of evolution towards an open and modular product architecture



Appendix 9: Summary of evolution towards a hybrid (platform) architecture