A Basic Economic Case For Reordering The Patent Market With Gain-based Remedies

by

Moshood Abdussalam

LLB (London Metropolitan University) LLM (Stockholm University)

Submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

University of Tasmania, October 2016
Statement of Co-Authorship  
(for inclusion in thesis)

**Publication 1:**  
'The Superiority of the Restitutionary Model in the Computation of Reasonable Royalties for Patent Infringement: A Comparative Discussion'  

**Authors:**  
Moshood Abdussalam *(Primary author)*, PhD Candidate, Law School, University of Tasmania  
Jane Nielsen *(Co-author)*, Senior Lecturer, Law School, University of Tasmania.  
Dianne Nicol *(Co-author)*, Professor of Law and Chair of Academic Senate, Law School, University of Tasmania.

**Location of publication in the thesis:**  
This publication is drawn from Chapter 5 of the thesis.

**Nature of contribution by co-authors:**  
The co-authors contributed to the formalisation, refinement and presentation of the paper.

**Publication 2:**  
'Rethinking the Standard for Ascertaining the Quantum of Disgorgement in Patent Law Contexts'. Accepted for publication in 2017 in the *Journal of Business Law*.

**Authors:**  
Moshood Abdussalam *(Primary author)*, PhD Candidate, Law School, University of Tasmania  
Jane Nielsen *(Co-author)*, Senior Lecturer, Law School, University of Tasmania.  
Location of publication in the thesis:  
This publication is drawn from Chapter 6 of the thesis.

**Nature of contribution by co-authors:**  
The co-author contributed to the formation of the idea, refinement and presentation of the paper.

**Publication 3:**  
(available at: http://jiplp.oxfordjournals.org/content/early/2016/06/27/jiplp.jpw034)
Author:
Moshood Abdussalam (Sole author), PhD Candidate, Law School, University of Tasmania

Location of publication in the thesis:
This publication is drawn from Chapter 6 of the thesis.

We the undersigned agree with the above stated, "proportion of work undertaken" for each of the above published (or submitted) peer-reviewed manuscripts contributing to this thesis:

Signed:

Prof. Dianne Nicol (Supervisor), Law School, UTAS

Prof. Margaret Otlowski, Head of Law School, UTAS

Date: 18-10-2016
Declaration of Originality

This thesis contains no material which has been accepted for a degree or diploma by the University or any other institution, except by way of background information and duly acknowledged in the thesis, and to the best of my knowledge and belief no material previously published or written by another person except where due acknowledgement is made in the text of the thesis, nor does the thesis contain any material that infringes copyright.

This thesis may be made available for loan and limited copying and communication in accordance with the Copyright Act 1968 (Cth).

Moshood Abdussalam
18th October 2016
Abstract

There is a sizeable disparity between the advantages enabled by patent law remedies and the underlying rationale for the patent system. Furnishing solutions to the problem of patent opportunism, which is a product of this gap, is the singular purpose of this thesis. The nexus between patent remedies and the utilitarian social welfare goals of the patent system appears to have been understudied in major patent law jurisdictions, including Australia, Canada and the United Kingdom (UK). In these jurisdictions, focus has been placed more on managing the problems of the patent system attributable to patent law remedies, rather than addressing, head on, the nature of those remedies. This is not, however, the case in the United States of America (USA) where judges, government agencies and academics have expressed ongoing concerns over the tendency of patent law remedies to impact negatively on the social welfare goals of the patent system. Notwithstanding these articulated concerns, the fundamental reason for this negative tendency has been poorly identified, and recommendations to reform patent law remedies in response to it have been equally inadequate.

Ted Sichelman’s work titled ‘Purging Patents of Private Law Remedies’ ((2014) 92 Texas Law Review 528) and that of David Opderbeck titled ‘Patent Damages and the Shape of Patent Law’ ((2009) 89 Boston University Law Review 127) are a timely and valuable response to this problem. Both Sichelman and Opderbeck rightly blame patent opportunism on the ideological dissonance between the purpose of the patent system and the regime of legal remedies applied towards patent infringement. However, both scholars differ in their postulations of solutions to the problem. Sichelman has not yet put forward workable recommendations on how to replace the current remedies (particularly monetary remedies) prevailing in most patent law jurisdictions, with a view to reflecting the utilitarian nature of patents. Opderbeck suggests solutions that are workable but improvable. This thesis adds to the works of Sichelman and Opderbeck by postulating gain-based remedies—chiefly disgorgement and restitutionary
reasonable royalties— as the most pertinent species of monetary remedies suitable to furthering the utilitarian nature and objectives of patent entitlements. It is the submission of this thesis that these species of remedies will be effectual in stemming the tide of patent opportunism by changing the incentives of economic entities within the patent market, and correcting problems that emanate from the patent market's illiquidity.
ACKNOWLEDGEMENT

I am beholden to my fantastic supervisors, Prof Dianne Nicol and Dr. Jane Nielsen, for their sedulous and quality efforts in overseeing my thesis from its conception till its culmination. I am particularly grateful to them for the resilience they showed at the difficult stages of my doctoral studies. They are genuinely wonderful persons. I am also grateful to Prof Gino Dal Pont for the significant academic assistance he provided me, which immensely impacted on the shaping of my thesis. I would also like to register my appreciation to the University of Tasmania for the academic scholarship provided to me, without which my PhD studies would have been impossible. I would like to acknowledge the kind gestures of Prof Brian Love, Santa Clara University; Prof Jane Stapleton, Australian National University; and Sir Robin Jacob, University College London, in giving me helpful feedback on the draft version of Chapter 4 of this thesis.

My profound gratitude extends to several nice persons I had the opportunity to meet and interact with in the course of my doctoral studies who have also, one way or another, shaped not just my thesis and also my person. Persons such as, the current Dean of UTAS Law School, Prof Margaret Otlowski, Prof. Don Chalmers, Dr. Jeremy Pritchard, Prof Jan McDonald, Rick Snell (the Deputy Dean of the Law School), Dr. Olasupo Owoeye, Dr. Jonathon Liddicoat, James Scheibner, Austin Nnajieze, Lucy De Vreeze, Kate Cashman, Joe Wenta, Jan Charbonneau, Kerryn Brent, Riki Mawad, Philipa McCormack and other great persons I have not mentioned. I must also express my special gratitude to the staff of UTAS Law Library, chief of who are: Deborah Bowring, Chris Hurburgh, Carolyn Jarvis and Peter Doyle.

My gratitude to family and friends for their support must also be expressed. I would like to give thanks to my mum, Alhaja B. T. Salami, for her tenacity in seeing my brothers and I through tertiary education despite the dire financial straits my family had fallen into upon the demise of my beloved father, Alhaji G. A. Salami. I would also like to thank my brothers, Jubril, Sheriff, Ibrahim and Abdulrasheed Abdulsalam, for the love shown and the sacrifices they have made towards
enabling me pursue my academic dreams. I would also like to express my gratitude to my cousins, Femi Banjo, Toyosi Awonuga, Yemi Odunsi, Ademola Oyeledun, Ayodeji Oyeledun, Omobo Oyeledun, Adesewa Oyeledun, Oluwatobi Oyeledun, Wonuola Oyeledun, Temitope Oyeledun, Kanyisola Oyeledun, Akin Oyeledun, Oluwaseun Oyeledun and Irewamiri Oyeledun. I must also mention my favourite friends Moshood Anifowose and Abisoye Idris.
Table of Contents

INTRODUCTION ............................................................................................................................................. 1

General Introduction ................................................................................................................................. 2
Responses to the State of the Patent Opportunism ............................................................................ 3
Original Contribution of this Thesis to Knowledge .............................................................................. 4
Scope and Methodology ............................................................................................................................. 5
Arrangement of Chapters ........................................................................................................................ 6

Chapter 1: THE PROBLEM FOR EXAMINATION ................................................................................. 9

1.1 Introduction ........................................................................................................................................... 9
1.2 The Problem of Patent Opportunism Stated ..................................................................................... 9
   1.2.1 The Economic Purpose of Patent Entitlements ........................................................................ 12
   1.2.2 The ‘Propertization’ of Patents and The Drift from the Utilitarian Foundation of the Patent System .......................................................................................................................... 16
1.3 The Patent Market, Its Failure and the New Institutional Economics Analysis .......................................................... 18
   1.3.1 The Concept of the Patent Market ......................................................................................... 19
   1.3.2 The Nature of Transaction Cost Economics and Its Application to the Patent Market .......................................................................................................................................................... 23
   1.3.3 The Advent of New Institutional Economics ........................................................................ 28
   1.3.4 The NIE ‘Recipe’ for Opportunism as a form of Market Failure ......................................... 29
1.4 Conclusion .......................................................................................................................................... 33

Chapter 2: THE NEXUS BETWEEN THE ILLIQUIDITY OF THE PATENT MARKET AND OPPORTUNISM ................................................................................................................................. 34

2.1 Introduction .......................................................................................................................................... 35
2.2 Part I: Information Costs ..................................................................................................................... 35
   2.2.1 What Are Information Costs? ............................................................................................... 35
      2.2.1.1 Patent Scope ............................................................................................................... 37
      2.2.1.2 Uncertainty of Patent Validity .................................................................................. 46
      2.2.1.3 Problems of Notice of Patent Existence .................................................................. 49
   2.2.2 The Implications of Information Costs for the Patent Market (Asset Specificity) ................. 54
2.3 Part II: Bargaining and Negotiation Costs ........................................................................................ 56
   2.3.1 The Macrscopic Level – Patent Thickets .............................................................................. 57
      2.3.1.1 Nature of the Science To which Patents Relate ......................................................... 58
      2.3.1.2 Patenting Strategies ................................................................................................. 61
2.3.2 The Microscopic Level ................................................................. 62
2.3.3 The Implications of Bargaining and Negotiation Costs for the Patent Market (Asset Specificity) ................................................................. 66

2.4 Part III: Overview on Prevailing Solutions to the Problems of Illiquidity.... 69
  2.4.1 ‘Creation of a firm’ ................................................................................. 69
  2.4.2 Government Intervention ....................................................................... 71
  2.4.3 Efficient Allocation of Rights ................................................................. 73
  2.4.4 The Need for Remedial Reform ............................................................... 75

2.5 Conclusion ............................................................................................... 76

Chapter 3: LEGAL REMEDIES AND OPPORTUNISM IN THE PATENT MARKET 78
  3.1 Introduction ............................................................................................... 78
  3.2 Part I: Ex-post Patent Market ................................................................. 80
    3.2.1 Legal Remedies as ‘Prices’ in the Patent Market ................................... 84
    3.2.2 Rationale for and Choice of Legal Remedies in Law .......................... 87
    3.2.3 The Nature of Patent Law Remedies in Australia, Canada, the UK and the USA ................................................................. 90
    3.2.4 Encapsulation of Part I ........................................................................ 93
  3.3 Part II: The Case for the Administrative Model of Remedies ............... 94
    3.3.1 Theoretical Foundations of the Administrative Model of Remedies in Patent Law ................................................................. 95
    3.3.2 Encapsulation of Part II ....................................................................... 101
  3.4 Part III: The Social Costs of the Bargaining Model as Identified by Proponents of the Administrative Model ....................................................... 102
    3.4.1 Classic Holdup ...................................................................................... 103
    3.4.2 Patent Holdout ...................................................................................... 109
    3.4.3 Encapsulation of Part III .................................................................... 112
  3.5 Part IV: The Deficiencies of the Case for the Administrative Model ...... 113
  3.6 Conclusion ............................................................................................... 115

Chapter 4: COMPENSATORY DAMAGES AND THE PATENT MARKET ........ 117
  4.1 Introduction ............................................................................................... 117
  4.2 Part I: The Ideological Gap Between Compensatory Damages and the
Purpose of the Patent System — the First Premise .................................................. 119
4.3 Part II: The Lottery Effects of the Methods of Computing Patent Damages—
The Second
Premise........................................................................................................................................... 123

4.3.1 What Are the Recognised Heads of Compensation in Patent Law? .... 125
4.3.2 The Difficulties in Computing Patent Damages—Primary Heads of
Compensation........................................................................................................................................ 127
  4.3.2.1 Proof of causation and loss where substitutability to user(s) is zero
.................................................................................................................................................. 131
  4.3.2.2 Proof of causation and loss in cases of perfect substitutability .. 134
4.3.3 The Commonwealth Model .................................................................................. 136
4.3.4 The ‘Non-Infringing Alternative Considerations’ Model.............................. 143
4.3.5 The US Conception of ‘NIA Considerations’ Model Is Best, But It Does
Not Obviate Systemic Opportunism.................................................................................... 151
4.3.6 Price Erosion and the Second Premise ............................................................. 153
4.4 Conclusion....................................................................................................................... 154

Chapter 5: REASONABLE ROYALTIES AND THE PATENT MARKET .......... 156
5.1 Introduction ...................................................................................................................... 156
5.2 Part I: The Competing Juridical Bases for Reasonable Royalties.................. 160
5.3 Part II: The Compensatory Model: Substance and Implications for the
Patent Market............................................................................................................................... 164
  5.3.1 The Pith of the Compensatory Model .............................................................. 168
  5.3.2 Pitfalls of the Compensatory Approach.......................................................... 178
5.4 Part III: The Restitutionary Model............................................................................ 184
  5.4.1 The Gist of the Restitutionary Model ............................................................... 186
    5.4.1.1 The Doctrine of Subjective Devaluation Applied to Patent
    Infringement.................................................................................................................. 188
  5.4.2 The Pertinence of the Restitutionary Model to Patent Law.................... 192
  5.4.3 Critiquing the Restitutionary Model ................................................................. 196
5.5 Conclusion.......................................................................................................................... 200

Chapter 6: DISGORGEMENT AND THE PATENT MARKET ............................. 202
Introduction

1. General Introduction

The patent system has always been perceived as a mixed blessing. This is because while it enhances social welfare by encouraging technological and scientific advancements, it also imposes costs on society in creating artificial scarcity of goods and services in that it causes prices to increase above marginal costs. It may also impose costs on society by retarding innovation because it has the effect of excluding unauthorised persons, who may be innovators, from gaining access to patented technologies. Although there had always been scepticism about the social benefits derivable from the patent system, the height of disillusionment with it in modern times is significant. The reason for this is mostly because patent holders are able to both monetize their patents and secure robust payouts, either from courts or in private settlements, in ways that overreach the intended purpose of the patent system. Arguably this state of affairs deals a serious blow to the social welfare objectives of the patent system.

This state of affairs, which is prone to significantly distort wealth redistribution and abrade social welfare enhancements, can be imputed, at least in part, to dissonance between the design of patent law monetary remedies and the purposes of the patent system. Patents are utilitarian entitlements meant essentially to enable inventors and their sponsors to recoup their marginal costs and capture a reasonable measure of the social surplus they have contributed to

---

1 See, for example, the old English patent case of Beaumont v. George (1815), 1 HPC 593.
society through their inventions. However, the monetary remedies currently applied to patents are those consonant with proprietary entitlements of a libertarian nature, designed to protect exclusivity in ownership. This cleavage between the nature of patent law remedies and the purpose of the patent system is the bane of the patent system. This is because while the patent system is aimed at incentivising inventive outcomes, the current regime of monetary remedies appears more concerned about ensuring the exclusivity of inventors in the commercialisation of their inventive outcomes. This results in what is known as ‘patent opportunism’ a condition whereby patentees get legal protection that is more than necessary to incentivise inventive outcomes. This condition poses serious implications for bargaining outcomes within the patent market because patentees are by this reason vested with undue bargaining powers. It should be noted, however, that potential patent infringers could also act opportunistically. Infringers could act opportunistically by securing undeserved benefits that make it difficult for inventors and their sponsors to recoup their marginal costs, which could dampen their incentives to invest in inventive activities. This is by infringing patents

2. Responses to the State of the Patent Opportunism

Responses to the problem of patent opportunism have hitherto been largely from transactional standpoints but this line of solutions has been manifestly inadequate in addressing patent opportunism. It puts forward suggestions on transaction-easing facilities such as patent pools, compulsory licensing and improvements in the assessment of patent application. Scholars and academics in different publications have equally advanced these solutions. However, these solutions,
while of merit, do not go beyond addressing the workability of the patent market, which is innately impossible to effectively manage. This is because patents, being intangible rights over inventions, are fraught with transactional difficulties that render the market for patents to be unworkable in the same sense that one would expect of a neoclassic concept of the term 'market'. Thus, striving to devise institutional measures to ease transactions between patentees and users of patented inventions would certainly not solve more systemic problems that beset the patent system.

Ultimately, the systemic problems of the patent system can be chiefly attributable to the rules for the enforcement of patent rights; chiefly patent law remedies. This is because patent remedies have significant impact on the state of the patent market and the patent system. The place of patent law remedies on the patent system and patent market can best be appreciated when these remedies are understood as factors that shape the incentives that motivate patentees and users of patented technologies to behave in ways that impact both negatively and positively on the conditions of the patent market and system. This position is shared by several academics in their scholarly works. The major shortcoming of these works of academics sharing the same viewpoint, however, is that they fail to address the need for patent remedies to correspond with the utilitarian nature of patents as legal entitlements. Thus, despite acknowledging that patent law remedies impact strongly on the incentives of both patentees and third parties in the patent market, these works regard patents as property rights and urge caution only in the award of remedies.

This approach however, is unlikely to bring about any significant positive change with respect to the ordering of the patent market because it not address place of patent law remedies on the ordering of the patent market and patent system. As things currently stand, there is a gap the nature of patents as utilitarian entitlements and the nature of legal remedies deployed towards the enforcement

of patents. This cleavage creates a situation whereby the social welfare benefits of
the patent system are retarded or at the risk of being rendered unattainable
because room for opportunism is created. However, if patents are treated as
utilitarian entitlements (which they truly are), then remedies suitable to the
utilitarian nature of patents would be applied towards the enforcement of patents
and the social welfare objectives of the patent system would be far less likely to
be defeated.

A case for the treating patents as utilitarian entitlements and accordingly applying
remedies suitable to the utilitarian nature of patents (that is, gain-based
remedies) forms the backbone of this thesis. It is submitted in this thesis that this
is the only way of ensuring that the social welfare objectives of the patent system
are not adversely affected.

3. Original Contribution of this Thesis to Knowledge

The limitations of prevailing attempts at dealing with the problems of the patent
system provided the motivation to pursue this thesis. Although the writings of
numerous scholars have assisted this research, two academic works that have
been primary sources of inspiration. They are Ted Sichleman's seminal paper,
‘Purging Patent Law of Private Law Remedies’9 and that of David Opderbeck titled
‘Patent Damages Reform and the Shape of Patent Law’. 10 Both scholars
comprehensively highlight the utilitarian nature of patent law. Most significantly,
however, they advance a case for the application of legal remedies that correspond
to the utilitarian intendeaments of the patent system, particularly denouncing
compensatory damages. However, while Sichelman rejects disgorgement as a
remedy and proposed a scheme of remedies, Opderbeck makes a case for the
exclusive application of disgorgement as a remedy.

This thesis focuses on patent law monetary remedies with a view to complementing the suggestions of Sichelman and Opderbeck on patent remedial reform. It proposes a design of monetary remedies that advances the utilitarian aims and nature of the patent system. The novel aspect of this thesis is that it proposes the application of gain-based remedies—disgorgement and a restitutioanl model of reasonable royalties—as the exclusive monetary remedies for patent law. It is the submission of this thesis that with gain-based remedies, inventors and their sponsors will be able to capture a significant measure of their social surplus contributions and also recoup their marginal costs, without incentives to invent being negatively affected. Also with this regime of remedies, infringer opportunism can be effectively discouraged without undue waste or sacrifice of resources.

4. **Scope and Methodology**

This thesis embodies a cross-jurisdictional analysis of patent opportunism, focusing primarily on legal materials and practices of major common law jurisdictions: namely Australia, Canada, the United Kingdom (UK) (referred to as the ‘Commonwealth’ jurisdictions) and the United States of America (USA). Notwithstanding the exclusive use and reference to materials and practices of these common law jurisdictions, the discussions of this thesis are applicable to other common law and non-common law jurisdictions.

This thesis is founded almost entirely on qualitative analysis, but with some empirical data references. The main sources of this thesis are jurisprudence and academic commentary from Australia, Canada, the UK and the USA. This thesis also relies extensively on law and economics literature, particularly in Chapters 1, 2 and 3. The law discussed in this thesis is current to October 2016.
5. Arrangement of Chapters

Chapter 1 introduces patent opportunism as the primary thematic concern of this thesis. It identifies patent opportunism as the pinnacle of patent market failure, and discusses the factors responsible for descent into patent opportunism. It also expands upon the concept that patent opportunism is caused by a gap between the utilitarian purpose(s) of the patent system and the property model of patent law remedies. Further, it identifies the analytical approach of this thesis as fashioned after the New Institutional Economics.

Chapter 2 deals with the illiquidity of the patent market which is caused by market failure factors such as transaction costs, bargaining and negotiation costs. In this chapter, it is submitted that patents are inherently illiquid entitlements and that the patent market is unworkable. It is further submitted that both private and public measures aimed at assuaging the illiquidity of patents, can have only marginal impact on reducing the problems of patent illiquidity, especially in circumstances of patent thickets (i.e. an array of overlapping or complimentary patents). Moreover, it is submitted that raising patentability requirement standards, while being a laudable route to reducing the grant of unmerited and socially costly patents, is not a foolproof measure. Finally, it is argued that a more effective avenue for ensuring that the illiquidity of patents does not culminate in opportunism, is to re-examine patent law remedies.

Chapter 3 deals entirely with the place of patent law remedies within the patent market, particularly the effects of patent law remedies on patent opportunism. It is submitted in this chapter that the behaviours of patentees and third party users within the patent market are shaped by the incentives and constraints created by patent law remedies. Therefore, patent opportunism is a function of the outcomes enabled by patent law remedies. Further, this chapter highlights the patent remedies policy debate, which pivots on the dichotomy between the bargaining model (‘property rules’) and administrative model (‘liability rules’), and the question as to which of the two is more suitable to patent enforcement. It contends that the administrative model, usually represented in monetary remedies, bears just as much propensity to create room for opportunism as the bargaining model.
For this reason, this chapter recommends an examination of the capacity of monetary remedies to enable opportunism.

**Chapter 4** discusses the application of compensatory damages to patent law. In this chapter, a case is made against the application of this monetary remedy to successful patent law infringement actions. This case is made on the basis of two arguments. The first is that compensatory damages provide legal protection that overreaches the purposes of the patent system, and as such inherently creates room for opportunism. The second argument is that, by simply adjusting the rules on the ascertainment of causation, the inherent opportunism facilitated by compensatory damages remains insoluble. This is because the nature of compensatory damages is at odds with the purposes of patent law.

**Chapter 5** addresses the reasonable royalties remedy and its place in the patent market. This chapter makes a case for the application of the restitutionary model to the computation of reasonable royalties, and for dispensing with the compensatory model. The line of reasoning employed in this chapter is that the application of the restitutionary model obviates the incidence of opportunism and suits the purposes of patent law as it: i) is suitable to the peculiarities of the patent market; and ii) simply determines reasonable royalties on the basis of the value of infringing activities to the infringer.

**Chapter 6** focuses on the disgorgement remedy. This chapter discusses the ways in which this remedy can be applied in a fashion that suits the utilitarian nature of patents and the peculiarities of the patent market. It is asserted that the term ‘profits’ ought to be perceived from an economic standpoint instead of a legalistic and accounting perspective. Also, it is argued that the quantum of disgorgement should be based on the incremental benefits gained by the infringer from infringing the patent in comparison with other benefits that would have been enabled by the prior art or public domain prior to the infringed patent being granted. Finally, it is argued that, in determining the deductible expenses of the infringer, only the incremental costs or necessarily increased costs to the infringer should be taken into account for deduction.
Chapter 7 is the conclusive component of the thesis. In this chapter, the case for gain-based remedies is concretely made. Before presenting the recommendations of this thesis, the postulations Sichelman and Opderbeck on remedial regime change are comprehensively analysed. Subsequently, the initial, provisional recommendations of this thesis that gain-based remedies should be exclusively applied to patents are systematically examined and affirmed. The ultimate recommendations of this thesis are essentially shaped by the conclusions of Chapters 4, 5 and 6.

This thesis makes a case for the exclusive application of gain-based remedies in the reinforcement of patents as legal entitlements. It is the submission of this thesis that the use of gain-based remedies will ensure the patent market, and ultimately the patent system, can function optimally so that the perennial problem of patent opportunism is significantly abated. This thesis focuses largely on how patent opportunism is facilitated or enabled through the application of monetary remedies to patent law enforcement. It argues that the simplistic and wholesale adoption of monetary legal remedies founded on property law rhetoric to patents is a major cause of patent opportunism.
Chapter 1

The Problem for Examination

1.1 Introduction

This chapter provides a background to the thematic concerns of this thesis. It introduces what this thesis presents as the foremost bane of the patent system, namely the problem of opportunism. The chapter is divided into two main parts. The first part defines and analyses the nature of patent opportunism. It attributes patent opportunism to the inapposite shift of patent law towards the rhetoric of property rights, both in status as a legal entitlement and in enforcement matters. By upsetting the cost-benefit balance of the patent system, patent opportunism is argued to advance the private gains of certain entities at the cost of the interests of the generality of society.

The second part of this chapter introduces the context of patent opportunism, particularly how it shapes and accounts for the interaction of patentees and interested third parties (the ‘patent market’). It assesses the components of the patent market (microscopic and macroscopic levels), and conditions for market failure. It thereby aims to further demonstrate that patent opportunism is the apotheosis of patent market failure.

1.2 The Problem of Patent Opportunism Stated

It is submitted that the chief systemic problem that has always assailed the patent system, undermines it and in fact appears inherent to it, is ‘opportunism’. Why and how this is so will be shown in the course of this thesis. The term opportunism has been described as ‘behavior that is technically legal but is done with a view to securing unintended benefits from the system and these benefits are usually
smaller than the costs they impose on others'. The problem of opportunism is generally manifested in the prospect of patentees being able to engage in 'rent seeking', which essentially means the possibility of patentees securing for themselves, through the assertion of their patents, a quantum of reward not intended by the patent system.

While opportunism has been a matter of significant concern in the USA since the nineteenth century, it has become more pronounced in recent times on account of a proliferation in patent litigation, fierce rivalry for the possession of patents and robust remedial payouts that patentees can secure. In other major patent law jurisdictions such as Australia, Canada, and the UK, among others, the patent marketplace is active but not characterised by these features to the same degree as the USA, especially the ‘robust remedial payout’ element. Yet even in these other jurisdictions there is capacity for systemic and innate patent opportunism, judging by their patent remedial frameworks, as the thesis will reveal.

The patent opportunism problem does not go only the one way; it is reciprocal. Infringers can, after all, also engage in it by infringing patents with impunity. Infringer opportunism diminishes the possibilities for inventors and their sponsors to recoup marginal costs expended in securing inventive outcomes, or make it impossible for them to capture a reasonable measure of the social surplus enabled through their inventions.

Opportunism in the patent system, as in any regime of legal entitlements, is inextricably connected with the remedial facilities of the system. This is because the legal remedies deployed towards the enforcement of patents inform the

bargaining powers of the patentees in transactions or interactions with third parties as regards their patented inventions. Thus, the more generous remedial facilities are towards patentees, the higher the likelihood of opportunism. Patent opportunism is often referred to as ‘holdup’ where the patentee seeks undue financial rewards, and ‘reverse holdup’ where the infringer acts freely in disregard of patent rights.

While opportunism is acknowledged as problematic across patent law jurisdictions in both judicial and academic circles, the key argument presented in this thesis is that it has been addressed from a flawed standpoint. This is why, it is submitted, that prevailing views on patent opportunism in the literature are of limited assistance in the conceptualisation and analysis of opportunism. There are, however, some notable exceptions, particularly Ted Sichelman and David Opderbeck, who provide a more robust analysis. The prevailing views that are considered as being of limited assistance in this thesis tend to focus on the quantification of remedial rewards (i.e. robust remedial payout) obtainable by patentees, while largely ignoring the utilitarian foundation of the patent system. Yet, as this chapter seeks to show, the ideological cleavage between the utilitarian foundation of the patent system and its remedial framework is the chief cause of patent opportunism. Therefore, this ideological gap that should be the prime focus for policy analysts and decision makers in relation to patent opportunism.

The prevailing views on patent opportunism primarily revolve around the question of whether a patentee’s remedial rewards should be fashioned on the

---


19 See, for example, Microsoft Corp. v. Motorola Inc., 696 F.3d 872 (9th Cir. 2012) (USA); see also HTC v Nokia [2013] EWHC 3778 (Pat) (UK).


basis of proprietary autonomy (that is, property/bargain-based rules), or
determined on the basis of a judicial assessment of value (liability/administrative
rules). These views, regardless of the angle from which they are fashioned,
consider patents as property rights and address their protection from a ‘private
entitlement’ standpoint. Room for opportunism is thereby created because a
remedial framework for protecting private or libertarian entitlements differs
materially in substance and effect from one designed to protect utilitarian
entitlements.

This thesis elaborates this difference, and ultimately proffers suggestions to solve
the problem of opportunism that emanates from it. What ensues in this chapter is
foundation to this argument, through an explanation of why it is that patents
should be viewed not as property rights, but as utilitarian entitlements. It also
explains how the ideological gap between patents as utilitarian entitlements and
the present design of patent law remedies is the chief cause of patent
opportunism. As backdrop to this discussion, it is important to first identify the
economic purpose of patents as legal entitlements.

1.2.1 The Economic Purpose of Patent Entitlements

Competing economic theories have dominated discussion on purpose of the
patent system, particularly the nature of protection intended by the patent
system. This is because of the reticence of legislation in most jurisdictions on the
exact purpose of the patent system has left a void filled by economists. The US is
no exception to this, even though in that country appears to be constitutional

23 Mark Lemley and Philip Weiser, ‘Should Property or Liability Rules Govern Information’
(2007) 85 Texas Law Review 783; see also Mark Schankerman and Suzanne Scotchmer, ‘Damages
199.

24 See A Samuel Oddi, Un-Unified Economic Theories of Patents—The Not-Quite-Holy Grail, 71
Notre Dame Law Review 267 (2014); see also Kitch, Edmund, ‘The Nature and Function of the
Patent System’ (1977) 20 Journal of Law and Economics 265; see also Kenneth Dam, ‘The
Review of the Patent System (Study No. 15, U.S. Senate, Committee on the Judiciary, Subcommittee
on Patents, Trademarks, and Copyrights, 85th Cong., 2d sess., Washington 1958), 56; see also
Journal on Telecommunications and High Technology Law 167.
mandate to protect inventors’ interest. Interestingly, however, even though there was a recent statutory overhaul of US patent law by America Invents Act of 2011 at a time when there was robust debate on the design of patent law remedies, that statutory overhaul did not touch on remedies directly.

Without pursuing a discussion on the history of economic thought in this thesis, this thesis submits that the application of libertarian legal remedies to patent law has its roots in classical political economic theory that prevailed during early patent law history. Classical economic theorists reasoned without advertence to the incremental nature of economic matters, particularly matters of utility, costs and benefits. It is for this reason that even though patents were in that period recognized as utilitarian entitlements, yet libertarian remedies such as compensatory damages applied towards enforcement. This is evident in the fact that the incremental social value of patents, on the one hand, and private value of patents (to patentees), on the other hand, factored neither in the evaluation of the patent system nor in the conception of patent remedies to complement the utilitarian nature of patents. Thus, this appears unfortunately to have laid foundation to the use of compensatory damages in patent law; a remedy aimed at enabling patentees recoup their ‘sunk costs’ by charging supracompetitive prices. For this reason patentees are not be limited to the recovery of their incremental (opportunity) costs.

An improvement to classical political economic theory came in form of ‘marginalism’—or neo-classical economic analysis—in the 19th century. The

25 Article I, Section 8(8) of the US Constitution: “Congress shall have power to.... (t)o promote the Progress of Science and useful Arts, by securing for limited Times to Authors and Inventors the exclusive Right to their respective Writings and Discoveries.
marginalism propounded a more systematic and forward-looking perspective to analysing economic concepts such as utility, costs, benefit by analysing economic matters ‘on the margin’ or incrementally using mathematical functions. This approach assesses human behaviour or decisions as shaped by consideration of factors such as utility, benefits or costs, measured ‘at the margin’ or incrementally. It also analyses the patent system and most importantly, issued patents, along similar line—i.e. marginally or incrementally. Unfortunately, although marginalism is a superior economic approach that provides better perspective for the modern justification for the patent system, it has not had impact on the design of patent law remedies.

Going by marginalist economic thought, it is right to hold that patent system is founded on the reasoning that society must suffer some forbearance, in the sense of sacrificing ‘free market’ practices, so as to earn dynamic outcomes from entities that are willing to assume economic risks to attain those outcomes. Thus, by granting a patent to an entity (the ‘patentee’), society is precluded from freely engaging in matters to which that patent relates, except with the permission of that entity. The assumption is that the social costs borne to secure disclosure of inventive technical information from inventors—a dynamic outcome—are generally worth the legal protection provided by patent law. In other words, the patent system is aimed at ensuring that the marginal social benefits derived from the patent system are at least equal to the marginal social cost borne in operating the system—i.e marginal social cost must equal marginal social benefit (MSC =

---


This is why Pigou states that ‘the patent laws aims ... at bringing marginal private net product and social net product more closely together’.33

Modern economists employ a juxtaposition of assessments based on two concepts – marginal social costs and marginal social benefits – in order to ascertain the efficiency or otherwise of a plan or an arrangement.34 Before delving into a discussion of these concepts it is important to note that economists make value judgments on the basis of the additional (i.e. marginal) quality of matters, should they be benefits or costs.35 This is why it is generally said that economists judge or assess matters on ‘the margin’.36 The marginal social costs of the patent system can be said to be the summation of the additional private costs borne by inventors to obtain and maintain a patent, and the additional external costs imposed on society as a result of enforcing patents.37 Thus, the more the social marginal costs of the patent system exceed its social marginal benefits, the more justified is an assessment of the patent system as socially costly or inefficient.

The likelihood of the social costs of the patent system exceeding its social benefits is often blamed on transaction costs, as highlighted in the second part of this chapter. However, it is opportunism that is the most critical cause of the social costs of the patent system exceeding its social benefits. Patent opportunism causes heightened social costs, due to a chasm between the utilitarian foundation of the patent system and the legal remedies applied towards reinforcing patents.

1.2.2 The ‘Propertization’ of Patents and The Drift from the Utilitarian Foundation of the Patent System

Before the year 1994, when the discussions and negotiations leading to the Trade Related Aspects of Intellectual Property Rights Agreement (TRIPS) were concluded, patent law was treated as largely a territorial matter.\textsuperscript{38} This was the case with regard to matters of patent subsistence (or grant) and general attitudes towards protection or enforcement.\textsuperscript{39} TRIPS has produced not only international uniformity on these components,\textsuperscript{40} but augmented a global and more pronounced conferral of property status on patents. In most patent regimes around the world, patents are now accorded property status.\textsuperscript{41}

Article 28 of TRIPS requires signatories to treat patents as a species of personal property.\textsuperscript{42} This property status, it has been said, helps fashion an artificial scarcity in the availability of usable knowledge:

\textldots the patent system creates property in inventions by facilitating and imposing control of the knowledge that defines the invention, therefore enabling limitations to be placed upon its use and supply.\textsuperscript{43}

It assists in solving the ‘appropriability’ problems likely to arise from ‘free-riding’ that may impede the creation of inventive knowledge if inventors and their sponsors are not able to recoup their investments and make appreciable profits for themselves.\textsuperscript{44} According to Paul Jensen and Beth Webster:

\ldots the value of a patent to its owner lies in its ability to prevent rivals from

\textsuperscript{39} Ibid.
\textsuperscript{40} Zahra H Shahzileh, and Ali Maghajan, ‘Patents: Territoriality vs Harmonization’ (2013) 1 International Journal of Humanities and Management Sciences 281.
copying their invention. Without the means to enforce a patent in a court of law (or through other quasi legal means), the owner of the patent cannot recoup the capital invested in its creation.45

Property, broadly being the state of exclusive ownership over subject matter, vests exclusive rights of control and use in those entitled to it.46 As such, to endow patented technical knowledge with property rights is to exclude non-owners from exploiting it commercially.47

The property status conferred on patents is reinforced by legal remedies that ensure that infringers are sanctioned for illegitimate exploitation of patented inventions. Many scholars believe that these legal remedies guarantee exclusivity in the control and use of a subject matter, and thus enable ‘appropriability’.48 They also believe that only when legal remedies ensure exclusivity in the control and use of patents can patents truly afford inventors and their sponsors the incentives to invent, disclose and commercialise inventions.49

The recognition of patents as property rights is particularly reflected in the nature of legal remedies applied towards the enforcement of patents. This is so because the legal remedies deployed to enforce patents align with those applied to protect property rights.50 These include injunctions, compensatory damages, disgorgement and equitable damages (the latter known in patent law parlance as reasonable royalties).51 These remedies are generally aimed at ensuring the exclusivity of patents or deterrence against interference with patents, as when applied to property rights. An argument that is core to this thesis is that the

51 See Chapter 3 of this thesis for a detailed treatment of patent law remedies.
application of this scheme of legal remedies derogates from the utilitarian foundation of the patent system – which proceeds by balancing social marginal costs and social marginal benefits – and thereby enables opportunism. Only in exceptional circumstances, usually in matters of injunctive remedies, are considerations bearing on the need to balance the interests of society and the property owner ostensibly taken into account.52

Patents, it is argued, are best treated as a form of government subsidy to inventors and their sponsors to encourage inventive activities.53 In pursuing this argument, this thesis starts with the premise that the modern patent system, by vesting rights in inventors over inventive ideas, creates a patent market.54 This market is one whereby intending users of patented technologies must bargain with patentees for access to those patented technologies, and is inherently characterised by transaction costs. Although the latter may dampen the social welfare gains derivable from the patent system, controlling them does necessarily solve the problem. This is because, even barring these other factors, the potential for heightened social costs is very much present due to the ideological gap between the patent system and the property rights model of remedies. This gap, as foreshadowed, results in patent opportunism. Focus must therefore turn to the nature of the patent market and transaction cost factors that affect its workability.

1.3 The Patent Market, Its Failure and the New Institutional Economics Analysis

The second part of this chapter traces the origins of New Institutional Economics (NIE) and its pertinence to analysis of patent opportunism. It concludes by

identifying transaction costs as the predominant factor in paving the way for opportunism.

1.3.1 The Concept of the Patent Market

That major patent law jurisdictions, as signatories to the TRIPS Agreement, treat patents as a form of proprietary entitlement, makes it necessary to examine concerns relating to access to patented technologies by third party users. Gans, Williams and Briggs postulate three vital features of property rights: a legal regime that enables the identity of the right owner to be ascertainable (i.e. establishment); a body of rules that enables the right(s) to be legally enforceable (i.e. enforcement); and a means by which interested third parties can contract or transact with a right owner to gain approval to use the right (i.e. exchange). It is this third element that is germane here because it specifically relates to the patent market.

The possibility of patent holders assigning (disposing of ownership by sale) or licensing (allowing use without transferring ownership) patents gradually leads to the emergence of patent markets. A market has been defined as ‘an arrangement by which buyers and sellers of a commodity or commodities interact to determine its price and the quantity of the subject matter of exchange’. A ‘patent market’ is accordingly an avenue whereby patent holders and intending users of patented technology seek each other out over access to patents.

The general economic concept of a market is therefore not circumscribed to a physical space where goods and services are exchanged. It extends to all cases of, and avenues for, repetitive exchanges in return for value or consideration. It is

important to note that the patent market can have two phases— *ex-ante* and *ex-post*— as informed by the conditions of that shape negotiation. The *ex-ante* patent market is one whereby transactions occur before infringement. Where, conversely, transactions occur after infringement or in the shadow of possible infringement litigation, such a phase is described as the *ex-post* patent market. In this chapter the descriptor ‘patent market’ is used to refer only to the *ex-ante* phase. This is because in general parlance a market is where economic entities meet, without the inducement of litigation, to transact.

Gordon’s 1899 work on the origins of compulsory licensing in Great Britain reveals a longstanding recognition of one of the major problems associated with exchanges of patented technologies: the reluctance of patent holders to license out to manufacturers. In contrast, investigations by Lamoreaux et al. into the origins of the patent market in the USA show that the bulk of patent holders in that jurisdiction in the 19th century (who were predominantly non-manufacturers) found it difficult to attract assignees and find licensees. They conclude that information on willing assignees and licensees and on the quality of the utility possessed by patented technology, were always in deficit. This resulted in the need for intermediaries who could bridge the gap between patentees and interested parties. The need to salvage this situation engendered the birth of the market for patents and concomitantly resulted in the creation of a novel specialisation for lawyers concerned with the commercialisation of patents. Lamoreaux et al. also recount how the patent market was characterised by

---


60 Ibid.


difficulties, such that a majority of patents were not licensed or commercialised at all.\textsuperscript{64}

In modern times, where there is a glut of patents,\textsuperscript{65} it is feared that difficulties in securing access to them might be exacerbated. Based on a study of prevailing events in the USA, Eisenberg opines there is widespread infringement of patents, lower rates of successfully concluded licensing contracts and a soaring increase in (threats of) litigation on patents.\textsuperscript{66} Lemley and Myhrvold describe the patent market as ‘blind’.\textsuperscript{67} They argue that an inability to agree over the value of patents, and thus agree on a fair royalty rate, has resulted in patent holders being unable to find would-be users. This, they reason, has left a great number of patents unlicensed and un-commercialised.\textsuperscript{68} Beside the difficulty in placing a mutually agreeable value on patents, Lemley and Myhrvold assert that ‘blind’ market situations may also be attributable to other bargaining and negotiation difficulties that thwart the efficiency of the patent market.\textsuperscript{69}

If patents in any given class are unlicensed and their underlying technologies left unexploited, this may suggest that there is no demand for them. But this cannot be fairly ascribed to an inefficient patent market. Rather, it speaks to better available alternatives or that the patented technologies are commercially irrelevant. It is only where patented technologies are infringed or avoided for fear of infringement liability that the situation can be described as a ‘market failure’. This is because it likely results in a less than optimal exploitation of underlying technologies, thereby causing the social costs of patent protection to be inordinate and diminishing the gain from the patent system. This sort of market failure may


be caused by transaction costs, which may tentatively be described as the costs of facilitating transactions over access to resources or entitlements—in this case patents.

As explained above, the social costs of patent protection are high when the resources sacrificed or opportunity costs incurred for the sake of protection exceed the marginal social benefit derived. Such a state is one of overall ‘failure’, as the patent market may be said to fail when the sum of allocations attained within the market is inefficient.\textsuperscript{70} Patented inventions are be said to be efficiently utilised when they are made available to entities who need them most to meet society’s needs in terms of the provision of goods and services\textsuperscript{71} (i.e. allocative efficiency).\textsuperscript{72} Further, when there is an efficient utilisation of patented inventions, then innovators or inventors can gain access to them for the purposes of technological or scientific advancement (i.e. dynamic efficiency).\textsuperscript{73}

Where allocative and dynamic efficiencies appear to be impaired, economists argue it is as a result of there being a degree of inefficiency or failure in the market. Such a situation generally gives rise to the imperative for third parties to gain access to the patented technologies.\textsuperscript{74} It is important to note that efficient utilisation of patents can only be attained with a workable patent market. It is only when the patent market functions properly are third party users able to access patents and put them to uses that advance social welfare. Inefficiencies in the allocations of resources, by reason of the patent system, are often attributable to transaction costs. Sir Robin Jacob succinctly encapsulates this:

The patent system is there to provide a research and investment incentive but it has a price. That price (what the economists call “transaction costs”)

\begin{itemize}
\end{itemize}
is paid in a host of ways: the cost of patenting, the impediment to competition, the compliance cost of ensuring non-infringement, the cost of uncertainty, litigation costs and so on.\(^75\)

Having identified transaction costs as a common cause of patent market inefficiency, it becomes necessary to address how transaction costs can be dealt with. In pursuing this end, the nature of transaction costs and its implications for the patent market are explored further below.

**1.3.2 The Nature of Transaction Cost Economics and Its Application to the Patent Market**

Since patented technologies are valuable resources to economic actors and also have huge economic implications, it is apt to discuss the patent market problem from an NIE standpoint. NIE is a specialisation of economics best suited to dealing with both the conflicts and relationship between property rights over scarce resources and access thereto (i.e. transaction costs), and how institutions can be re-arranged to manage this conflict and relationship.\(^76\) It provides incisive and thorough analysis, and insight into functioning of markets, including those peculiar to patents. To understand the NIE analysis, however, it is important to first trace its origins to another school of economic thought known as Transaction Cost Economics (TCE).

TCE is concerned with the scientific study of the management of negative externalities (i.e. negative consequences or implications of economic activities that flow or accrue to third parties), and how contracts could be used to deal with them.\(^77\) Externalities may be positive or negative.\(^78\) Where externalities are

---

\(^75\) *Aerotel Ltd v Telco Holdings Ltd and others* [2007] RPC 117, 130.


\(^77\) Łukasz Hardt, ‘The History of Transaction Cost Economics and Its Recent Developments’ (2000) 2 *Erasmus Journal for Philosophy and Economics*, 29, 30 (‘It should be clear therefore, that TCE, understood as the study of the economic study of “costly exchange”, existed a long time before becoming a research program within the framework of economics’).

positive, there are beneficial consequences accruing to third parties from economic activities or transactions.\(^79\) For example, where Family A plays recently released melodious music to visitors on their premises, and their neighbour, Family B overhears and savours the music so played, the playing of the music comes with a positive externality to Family B. Similarly, the patent system can bear positive externalities, providing incentives to develop new goods and services that are beneficial to society.

Negative externalities are the non-beneficial consequences or inconveniences resulting from actions or decisions of economic actors to third parties. For example, if the music played by Family A is loud and distressing to Family B, it amounts to a negative externality to Family B. Similarly, the patent system comes with its own negative externalities. Patents may cause an artificial scarcity of patented goods and services, and also encumber access to patented technologies.

It is generally believed that economic inefficiencies stem from negative externalities not being internalised or accounted for by the parties from whom they emanate.\(^80\) By not bearing the costs of the inconveniences imputed to third parties, creators of negative externalities are emboldened to continue creating such negative effects. Market failure or market inefficiency will likely ensue because potential gains derivable from a given (market) situation cannot be actualised.\(^81\) The patent system bears its own type of negative externalities in the form of increased prices and scarcity in the availability of goods and services, these being inevitable corollaries of having a patent system.\(^82\)

Negative externalities are perceived to be precipitated and perpetuated by transaction costs.\(^83\) This is so because if the creators of negative externalities and those who suffer those externalities can bargain for the purpose of averting or

curtailing them, certain efficient compromises could be reached.\textsuperscript{84} However, any failure to agree in this context perpetuates negative externalities.

Although private compromises of this kind could go some way in managing or abating inefficiencies arising from negative externalities, they are difficult or impossible to achieve when transaction costs enter into the picture. Furubotn and Richter define `transaction costs', in the market context, as the `costs of defining and measuring resources or claims, plus the costs of utilizing and enforcing the rights specified'.\textsuperscript{85} Transaction costs have also been defined as the costs required to:

\begin{quote}
discover who it is that one wishes to deal with, to inform people that one wishes to deal and on what terms, to conduct negotiation leading up to a bargain, to draw up the contract, to undertake the inspection needed to make sure that the terms of the contract are being observed, and so on.'\textsuperscript{86}
\end{quote}

This explains Coase’s argument that externalities are of mutual origin. A patent-based example illustrates the point. Suppose that A (an inventor) has a patent on an electronic component applied by B (a phone manufacturer) in the making of phones. B exploits A’s patent without A’s approval, and A is unable to stop B or to make B furnish adequate recompense for the infringement. B’s action will amount to a negative externality to A, as it could reduce the amount of rent A can capture from the patent; this could in turn reduce A’s incentives to invent.

A’s patent on the technology to which the applied component relates can equally be described as a source of negative externality to B. This is because it is meant to exclude B from applying the patented technology in the manufacture of B’s phone, which could cause market entry barriers to B. These negative externalities are thus mutual in origin. Following Coase’s reasoning, if A had not disclosed his or her invention in applying for a patent, B may not have been able to access the

technology to which the patent relates. On the other hand, had B not infringed, A would not have sought to enforce the patent against B.87

Notwithstanding the merits of Coase’s view that externalities are mutual in origin, the legal system in its calculation of efficiency, wealth distribution and justice allocates rights to certain persons while imposing corresponding duties on others. An invasion or breach of those rights amounts to an externality (or wrong) in judgement of the law.88 Thus in the patented phone component illustration, despite the likely reciprocal effects of A and B’s actions, the patent system is designed to advance A’s interest over B’s. As such, B’s action will be discouraged, while A’s right would be vindicated. This is with a view to enhancing innovation—the cardinal rationale of the patent system.

Coase also postulates that regardless of the initial allocation of rights or entitlement (that is, that a person is conferred with a patent), economic efficiency can still be attained if transacting parties could reach agreement to trade positions.89 Returning to the patented component example above, A (the inventor) and B (the infringing manufacturer) could have reached a bargain. B could have sought a licence from A, with A approving exploitation by B in return for a consideration or price (i.e. royalty). Terms that would govern B’s conditions of use would be determined by A and B. These terms would regulate matters such as the location of use, the quantity of phones to incorporate the component and the duration of such exploitation, amongst others.

The parties can vary the terms of their bargain as it suits them, such that the incentive to invent for A is not eroded and B is not barred from further exploiting the invention or legally sanctioned for doing so. The value of the outcome of the bargaining between such parties can be regarded as a cooperative surplus.90 The dividends of cooperative values to be shared by the parties may, however, depend

87 Ibid, 13.
on several factors, such as bargaining strength, negotiation tactics and agreed price.\textsuperscript{91}

Yet it remains possible that likely gains that inhere in given market situations may not be actualised or optimised because of transaction costs. The potential merits derivable from bargaining between parties over rights are contingent on the weight of transaction costs.\textsuperscript{92} When Coase theorises on the merits of bargaining over rights, he hypothesizes a situation where transaction costs were zero (i.e. transaction costs do not exist), but admitted that such an ideal situation is ‘a very unrealistic assumption’ as transaction costs are generally pervasive.\textsuperscript{93} These costs could be so monumental as to dampen the total efficiency of the bargain when eventually reached or even thwart a bargain in any event.\textsuperscript{94}

This is why Zerbe and McCurdy argue that the presence of negative externalities alone is not sufficient reason to assume that there is market failure.\textsuperscript{95} Externalities are ubiquitous and prevalent in everyday life and, were externalities alone enough to cause market failure, the majority of social engagements and activities would be encumbered. They assert that the singular most defining factor of market failure is transaction costs.\textsuperscript{96} It becomes clear from the foregoing that although externalities are rife, transactions or bargaining could assuage them, but the possibility of completing such transactions successfully is beset by transaction costs.

\textsuperscript{91} Ibid.
\textsuperscript{95} Ibid.
1.3.3 The Advent of New Institutional Economics

The understanding that transaction costs could thwart bargaining and thus lead to market failure, such that potential economic gains are not actualised, has given impetus to the need to reduce or ‘economise’ transaction costs. This need resulted in the advent of NIE, by transforming TCE into an area of study that is concerned with the creation and use of institutions to govern transaction costs.\(^{97}\) NIE, just as TCE, treats transaction costs as cogs in the wheel of economic engagements, just as physicists consider friction a constant feature in the operation of machines.\(^{98}\)

The key difference between NIE and TCE is that while the latter is limited to looking up to mutually arranged contracts to solve problems of externalities, the former focuses on alternative arrangement of institutions to deal with (negative) externalities.\(^{99}\) To illustrate how NIE works, let it be assumed that there is an arrangement of rights such that an entity has patents over inventively purified and isolated genetic materials. However, this arrangement of rights causes the cost of healthcare for certain health problems to be inordinately high. Proponents of NIE would inquire into whether an alternative arrangement – such as tax credits, rewards or other incentives – should be pursued to replace the grant of patents on such inventions.

NIE conceptualises and analyses economics from two levels: macroscopic and microscopic.\(^{100}\) The macroscopic level, also known as the institutional environment, concerns the ‘rules of the game’\(^{101}\) that provide the background norms, principles, law and institutions for governing behaviour and arrangements.

---


of economic actors. The microscopic level, also known as the institutional arrangements, can be described as concerning the (sum of) individuals and organisations, or ‘players of the economic game’. Expressed another way, the macroscopic level consists of the sum or assemblage of laws and institutions that governs society or a given regime, while the microscopic level consists of members of society or actors in a given regime (real, unincorporated and corporate), which are governed or regulated by the former. These two levels of analysis are inseparable because laws and institutions do not exist in vacuum without objects or addressees and, on the other hand, society would amount to nothing meaningful without proper institutions for governance.

1.3.4 The NIE ‘Recipe’ for Opportunism as a form of Market Failure

Having introduced NIE, it becomes necessary to identify the NIE list of factors likely to culminate in a descent into market failure. NIE theorists employ the microscopic and macroscopic levels of economics to analyse market failure. They posit that a poor coordination of both levels has the propensity to culminate in market failure. The aim here is to show how this combination of both levels could cause the patent marketplace to fail. For example, in relation to the macroscopic level, economic actors are usually characterised by cognitive and decisional limitations and the desire to take advantage of other people and situations. With regard to institutional factors, laws may be such that they poorly define rights and provide remedies to rights holders that are over-compensatory or under-compensatory. An interaction of these levels causes market failure. The following

---

factors, attributable to Oliver Williamson’s conception, and falling along the macroscopic-microscopic divide as espoused by NIE theorists, account for market failure:

1) **information costs**: these are the costs of identifying and measuring claims over resources. As described by Henry Smith, information costs include ‘the costs incurred by third parties in processing information about the scope, nature and validity of those rights’. They relate to the macroscopic level of NIE analysis. In Chapter 2, information costs are discussed in detail with a view to showing that they are inherently peculiar to the patent system. That chapter also shows that improving patentability standards, conceiving better theories for patent scope delineation and ameliorating measures to create awareness of patent existence are inherently inadequate to eliminate patent information costs, given the probabilistic nature of patent rights.

2) **bargaining and negotiation costs**: these are the costs of finding whom to deal with, and also costs of negotiating, drafting and concluding contracts. These result from an intersection of the microscopic element of market failure with macroscopic level of NIE analysis. As is also shown in Chapter 2, bargaining and negotiation costs are an admixture of both levels of NIE analysis. In that chapter it is argued these species of costs are inexorable because of the non-transparent and illiquid nature of the patent market.

---


3) **asset specificity**: an asset is said to be ‘specific’ when investments made towards it assume higher value or significance under a given relationship or market situation than outside it.\(^{110}\) It relates to the vulnerability of assets invested towards a given end, which cannot be redeploled towards another without a loss in value.\(^{111}\) For example, assume that A, in the hope of securing a patent over a technical problem of interest, invests time, resources and skills in the development of an invention and eventually secures a patent. Upon the patent being obtained, B and C exploit it without A’s approval. A cannot then undo the disclosure of his or her invention even if A chose to opt out of patent protection. Thus A’s invention is better off under patent protection than outside it after disclosure towards patent application is made. Therefore, A’s invention becomes asset specific to the relationship between A and other would-be users of A’s invention within the patent system.

A countervailing example of asset specificity is where, upon a mistaken understanding of facts, B (due to information costs), with or without any misleading information or misrepresentation from A, invests in the commercialisation of goods falling within the scope of A’s patent. Such investment becomes asset specific to the relationship between A and B, as B cannot rightfully enjoy his or her investment except with A’s permission. Asset specificity is a factor relevant to the macroscopic level of NIE analysis because it relates to the enforcement of entitlements. Asset specificity and its implications are elaborated in this thesis in Chapters 2 and 3 and further illustrated in subsequent chapters.

It is broadly perceived that information, negotiation and bargaining costs, as well as asset specificity, can still be overcome through gainful bargaining between parties.\(^{112}\) Where, however, such bargaining is unlikely, then opportunism—the


\(^{112}\) Ibid.
pinnacle of market failure—sets in. This is because opportunism relates to the ability to take advantage of transaction costs and asset specificity. As such, it is relevant at both the microscopic and macroscopic levels. Williamson defines it as ‘self-interest seeking with guile’, as it ‘includes but is scarcely limited to more blatant forms, such as lying, stealing and cheating’. Williamson considers also that it includes all subtle kinds of deceit characterised by taking advantage of incomplete or distorted disclosure of information to the disadvantage of others. However, his postulation that guile or deceit is fundamental to opportunism appears incorrect, as information costs alone could expose an entity to asset specificity, the condition that predisposes one to opportunism. This is point is returned to in Chapter 2 of this thesis.

Having highlighted the conditions that have propensity to culminate in patent market failure, it becomes imperative to shift into discussing these factors in greater detail, with a view to advancing suitable solutions. To this end, it is apposite to first address the factors of information, bargaining and negotiation costs that are responsible for the illiquidity of the patent market, with the objective to showing the nexus they have with opportunism. The patent market is often described as illiquid because of the impact of transaction costs renders patents to be less tradable. The upshot is that the patent market does not function in the neo-classical sense of the concept ‘market’. The illiquidity of the patent market and nexus with patent opportunism is the focus of the ensuing chapter.

114 Frank Stephen, The Economics of Law (Billing and Sons Ltd, 1988) 185.
116 Ibid.
117 Ibid.
1.4 Conclusion

This chapter served to provide a thematic prelude to the research concerns of this thesis, the primary one being the problem of patent opportunism. Patent opportunism enables patentees to secure for themselves legal protection that goes beyond the purposes of the patent system. It also impedes the possibility of patentees capturing a reasonable fraction of the social surplus they have contributed to society through their inventions.

The chapter identified patent opportunism as a function or product of the ideological gap between patent law remedies and purposes of the patent system. This ideological gap is attributable to the patent system being meant to enhance social welfare—a utilitarian objective—to be contrasted with the present conception of patent remedies fashioned after a property rhetoric.

The second part of the chapter identified the patent market in which patent opportunism occurs. It revealed that the workability of the patent market is often negatively affected by transaction cost factors, providing a springboard for the discussion of the connection between transaction costs and patent opportunism in Chapter 2.
Chapter 2

The Nexus Between The Illiquidity of the Patent Market and Opportunism

2.1 Introduction

In the previous chapter, beyond patent opportunism, transaction costs were identified as responsible for the illiquidity of the patent market, particularly in its *ex ante* phase. As explained in that chapter, the illiquidity of a patent market relates to the difficulties that players or actors within the market face in transacting or dealing with one another in a fashion that brings about efficient outcomes.

In this chapter, the illiquidity of the patent market is analysed with a view to setting a stage for more elaborate treatment of patent opportunism in the next chapter. The ultimate purpose of this chapter is to show that the insolubility of transaction costs and the illiquidity problems peculiar to patents renders the reliance on transaction-facilitating measures almost ineffectual in stemming patent opportunism.

In pursuing the argument that transaction-facilitating measures are ineffectual in averting opportunism, this chapter is divided into three parts. **Part I** addresses information costs. In this part the features of information costs—patent scope indeterminacy, invalidity of patents and notice of existence problems—are discussed with a view to showing they negatively impact on the illiquidity of the patent market. In **Part II** the problems of negotiation and bargaining costs are examined. This Part demonstrates how a combination of the microscopic and

---

macroscopic levels of the patent market, explained in Chapter I, contribute to the illiquidity of the patent market. In Part III, an overview of Coasian solutions aimed at mitigating the problems of illiquidity is presented. Ultimately, it is submitted that these solutions are inherently unable to deal with the illiquidity problems that characterise the patent market.

2.2 Part I: Information Costs

This part builds upon the discussion in the previous chapter. It analyses how information costs affect transactions or exchanges in relation to patents. The chief components of information costs in the patent market context are patent scope, validity and notice of patent existence. This part, in focusing on information costs, is aimed at building towards Part II, which deals with bargaining and negotiating costs. The reason for this is that information costs can have adverse bearing on bargaining and negotiations costs.

2.2.1 What Are Information Costs?

Information costs are independent components of transaction costs. As explained in Chapter I, information costs are the costs of delineating the boundaries of property rights, particularly those costs incurred by third parties in ascertaining the scope, nature and validity of entitlements. Unauthorised third parties are meant to be excluded from entitlements, but in circumstances where parties cannot ascertain the existence of such entitlements, or adequately identify their bounds and validity, they are likely to encroach (unwittingly) upon them, or find it difficult to buy or license access to them. Menell and Meurer consider problems

---

that arise from information costs to be ‘notice externalities’. Notice externalities can have implications for both the holder of rights and third parties. The rights holder is not able to exclude unauthorized persons, and correspondingly, unauthorized persons are unaware of the limitations with respect to their use of those entitlements.

Informational goods or entitlements like patents are inherently susceptible to notice externalities. This is largely because informational goods, unlike traditional property rights such as chattels and real estate, are non-excludable and non-rivalrous. It is for this reason that Arrow posits that if information is not legally protected as an entitlement, creators of new ideas would not have the incentive to disclose, for once information is disclosed it becomes difficult to preclude others accessing and using it. This is popularly recognized as ‘Arrow’s information paradox’.

Notice externalities peculiar to patents can best be understood on account of the nature of patent claims. Patent claims establish the boundaries of granted patents. The legal and economic importance of patents dwells not merely in the patent grant, but in the wording of the claims. Patent claims are usually couched in a single sentence, identifying what the inventor considers to be his or her invention and seeks to have legal protection over. Generally, a single patent will

---

5 Ibid.
include a series of claims, ranging in scope from broad to narrow. There would usually be a broad range of considerations that apply to patent claims, but what is most important are the validity of the patent’s subsistence and the extent of protection provided by law.\textsuperscript{13}

It is through the claims that patent validity is determined on the basis of novelty, non-obviousness, sufficiency and industrial applicability.\textsuperscript{14} This is why a maxim by Giles Rich that ‘(t)he name of the game is the claim’ has a universal cliché in the field of patent law.\textsuperscript{15} To discuss information costs relating to patent market failure, it is important to understand how the matters of patent scope, validity, nature and knowledge of patent existence can affect the workability of the patent market.

\subsection*{2.2.1.1 Patent Scope}

The scope and limitations of patent rights are a function of claim construction and interpretation.\textsuperscript{16} The interpretation of patents claims is notably fraught with uncertainty.\textsuperscript{17} Building on the theory of property rights as a form of legal communication between property owners and third parties, as espoused by Smith,\textsuperscript{18} Janis and Holbrook theorize that patent law is a form of legal communication.\textsuperscript{19} They posit that there are two essential parameters in any situation of communication: proximity (the nexus between the speaker and the audience or between the law and its subjects), and complexity (the ease with which information or legal rules can be understood).\textsuperscript{20} They assert that the patent

\begin{itemize}
\item \textsuperscript{13} European Central Bank v DSS EWCA Civ 192 (19 March 2008) paragraph 5.
\item \textsuperscript{14} Moshood Abdussalam, 'Identifying 'the invention' in Inventorship Disputes' (2014) 11 SCRIPTed 44.
\item \textsuperscript{15} See Giles Rich, Extent of Protection and Interpretation of Claims—American Perspectives, (1990) 21 International Review of Industrial Property and Copyright Law 499.
\item \textsuperscript{16} Lord Russell of Killowen in Electric and Musical Industries Ltd v Lissen Ltd (1938) 56 RPC 23, at 39; See also Robert Merges, Peter Menell, Mark Lemley, Intellectual Property in the New Technological Age (Wolters Kluwer, 2007) 251.
\item \textsuperscript{20} Ibid, 76-82.
\end{itemize}
audience is heterogeneous, as it includes a wide range of participants: technically skilled persons, lawyers with and without technical training, entrepreneurs and other economic actors.\textsuperscript{21} In bridging the gap between patent law and its diverse audience, the conception of a fictional or heuristic person of statutory construct regarded as a skilled addressee or person having ordinary skill in the art (‘the skilled person’) emerged.\textsuperscript{22}

It is important to understand the nature and dynamics of the skilled person, because the scope of patent claims is ascertained and viewed through their lens. The skilled person is considered the hypothetical benchmark used in ascertaining certain technical qualities of patents.\textsuperscript{23} He/she is evoked in determining if the scientific or technological idea for which patent is sought is novel, inventive and if it adequately supports claims.\textsuperscript{24} The skilled person is also used in the construction of patent claims, although Justice Middleton,\textsuperscript{25} a judge of the Australian Federal Court, argues no such statutory duty is imputed upon the skilled person.

When the skilled person is used in patent claim construction, the question arises as to whether he/she is perceived as a technician or technician-lawyer. According to Fromer, the answer to this question is that patent law has two major audiences: the commercial audience to whom the claims are addressed and the technical audience to whom the specification is addressed.\textsuperscript{26} Holbrook argues that when it comes to claims interpretation, it should be borne in mind that patents are technical information that should be interpreted in a technical but not legalistic fashion.\textsuperscript{27} He argues that inventors do not concern themselves with law when they invent but with the dynamics of science and technology.\textsuperscript{28} Therefore, to hold the

\begin{flushleft}
\textsuperscript{21} Ibid, 84.
\textsuperscript{28} Ibid.
\end{flushleft}
patentee to legalism, from a commercial viewpoint, is stifling and insensitive to
the nature of science and technology. However, Golden disagrees with this
position. He is of the view that patent law, being a subject of diversified audience,
needs to have an intermediary who can uniformly cater for the diverse interests
or concerns.  

Golden posits that the drafting and interpretation of patents is a specialized
matter, not for technicians, nor laypersons, but for ‘patently’ enlightened persons
possessing both a technical background and an understanding of patent law, as
applied to commercial ends. In essence his argument is that while judging patent
scope from the standpoint of the technician would be subjective and personal, that
of the technician-(commercial) lawyer will be objective.

It is submitted that although the skilled person is the heuristic person through
whose judgment patent scope is determined, that person is always an appendage
of judicial discretion, which is contingent upon the nature of the science governing
the patent in issue. How the combination of the impacts of judicial discretion and
the nature of science to which patents relate affect the skilled person’s
hypothetical construction of patent claims is at this juncture discussed.

i. Judicial Discretion

The scope of patents cannot be determined in abstract except though proper
judicial assessment. Even a court grappling with a question of patent scope
might struggle with providing clear answers. In fact, empirical studies conducted
by patent law scholars in the USA show that patent claim interpretation or
construction is largely dependent on divergent policy considerations of judges

29 John Golden, ‘Construing Patent Claims According to Their Interpretative Community: A Call
30 Ibid.
Synthesis and Structured Framework’ (2011) 25 Berkeley Technology Law Journal 713; see also,
Law Review 1353-1364.
thus reflecting the influence of judicial discretion.\textsuperscript{32} This conclusion is equally applicable to all other patent law jurisdictions.

While in the USA an approach to patent interpretation known as the doctrine of equivalents (DOE) is used, in Australia and the UK the approach used is simply one of purposive interpretation.\textsuperscript{33} The DOE enable a broad interpretation of patent scope such that that immaterial variants that do not fall within the literal reading of the patent claims, but are technical equivalents attributable to inventive concept, are caught within the scope of the claim.\textsuperscript{34} There has been timeless criticism of the doctrine and even judicial attempts at limiting it on the reasoning that it undermines the informational purposes of patent claims.\textsuperscript{35} However, Supreme Court of the USA has recently, categorically confirmed the continued validity of the doctrine in \textit{Festo Corporation v Shokestu Kogyo Kabushiki}.\textsuperscript{36} The basis for the doctrine of equivalents is largely three: linguistic inadequacies or limitation, patent application or prosecution errors, and unforeseen technological developments.\textsuperscript{37}


\textsuperscript{36} 535 U.S. 722 (2002)

On the other hand, the purposive interpretation approach, used in the UK and Australia, as described by Lord Hoffmann, is concerned with “what the person skilled in the art would have understood the patentee to be using the language of the claim to mean”. As such, this approach differs from the USA’s DOE in that it is ‘claim-centric’ (i.e. strictly focuses on the wording of the claims) and does not involve an expansive interpretation intended to accommodate immaterial variants that third parties may advance. The purposive doctrine and its application have been criticized as courting literalism. Literalism involves giving the patent application its verbatim or (strict) textual meaning, which could enable third parties ride on the coattails of the invention by easily circumventing the patent protection through immaterial variants that do not fall within the textual definition of the claims.

Even where the purposive approach is applied, the impact of policy consideration of judges still influence claim construction outcomes. The judicial history of the popular UK case of *Kirin-Amgen v Hoechst Marion Roussel* brings to the fore how policy considerations come to play in patent scope determination. The patentee in the legal dispute, Kirin Amgen, had claimed a method for the production, through recombinant DNA technology, of the erythropoietin (EPO) hormone. This process involved gene activation using an exogenous medium. The defendant’s production of the EPO hormone involved gene activation through an endogenous medium. The question was whether the defendant’s process infringed despite the different channels used by both parties. Neuberger J, deciding at first instance, found infringement by the defendant. After considering that the patentee’s inventive concept had enabled the production of EPO in accordance with biotechnological methods which was previously impossible to achieve, he continued:

The teaching of the 605 patent and TKT’s technology involve many of the same essential features. They employ the same EPO encoding sequences;

...they involve expressing the same EPO artificially; they do so in a eukaryotic cell; they employ an exogenous promoter; the biotechnological/chemical way in which they express EPO is substantially the same. Neither can be achieved without the essential disclosure—the contribution to the art—of the 605 patent itself...42

However, on appeal, the House of Lords disagreed with this broad interpretation of the patent’s scope and limited the claim to ‘exogenous’ media. Lord Hoffmann said ‘[h]owever, once the monopoly had been allowed to escape from the terms of the claims it is not easy to know where its limits should be drawn.’43 It submitted that the difference in the views of Justice Neuberger in the Court of Appeal and the House of Lords on the scope of the patent in issue was simply informed by the different policy reasoning of the judges on what the scope of a patent should be. This is because by tying the scope of the patent to its inventive concept, Justice Neuberger intended that inventors should be able to enjoy exclusive proprietary interest in the pith of their inventive concept. However, the Court of Appeal and House of Lords were more concerned about certainty in the delineation of the patent scope.

Patent law scholars, economists and judges differ on how patent scope should be determined. Some have reasoned that by broadly interpreting patent claims the patentee is able to get fair compensation for their technical contribution and the incentive to further engage in inventing is undisturbed.44 Others reason to the contrary that allowing broad scope has the propensity to undermine certainty and could even have a chilling effect on follow-on innovation.45 In summary, whatever the implications of broad or narrow interpretation might be, the interpretation of patent claims is a function of judicial discretion.

42 Ibid, at paragraph 625.
ii. The Skilled Person and Nature of the Science Governing the Patent in issue

Judicial discretion on patent interpretation is however dependent on the two other inseparable factors: the heuristic skilled person and the nature of the science to which patents relate. As rightly observed by Chiang, patent claims are abstractions from the (sum of) inventive concept(s) disclosed in the specification.46 However, certain types of patents, due to the nature of their technical origin, lend themselves to abstraction more than others. For this reason, scholars, most notably Burk and Lemley, classify patents as either predictable or unpredictable/experimental science patents.47 This categorization is judicially acknowledged, as is shown below.

Experimental or unpredictable science patents are generally those patents that require a degree of experimentation to put them to work and do not lend themselves easily to speculation: they therefore have low levels of abstraction.48 Generally, patents in the areas of chemistry,49 biomedicine/biotechnology,50 nanotechnology51 and other molecular sciences fall under this class.52 In the US case of Re Fisher this reasoning was clearly expressed by the court:

---------------------------------------------

49 Ibid.
In cases involving unpredictable factors, such as most chemical reactions and physiological activity, the scope of enablement obviously varies inversely with the degree of unpredictability of the factors involved.53

Another notable US case in which the similar position was expressed is Re Wright:

This precept recognizes that one skilled in these chemical and biological arts cannot reasonably predict how different chemical compounds and elements might behave under varying circumstances. Thus, in so called “chemical” patent law practice, the claims of a patent are limited by the scope of what the disclosure reasonably teaches to one skilled in the art.54

In the UK case of Kirin Amgen v Hoechst,55 Lord Hoffmann expressed a similar view, emphasizing that the nature of unpredictable science patents is that they do not lend themselves to generous interpretation:

No doubt there are other cases, not involving figures or measurements, in which the question is whether a word or phrase was used in a strictly conventional or looser sense ... No one suggests that ‘an exogenous DNA sequence coding for EPO’ could have some looser meaning which includes ‘an endogenous DNA sequence coding for EPO’. The question is rather whether the person skilled in the art would understand the invention as operating at a level of generality which makes it irrelevant whether the DNA which codes for EPO is exogenous or not.56

Lemley and Burk theorize that enablement or disclosure and non-obviousness have a bearing on patent scope.57 Accordingly, experimental science patents come with heightened enablement requirements due to the degree of experimentation required to get the invention to work, but they have lower non-obviousness

54 F.2d 1557, 1564 (Federal Circuit 1993).
57 Ibid.
requirements. This means that the patentee would be able to claim only to the level of disclosure he or she has provided.

On the other hand, predictable science patents enjoy broader levels of abstraction. This is largely because they lend themselves to a degree of speculation. Based on the fact that the skilled person, equipped with sufficient information, would have little difficulty in putting them into effect, their claims can be drafted in ways that can capture after ARISING advancements. In the US, the Federal Circuit noted in Spectra-Physics v Coherent that:

If an invention pertains to an art where the results are predictable, e.g mechanical as opposed to chemical arts, a broad claim can be enabled by disclosure of a single embodiment....and is not invalid for lack of enablement simply because it reads on another embodiment which is inadequately disclosed.

Lemley and Burk further argue, with regard to predictable arts patents, that while the disclosure requirements are lower as the skilled person can speculate on how to work the invention with lesser difficulties, the obviousness requirement is higher as the skilled person is likely to speculate on a solution.

Bessen and Meurer have argued that the costs that arise with the ascertainment of the scope of experimental science patents, especially in chemistry and pharmacy, are usually lower given that claims based on molecular or chemical structures are usually better defined. This is not necessarily always true,

58 Ibid; see also, per Lord Hoffmann in SmithKline Beecham Plc’s Patent (2006) R.P.C. 10, at paragraph 64.
59 Ibid.
60 Ibid; see also, Sean Seymore, ‘Heightened Enablement’ (2008), n48.
62 827 F.2d 1524 (Fed. Cir. 1987).
63 Ibid, at 1529.
however, since generally patent claims, irrespective of the field, cannot be
determined in abstract. There have been cases, even without applying the
purposive approach, where patent claims have gone beyond their textual
definition as contained in their claims.66 This is particularly why Jacob LJ said in
Technip France SA’s Patent:

Questions of construction seldom arise in abstract. That is why most
sensible discussions about the meaning of language runs on the general
lines “does it means this, or that or the other?” Rather than the open ended
“what does it mean”.67

The analysis here shows that the scope of patents is dependent not just on judicial
discretion, but also the nature of the science to which patents relate. This clearly
shows that information costs are natural to patents by reason of the
indeterminacy of patent scope, which contributes to the illiquidity problems that
beset the patent market.

2.2.1.2 Uncertainty of Patent Validity

The indeterminacy of the validity of patent claims is another factor that adds to
information costs. Even if patents are granted on the basis of satisfying
patentability requirements, they nonetheless are subject to revocation or
amendment.68 The validity of patents is thus fraught with uncertainty, as patent
offices issue patents that can end up judged by courts as invalid. This situation of
uncertainty has caused Lemley and Shapiro to label patents as ‘probabilistic
rights’.69

One factor responsible for the probabilistic nature of patents is that information
is indivisible— an outcome of a recombination of previously existing information

66 For example, see Generics [UK] Limited trading as Mylan v Yeda Research and Development
68 Mark Janis, ‘Rethinking Reexamination: Toward a Viable Administrative Revocation System for
U.S. Patent Law’ (1997) 11 Harvard Journal of Law and Technology 3-117; see also, Phillip Leith,
69 Mark Lemley and Carl Shapiro, ‘Probabilistic Patents’ (2005) 19 Journal of Economic
Perspectives 75-98.
and complementarity between available information. 70 It is by virtue of indivisibility that it is impossible to unequivocally discern how ideas differ in terms of quality. 71 In addition to the indivisibility of information, judicial discretion is responsible for the indeterminacy of patent validity. The deciding judge may bring several factors to bear, depending on the type of patentability requirement in issue. For example in determining enablement, the court could be lenient or strict, depending on the court’s consideration of the degree of experimentation required by the skilled person to put the invention to work.72

The same could be the case in non-obviousness determinations. The court’s consideration of whether the skilled person would have successfully mosaicked several pieces of technological or scientific information in the search for a solution is sometimes also a matter of judicial discretion. 73 Judicial discretion is particularly prominent when secondary factors, such as commercial success and long-felt market needs, are used to determine non-obviousness.74 In addition, as Powles argues,75 patentable subject matter can be expanded by judicial discretion.

71 Ibid.
75 Julia Powles, ‘Industrial Applicability of Bioscience Inventions in the Supreme Court’ (2012) 71 Cambridge Law Journal 50-52 (‘To avoid chilling such investment, a low threshold for industrial applicability (and other validity requirements) inevitably follows’).
International trade and political factors also shape judicial discretion in respect of the validity of patents. Two factors inform relevant judicial positions on the assessment of validity:

a) the nature of patent laws of a state mirror the extent of transnational investment and trade volumes the state can attract. In other words, courts may be lenient in their positions on patent validity in order to enable patent protection for inventions, usually of controversial nature, applied for by foreign investors or multinational companies.

b) patents serve as alternatives to tariffs in that they discourage importation and enhance the fortification of local industries. This can be easily explained by stating that the bar on patent validity can also be lowered so as to provide patent protection on debatable patent subject matter with a view to encouraging local manufacturers or industries.

These reasons partially account for a proliferation of patents of dubious validity, such as those on software and biological matters. This was well encapsulated by Sir Robin Jacobs in *Aerotel Ltd v Telco Holdings Ltd*:

... there is pressure from would-be patentees on patent offices. People are applying for what are, or arguably are, business method and computer program patents in significant numbers. .......This pressure in part stems from the fact that, following the *State Street Bank* case ...people have started getting patents for these subject matters in the USA. Since they can get them there, they must as a commercial necessity apply for them everywhere. If your competitors are getting or trying to get the weapons of business methods or computer program patents you must too. An arms race in which the weapons are patents has set in. The race has naturally spread worldwide....

77 Ibid.
79 Ibid, paragraph 25.
However, it appears that globally, there has been contraction in the grant of patents of dubious nature, with judicial caution expressed in US cases such as *Bilski v Kappos* 80 and *Alice Corp. v. CLS Bank International*. 81 The simple implication of these cases is that the threshold for patent eligibility of software has been raised, which makes securing the grant of such patents more difficult than was previously the case. Another case of significance is *Mayo Collaborative Services v Prometheus Laboratories, Inc.*, 82 decided in 2011, where the US Supreme Court decided against patent validity of claims relating the ascertainment of medical correlation with a view to knowing the amount of dosage to administer to patients. The court reasoned that the correlation between the administration of a drug and its result on a patient was entirely a natural phenomenon over which a patent should not be granted. The likely effect of this case is that it is apt to dampen the grant of patents on dosage regimen. 83

Scholars have rightly suggested that patent offices are constrained at the time of grant by informational and budgetary considerations, and this adds to the validity conundrum. 84 Patent applications, they say, are *ex parte* and the decision to be made by the patent officer on whether a patent is valid or not, is contingent upon available supporting or adverse information. 85 Additionally, the avalanche of applications in ratio to the examiners at patent offices places budgetary limitations on patent offices, as time, personnel, and other resources have to be optimized to assess applications. 86 This is why it is surmised that patent examiners have little incentive to thoroughly examine patents, but shift the
responsibility of weeding out underserving patents to the courts.87 Lemley, for this reason, described patent offices as being ‘rationally ignorant’.88 In summary, the uncertainty of the assessment of patent eligibility contributes to the information costs problem.

2.2.1.3 Problems of Notice of Patent Existence

Another issue apt to compound information costs is knowledge or notice of the existence of patents. Issued patents are usually recorded in the patent register along with the name of the patent owner(s).89 Where the patent is subsequently assigned, it is usually required that assignees of patents also record the assignment. 90 In some jurisdictions, it is required that exclusive licensees also record their interest in the patents.91

The purpose of recording patent title is to provide notice to third parties so that they become aware of who they may need to secure permission from in order to make use of the patented invention, and also to trace title and other legal interests registered against patents.92 The practice among industrial players is to expend resources in the search for patents that may affect their business activities, so that they are able to decide whether to design around or seek licences for those patents.93 This process is colloquially known as ensuring ‘freedom to operate’.94

90 Section 14 Patents Act 1990 (Cth); Section 33 UK Patents Act 1977.
93 Ibid.
However, where search costs are high, it may not be economically efficient to expend resources searching for patents.95 Lemley argues that since the scope and validity of patents are probabilistic, and the costs of searching for patents may not be socially optimal, a great number of economic players affected by patents simply ignore them.96 Thus, parties simply infringe patents, carelessly or inadvertently, and wait to receive responses from patent owners.97 Mulligan and Lee observe that certain patents are more easily searched than others. They hold that software patents and other predictable arts patents are usually difficult to search for, but that chemical and pharmaceutical patents, and some unpredictable arts patents can be more easily searched for in the register as they are usually represented in chemical formulae.98 But even if that is true, the scope of the patents cannot usually be delineated in the abstract.

Notwithstanding the fact that patentees and interested persons are generally careful to register their titles and interest, it is believed that there is widespread infringement of patents due to notice deficit.99 This can be attributed to several factors such as increased patent races, independent invention, indefinite patent scope, and the large number of patents currently in force.100 An interaction of these factors can impose expensive and inordinate search costs. Due to patent races, interested researchers or research teams could end up contemporaneously contriving inventions with similar inventive concepts to those patented.101 In such

a case, being the first to file permits a patentee to exclude others from the exploitation of the inventive concept(s).102

Patent infringement is founded on strict liability. The negligence or lack of intention of infringers does not exculpate them from infringement liability. Lord Hoffmann clearly sums up the strict liability nature of patent law in *SmithKline Beecham Plc’s (Paroxetine Methanesulfonate) Patent*,103 saying that:

But patent infringement does not require that one should be aware that one is infringing: “whether or not a person is working [an] . . . invention is an objective fact independent of what he knows or thinks about what he is doing”: *Merrell Dow Pharmaceuticals Inc v H N Norton & Co Ltd* [1996] R.P.C. 76, 90.104

Liability persists even in circumstances where the alleged infringer honestly believed the technology was not under patent protection. Blair and Cotter, however, opine that although liability for unauthorized use of patents may be classified as strict liability, it is not so *stricto sensu*—as it is modified strict liability.105 The basis for this reasoning is that patent liability can be mitigated where the alleged infringer was not aware of the patent’s existence due to the fact that the patentee failed to mark his or her patented goods when offered for sale. In Australia106 it is simply required that the patentee marks their patented merchandize as ‘patented’, but in the UK107 and the USA108 it is further required that the patent number must also be provided on the goods to bring effective notice to third parties. In circumstances of failure to do this, entitlement to

104 Ibid paragraph 22.
106 Section 123 Australian *Patents Act* 1990.
108 Section 287 of the US Patent Act 1952
damages may be limited to infringing goods sold after the alleged infringer had actual notice of the patent.109

Essentially, then, patent marking provides constructive notice to potential infringers of the existence of the patent.110 However, it appears that in Australia and the UK the alleged infringer cannot take advantage of this limitation on damages on all occasions. In these jurisdictions, even in the absence of patent marking, it is open to the patentee to argue that the degree of patent awareness borne by the infringer, or proximity with the patent landscape is such that they ought to have known of the existence of the patent.111 In the USA, in contrast, possession of knowledge by the infringer that a product is patented makes no difference to the disentitlement of a patentee from damages where there has been a failure to mark the goods as required.112 It should be noted that process patent owners who do not produce goods are excluded from the obligation to mark.113 In Canada, there is no requirement for marking.114

The reasoning behind this requirement to mark is obviously flawed in modern times where many products are constituted of a multiplicity of patented components. This is why the idea of virtual marking—a method whereby a list of all the patented technologies embodied in a product are enumerated on a given website or database—has been considered as an alternative to traditional marking methods of inscribing or labelling goods with patent numbers.115 Virtual

---

110 Ibid.
112 Section 287(a) of the US Patent Act 1952
marking has become legislatively recognized in the USA, but is yet to be so recognized in Australia and the UK.

The idea of virtual marking works well not only in circumstances where a variety of patents are applied towards the making of goods, but also where they relate to services and software. However, one noticeable flaw of both forms of marking (virtual and traditional) is that they are aimed at giving notice to entities that deal in goods and services in competition with those of a manufacturer-patentee. They are not able to give effective constructive notice to entities that apply patented technologies to goods and services not in competition with the manufacturer-patentee.

An observation that can be drawn from the discussion above on information costs is that however much patent rights are modelled after traditional property rights they fail to properly function as such. Although patents can be transferred to third parties, such transferability is marred by uncertainty, particularly relating to scope and validity. Information costs that arise from patents can have serious socio-economic implications. Relevant patents may be difficult to sieve out from a huge list. This is so even with the presence of technologies that can help identify patent concentration in any given field of economic engagement; a process known as patent-mapping.

Even if relevant patents can be identified, their scope may not be clear and consequently, their validity may be debatable. Lemley and Cotropia articulate that the majority of patent liability is not founded on literal infringement or copying.

---

Rather, it is based on non-literal infringement of claims and this could reflect the facts that interested entities try to design around patent claims or that inventions are independently invented.\textsuperscript{121}

2.2.2 The Implications of Information Costs for the Patent Market (Asset Specificity)

The information cost factors discussed so far have the propensity to create notice externalities and thus make the patent market unworkable for both patentees and users of patented technologies, and ultimately expose both to asset specificity. Asset specificity, as already explained in Chapter I, relates to a situation where assets or investments are of higher value in situations or relationships between given parties, than outside those relationships, usually due to irreversible costs already incurred by one party.\textsuperscript{122} Asset specificity, to the infringer, arises consequent upon patent infringement, in circumstances where investment towards the exploitation of patented technologies cannot be inexpensively deployed to alternative uses. However, to the patentee, investments are made towards inventive ends or the acquisition of a patent such that rewards of such investments are difficult to enjoy due to infringing activities that are difficult to monitor. Such patentee may have to rely on (threats of) litigation, in cases of infringement, or be compelled to engage in costly promotional activities to attract interested parties. This may erode the incentive to invent.

In the same vein, a third party user of patented technology can also be exposed to asset specificity due to information costs. This is likely to happen where such third party invests in the exploitation of patented inventions owing to the indeterminacies caused by notice externalities; investments which they otherwise would likely have avoided. In circumstances where such investments are made and are not easily deployable to non-infringing ends without significant costs,


\textsuperscript{122} See Chapter 1; see also, Christopher Cotropia, ‘Patent Claim Interpretation and Information Costs’ (2005) 9 Lewis and Clark Law Review 58-91; see also, Herbert Hovenkamp, ‘Notice and Patent Remedies’ (2011) 88 Texas Law Review (See Also) 221-233.
then asset specificity arises. The state of asset specificity makes it possible for patentees to engage in opportunism by demanding more than they are entitled to as reparations for patent infringement.

The adverse consequences of asset specificity can be mitigated if parties can contract to exchange economic values. It cannot, however, be denied that asset specificity can attenuate the bargaining strength of a party and thus give the other party a better bargaining position. Notwithstanding the possibilities of decimated bargaining strengths, parties can still bargain to exchange, with the implication that one party gains more than deserved—amounting to opportunism. It therefore becomes necessary to discuss bargaining and negotiation costs as they relate to the patent market.

2.3 Part II: Bargaining and Negotiation Costs

As already explained, if parties with mutually conflicting interests in relation to entitlements or resources (i.e. an owner and a potential user) could reach an agreement over access to those resources, there would be a chance of allocative and productive efficiency. The party who attaches more value to resources would get the desired access and most likely put them to the best use possible. However, the likelihood of achieving such beneficial outcomes becomes bleak where the parties cannot reach an agreement, which can arise when the bargaining and negotiating costs associated with attempting to do so become prohibitive. As already explained in Chapter 1, negotiation and bargaining costs are the costs of finding a willing party to contract with, and also drawing and concluding the contract.

Bargain and negotiations may occur ex ante, before infringement (i.e. before asset specificity), for example, through proper contracting steps or, ex post, after infringement (i.e. after asset specificity) after one party has made irreversible

investments. Where negotiations are conducted *ex post*, they are usually aimed at reducing the implications of asset specificity. In such circumstances, a party affected by asset specificity would tend to have a diminished bargaining position. The prospects of bargaining are, however, conditional upon the mitigating effects of the microscopic and macroscopic levels of economic ordering, whether before or after asset specificity occurs. The macroscopic level factors comprise a combination of information costs, as discussed in **Part I**, with patent thickets caused by fragmented ownership and discrete patent rights. However, the microscopic level factors can be ascribed to a range of matters such as bounded rationality, information asymmetries and strategic behaviour, and irrational judgments. Each factor, as it impacts upon bargaining and negotiation costs, will at this juncture be addressed.

### 2.3.1 The Macroscopic Level – Patent Thickets

The macroscopic aspect of bargaining and negotiation costs is an extension of the information costs already dealt with in Part I. While information costs relate to the ascertainment of scope, validity and existence, this macroscopic level of bargaining and negotiation costs relates to collecting or bundling of needed patent rights together, commonly referred to as the patent thicket problem. Shapiro has put forward what is now widely accepted as the classic definition of patent thickets. He considers patent thickets to be: ‘a dense web of overlapping intellectual property rights that a company must hack its way through in order to actually commercialize new technology’.\(^{124}\) But a patent thicket is not in itself a problem if identification and access to patents were (relatively) easy.\(^{125}\)

According to Egan and Teece, a thicket (i.e. an array of rights) is only problematic when characterized by anti-commons.\(^{126}\) Three kinds of patent thickets, as


identified by Egan and Teece, are of relevance. They are: technologically disparate, complementary and overlapping thickets. Technologically disparate patents thicket arises when a given technological field is awash with patents that have no functional relationship(s), but licence to use them may still be needed for the production of goods and services. In contrast, technologically complementary thickets relate to patents that have functional relationships, such that different patents need to be combined to achieve a given technical outcome or effect. Overlapping patents thicket arises in situations where a subsequently granted patent abuts on the scope of a prior granted patent, such that if the patents were differently owned the subsequently granted patent may not be used without the approval of the original patent owner, nor may the original patent owner use the subsequent patent without permission. Overlapping thicket give rise to what is known in patent law parlance as ‘blocking patents’. Patent thickets can be imputed to two central factors: i) the nature of the science or technological art to which patents relate and ii) patenting strategies.

2.3.1.1 Nature of the Science To which Patents Relate

As already discussed in Part I in reference to information costs, patents can be largely classified as predictable or unpredictable/experimental. These classifications determine the possibilities of disparateness, complementariness and overlap of patent thickets. These classifications bear upon the extent to which a given art lends itself to speculation or predictability. With respect to the overlapping thicket problem, it has been opined that since experimental science patents are likely to have lower inventive step but higher disclosure requirements, there is the likelihood that improvements upon patented matters in that field


127 Ibid.
128 Ibid.
129 Ibid.
would more easily secure patent grant, increasing the risk of overlap. On the other hand, it has been theorized that predictable patents, involving inventions that have higher inventive step but lower disclosure requirements, are such that improvements will not easily secure patent protection, particularly as they will most likely be obvious.

The consequence is that in the experimental science field, a secondary patent can be obtained much more readily than in fields that are more traditional. Such secondary patents may relate to a new use, not being an inherent or secret use, of the already patented matter. These kinds of secondary patents are known as ‘second-use’ inventions. However where the secondary patent is obtained as an inventive extraction from a prior disclosure, which itself may be patented, such invention is known as a ‘selection’ invention.

Grubb and Thomsen note that a selection invention arises when a novel compound or relatively small group of compounds emerge from larger groups previously that have been previously disclosed in broad terms. In both situations, overlapping patents are likely to arise, as one patent (claim) abuts on the other. For example, to illustrate a ‘second use’ invention situation, let it be assumed that Chemical A is invented to be used as textile bleach, but thereafter another person realizes that Chemical A could be used in a particular manner as a preservative for pulp papers. This resulting secondary use can be patented if proven inventive. However, such a secondary use patent would be said to overlap with the original Chemical A patent.

---

133 Ibid.
137 Ibid.
138 Cf. Lord Hoffmann’s criticism of MOBIL/Friction reducing additive Decision G02/88 [1990] E.P.O.R. 73, a European Patent Office case which influential to laying down the rule on second use
In the case of predictable patents, chances of overlapping thickets do arise but not to the extent as for unpredictable patents. When overlaps do occur, they arise as a result of improvements to originally granted patents. The improvements usually come in the form of added matter to the originally patented subject matter, giving it enhanced functionality. The inventor of the added matter will usually not be able to exploit the patent without the approval of the original inventor, as the improvement builds upon the original patent and may not be used without legal access to it.

In contrast to the overlapping thicket problem, there is likely to be more of a complementary patent thicket problem arising from predictable patents than from unpredictable patents. This is because most predictable patents are mechanical or modular in nature and are used in combination with other components in other to create a marketable product. Examples are electronic devices, automobiles and varying contraptions resulting from the complementary assemblage of components. Unpredictable or experimental patents have traditionally been applied in the making of discrete products, such that just one patent is enough to make a product. In such circumstances, there is low risk of a complementary patent thicket.

In recent times, however, this situation has changed due to the complexities of modern science and convergence of hitherto unconnected areas of experimental patents.

---

140 For a good example, see Westinghouse v. Boyden Power Brake Co., 170 US 537, at 573.
science. For example, an amalgam of patented matters in pharmacy, biotechnology and material chemistry (e.g. the crystallization of drug formulation) may need to be fused together to produce a marketable drug. Biotechnology is an area that is fast becoming wrapped in complementariness, particularly with the upstream/downstream dichotomy that arises from the different applications to which basic biotechnology patents can be put. In some other cases however it is possible that neither complementariness nor overlap of patents arises as a problem, but then there is a need for a number of relevant patents to be combined to achieve the production of goods.

2.3.1.2 Patenting Strategies

Hall and Zeidonis theorize that the plethora of patents available in different areas of science and technology may be attributed to a patent paradox. By this, they mean that rather than an increase in patent possessions being a result of heightened inventive activities, it is a product of meticulous patent strategies. Patenting entities stretch inventive concepts to enable secondary and follow-on patents on inventive concepts technically identical or similar to already patented

---

149 Ibid.
subject matter. A collection of patents owned by an entity is popularly referred to as a portfolio.

Wagner and Parchomovsky assert that in contemporary patent milieu, most individual patents have miniscule or no value. Thus, where a patent portfolio consists of technologically related patents, they serve as super-patents that broaden and strengthen offensive, signalling and most defensive uses of the patents collectively. Where, however, a portfolio consists of technologically unrelated patents, diversity in possession would exist but the desired ‘super-patent’ value would be lacking.

Portfolios result from a range of patent filing strategies. Sternitzke refers to these as fencing, blanketing or flooding, and surrounding. He provides the following succinct description for each:

...blanketing or flooding, where certain technological space is covered by various patents in a rather unsystematic way; fencing- i.e., filing multiple patents that describe different technological solutions for similar functional outcomes.....; surrounding, in which a basic patent is surrounded by a competitor’s picket fence, and patent network, such as a certain setup of a portfolio to enhance its overall strength. There are patent portfolios in both experimental and predictable arts patents.

This shows that the quest for the possession of patents inspires entities to exploit all legal avenues to expand their acquisition of patents. This reflects an interplay

---


153 Ibid.

of human and institutional factors that results in the complication of the patent marketplace by negatively contributing to transaction costs on account of compounding the patent thicket situation.

2.3.2 The Microscopic Level

Having discussed the macroscopic aspect of bargaining and negotiation costs, focus now turns to the microscopic aspect. This aspect is occasioned by behavioural and social conditions such as bounded rationality, informational asymmetries and strategic behaviour, irrational considerations and the difficulties of finding parties to transact with.

Bounded rationality relates to the cognitive and decisional limitations of humankind, particularly our limited capacities to collect all necessary information, process it effectively and make quality decisions. The effects of bounded rationality are best perceived in agreements on the price to be paid in the acquisition or licensing of patents. Clarkson explains that there are several customary approaches to determining the price of IP assets but that these assessments are based on parameters which are largely uncertain, nebulous and disputed by parties due to bounded rationality on one or both sides. Therefore, in most cases, parties end up with ‘satisficing’ or heuristic prices, which is apt to cause patents to be overvalued or undervalued.

Merges explains that, where there is a divergence in the valuation of patents, the prospects of transactions occurring will be low. This is likely to culminate in the protraction, if not frustration, of agreements. Bounded rationality, Merges also posits, has led to major historical breakdowns in bargaining. He cites the examples

---

157 Ibid.
158 Ibid.
160 Ibid, 78-91.
of the Marconi and De Forest companies, which failed to work in collaboration with each other over their respective patents, leading to stagnation in the advancement of radio technology until government intervened. He also cites the example of the historic Bessemer and Mushet blocking patents, which retarded the advancement of steelmaking technologies in the USA. These incidents, he opines, could have been avoided had the different patent owners agreed to work together. In short, bounded rationality is apt to upset the smooth functioning of the patent market by causing both sides of the bargain, the patentees and users, to lose sight of the likely cooperative surplus that can accrue to them through bargaining.

Asymmetric information and strategic bargaining relates to a situation in which there is a cleavage or imbalance in the quality of information possessed by parties in a given situation. Akerlof theorizes that if a party to a proposed transaction possesses more information than other parties, the party with the informational advantage is most likely to capitalize upon that edge to his own benefit. This, in most cases, would be detrimental to the interests of the other party.

Other scholars have supported the proposition that where there is information asymmetry, the party with the information edge is likely to act strategically to his favour and get the best deal for him or herself. This is certainly plausible. It has been argued, with respect to patents, that the patentee is likely to possess better information about the value and technological merits of his or her patents. It follows that ‘(t)he licensee as the less-informed party has to be educated as to the value of the innovation.…’ Others, however, argue that intending users are more likely to possess better information about the value and technological merits of inventions, rather than the inventors. Reepmeyer et al. studied licensing in the

---

161 Ibid.
162 Ibid.
165 Ibid.
pharmaceutical industry.\textsuperscript{166} They assert that licensees are more likely to possess better competence in developing inventions, dealing with regulatory bodies, carrying out clinical testing in a timely manner, producing market quantities and improving marketing networks.\textsuperscript{167} Consequently, they are in a better position to assess the value of patents.

The upshot of information asymmetry in the patent market, as regards transaction costs and illiquidity, is that both patentees and users of technology might hold out. In either case holding out could interminably delay or frustrate the conclusion of a bargaining process. This could have the effect of depriving not just both parties the dividends of cooperative surplus, but also dampening social welfare.

Irrational considerations, as Merges asserts, being factors that are not founded on bounded rationality but on sheer irrationality, could also stifle bargaining. As Merges notes: ‘This is not to suggest that a party pursuing this course must be mentally ill or deficient; instead, it is meant to be a catch-all phrase to include motives such as spite, pride and anger’.\textsuperscript{168} There are various examples of irrational considerations: patentees may have unrealistic expectations of the value of their patents; third parties may choose deliberately not to license-in patents, even when they are aware of their existence;\textsuperscript{169} patentees may, for reasons of pride or commercial vendetta, refuse to license.\textsuperscript{170} The obvious implication of irrational considerations is a retardation of the smooth conclusion of patent transactions.

Difficulties in finding willing parties can be another crippling problem at the microscopic level. Agrawal et al. argue that ‘market thickness’, which relates to the ‘opportunities to trade with a wide range of potential transactors’, is one of the

\textsuperscript{167} Ibid.
biggest problems confronting the market for ideas, particularly patents. Based on a survey of firms licensing out patents in Europe and Japan, Zuniga and Guellec conclude that an inability to find parties intending to license had hampered the development of a patent market or the effective licensing of patents. Gans and Stern argue that the presence of market intermediaries or private market mechanisms that can help bridge the gap between patent owners, and willing buyers and licensees could help improve the market thickness deficit. In the USA, there appears to be a fast-growing and somewhat settled practice of technology transactions being facilitated by patent brokerage services, where firms such as Ocean Tomo, and Intellectual Ventures aggregate patents and offer them for sale/auctions and licensing. This has been emulated in other jurisdictions, including as the UK and Australia. However, information costs appear to be the predominant factor responsible for this phenomenon, even with the growing presence of market intermediaries.

In sum, these microscopic factors can bring about a frustration of the patent market, either alone or in combination with one another. Where they do arise, asset specificity might be difficult to overcome and as such the potential for opportunism becomes increasingly possible.

2.3.3 The Implications of Bargaining and Negotiation Costs for the Patent Market (Asset Specificity)

Where bargaining and negotiation costs are so overwhelming that they make transactions unattainable, then asset specificity lingers and chances for opportunism loom. One major manifestation of asset specificity in cases of patent thickets is royalty stacking. Royalty stacking, as defined by Lemley and Shapiro, embraces ‘situations in which a single product potentially infringes on many patents, and thus may bear multiple royalty burdens’. The reason for this is that several patents have to be bundled together to make a product, each requiring its own licensing contract to be drawn up. Where royalty stacking occurs, two market drawbacks tend to arise: double marginalization and Cournot monopolies/oligopoly. Double marginalization results when the high costs of procuring inputs ultimately leads to expensive final products. If the costs of patent licensing were high, the producer would have to sell the final product at a high price to recoup the cost of production.

On the other hand, Cournot monopolies/oligopoly arise when complimentary inputs are needed to be conjunctively used but one or more of the input owners (threaten to) hold out in a bid to secure a higher reward on their input. Amir and Gama describe Cournot monopolies/oligopoly as arising where ‘n firms sells n different products that are useless unless they are used together...’ These products (or inputs) would need to be combined together by the buyer to create a finished product. Mossoff provides insight into the thicket experiences of the US sewing machine industry of the 1850s, particularly how serious transaction

---

breakdowns resulting from the thickets led to innovation stagnation in the sewing machine industry of that era.182

The tendency for the royalty-stacking problem to arise can be illustrated using the following hypothetical situation. If Element A and Element B were to be combined in order to make a device, let it be assumed that the prevailing market price of both elements is similar. If a contract is concluded over access to Element A, the owner of Element B may choose not to conclude a contract with a view to holding out until the manufacturer of the device agrees to a higher price for Element B.

If contracts are not concluded due to patent thickets, either of the following is likely to arise: patents are likely to remain unused because of the fear of infringing and legal implications of infringing;183 or manufacturers who are in need of the patented technologies but are unable to obtain legitimate access through licensing contract wilfully infringe those patents.184 In these circumstances, where patents are infringed detection and monitoring are often difficult to attain.185 In other cases, patent thickets may cause manufacturers to innocently infringe patents due to the non-optimal search costs that accompany them.186 These possibilities could spell asset specificity problems for both holders and users of patented technologies.

It is believed that a proper market for patents, in the neoclassical sense of the concept ‘market’, cannot be said to exist due to information costs that beset patents.187 According to this belief, a true market is one wherein common

practices, rules, standards and conventions govern conduct and transactions.\textsuperscript{188} This uniformity in factors is only likely to be found in the market of homogeneous goods, where complete information and transparency is likely.\textsuperscript{189} These factors cannot be easily located in the patent market or other markets of informational goods. It is for this reason that transactions over patents would naturally be mired in bargaining and negotiation costs.

On this basis, it seems fair to conclude that the microscopic conditions of transaction costs will only end up obfuscating a patent market that is already inherently beleaguered by information costs. The implication is that the patent market will inherently be illiquid and dysfunctional and, as such, likely to expose parties to asset specificity.\textsuperscript{190}

\textbf{2.4 Part III: Overview on Prevailing Solutions to the Problems of Illiquidity}

As can be gathered from the foregoing discussion, illiquidity of the patent market, which has the propensity to expose entities to asset specificity, is attributable to two major strands of transaction costs: information costs, and negotiation and bargaining costs. Coase, in his celebrated paper \textit{The Problem of Social Cost},\textsuperscript{191} reasoned that when transaction costs are apt to impede successful exchanges between entities, three major arrangements could be pursued in response. They are: a) the creation of a firm;\textsuperscript{192} b) government intervention;\textsuperscript{193} and c) efficient allocation of rights.\textsuperscript{194} These suggested solutions find equivalence in prevailing

\textsuperscript{188} Ibid.
\textsuperscript{189} Ibid.
\textsuperscript{192} Ibid, 16.
\textsuperscript{193} Ibid, 17.
\textsuperscript{194} Ibid, 42-44.
arrangements within the patent system devised to abate the effects of transaction costs, and reverse the illiquidity that besets the patent market. In fact, measures on how to improve these three arrangements formed the nucleus of the Australian Productivity Commission’s 2013 Report on Compulsory Licensing of Patents, with the aim of dealing with patent market illiquidity.195

2.4.1 ‘Creation of a firm’

Where transaction costs makes it difficult to reach bargains, it could be more advantageous to create a firm which drives down such costs.196 According to Coase, a firm is created when a long-term contract is reached between buyers and sellers thereby enabling the buyer to demand for the supply of goods or services from the seller at any point in time.197 In other words, a firm is created when a short-term contract would be unsatisfactory.198 By creating a firm, transaction costs, which could impede smooth access to resources needed for production, are obviated because the buyer is able to incorporate the seller within his or her organization through the long-term contract.

By incorporating the seller within the buyer’s organization, market transactions or exchanges are substituted with a coordination of resources by decisions within the firm.199 A number of private arrangements within the patent market are equivalent to Coase’s firm, the most popular of which are patent pools. A patent pool may be described as a package or collection of several different patent licences, usually comprising related technologies, which enables different users to gain access to such patent rights.200 Other similar measures include cross-licensing or joint licensing, patent clearinghouses and licences of right.201

197 Ibid, 391.
199 Ibid.
200 Ibid.
Verbeure rightly enumerates the merits of private arrangements of this kind: to be: a reduction in litigation; enablement of efficient use of patented inventions; abating royalty stacking problems; and maintaining the autonomy of patent owners, as it helps avoid the need for compulsory licensing. However, she equally identifies the flaws of these kinds of arrangements to be: high costs of setting up such an arrangement in terms of funds and time; the shielding of invalid patents; enablement of inequitable remuneration to patentees; and collusion and anti-competitive practices among pooling patentees.

Kieff also notes that arrangements of this kind are usually exclusive to members and closed to outsiders or non-members. This is not true in all cases, however, as some arrangements are designed so that they are openly available to entities that are able to afford licence fees. While private arrangements of this kind can help assuage transaction costs, they have a more serious flaw that goes to the root of their capability of significantly reduce transaction costs. This is the fact that such private arrangements are usually voluntary and for this reason certain patentees may hold out, causing the arrangement to be an incomplete assemblage of technically related patents.

2.4.2 Government Intervention

As rightly reasoned by Coase, resorting to the creation of a firm might be costly in administrative and resource terms, such that pursuing a private arrangement to avoid transaction costs might have little or negligible marginal value over usual

---


203 Ibid.


market exchange. An alternative solution might be found in government intervention in the market. According to Coase, the government functions as a super-firm in that it coordinates recourses more expediently because it can avoid the market and its attendant delays by making peremptory orders that bring about immediate outcomes. As Cooter et al. explain, in circumstances where parties are unable to reach private settlements to share stakes or gains, it is only right for government to intervene to dictate such allocations for them as it thinks fit.

In the patent market, government interventions to abate transaction costs are usually manifested in measures such as compulsory licensing. Compulsory licensing may be described as an institutional innovation, administered by the judiciary and other relevant State agencies, which enables third parties to gain access to patented inventions upon the satisfaction of certain standard requirements. Although England already had compulsory licensing scheme in 1883, a global adoption of compulsory licensing is generally traced to a 1925 as marked by a revision of the Paris Convention for the Protection of Industrial Property (Paris Convention). This revision of the Paris Convention permitted States to award licences to third parties against the wish of patentees in circumstances where it was considered necessary to encourage local working of the patents. The Paris Convention provides an alternative to the confiscation of patents by governments in circumstances where patentees did not locally practise

207 Ronald Coase, The Problem of Social Cost (1960), n191, 16-17
208 Ibid, 17.
the patents, by allowing only compulsory licensing of such patents at a fee nominated by the State or its agencies to compensate for such mandated use.214

Article 31 of TRIPS serves as the modern international authority for the award of compulsory licences. It lays down the accepted conditions upon which compulsory licence may be awarded, including: non-working of patent to satisfy local market needs; national emergency; public interest; and where an improvement or dependent patent (‘second patent’) cannot be practised without access to a given patent (‘the first patent’).215 The likelihood of applications for compulsory licences being granted is considerably low. This is largely because of two factors:

a) courts view patents as a form of property right and compulsory licences as a derogation from the patentee’s autonomy;216 and
b) courts have a systematic determination process and as such take time to conclusively reach a decision on such applications.217

While compulsory licensing can in certain situations help avert the effects of transaction costs, it is generally likely to have little advantageous value in circumstances of patent thickets occasioned by diverse patent owners. Where a mass of needed patents are owned by one entity or a few entities, then compulsory licensing may aid in avoiding the effects of transaction costs. However, where a patent thicket consists of diversely owned patents, interested parties would incur inordinate expenses in pursuing discrete applications, and also expending time and resources towards that end. The result is that compulsory licensing measures,

just as with the ‘creation of a firm’ solution, are an incomplete facilitation of access to patent rights.

2.4.3 Efficient Allocation of Rights

Whether the efficiency of outcomes resulting from transactions would be considerable or inconsequential depends on the original allocation of rights.\(^{218}\)

Thus Coase notes:

> In these conditions the initial delimitation of legal rights does have an effect on the efficiency with which the economic system operates. One arrangement of rights may bring about a greater value of production than any other. But unless this is the arrangement of rights established by the legal system, the costs of reaching the same result by altering and combining rights through the market may be so great that this optimal arrangement of rights, and the greater value of production which it would bring, may never be achieved.\(^{219}\)

According to Coase, before we embark upon electing allocation of rights it is essential to compare the marginal social product (MSP) of possible alternative regimes of allocations to see which enhances social benefits more effectively. This can be done by taking Regime A, and assessing its MSP (i.e. private benefits to owner of rights \textit{minus} the opportunity costs suffered by society in enforcing that right) and comparing the result with the MSP of a Regime B, or C or even more alternatives.\(^{220}\) Then the MSPs of these regimes must be compared to see which yields the best MSP before electing a regime upon which to found an allocation of rights.

This can be illustrated by giving a patent-based example. Let it be assumed that a patent with a patentability defect (i.e. with doubtful validity) is erroneously granted and is available to users at 10% of a user’s net profit. Let it be further assumed that the said defective patent consists of validly patentable subject

\(^{219}\) Ibid, 16.
\(^{220}\) Ibid, 40.
matter and non-patentable subject matter. If the defective elements were eliminated and the patent re-issued, assume that the fair market value of the patent royalty would be 5% of the user’s net profit. It would be clear that with ‘10% of net profits’ royalty; the social cost imposed can be valued at ‘5% of net profits’; being the difference between the inflated royalty fee and the fair royalty fee (10%-5%). This is on account of the opportunity cost of the ‘5% net profit’ inflated value attributable to the non-patentable elements. However, with an adjustment to a ‘5% net profits’ royalty, where the defective patentability factors are removed, the marginal social costs of that rate finds confluence with marginal social benefit enabled by the patented invention.

Putting this example in context, and relating it to the application of Coase’s suggestion that legal entitlements should be arranged in ways that they bring about efficient outcomes, it would be better for any patent system to strive at ensuring that only valid patents were issued. This is because it is only by ensuring that valid patents are issued that the marginal social benefits and the marginal social costs of patents meet or converge. However, if we sought only to ensure that transactions were successfully reached, as aimed with ‘creating firms’ and ‘government interventions’, without being mindful of the inherent social costs of the products of such transactions, then society might overall lose in social welfare terms. Thus ensuring that the initial allocation of rights is efficient evokes the need for a set of ex-ante measures essential to improving patent entitlement standards by raising patentability requirements.

As strongly argued by Olson\textsuperscript{221} and by Chiang\textsuperscript{222} the patent office and courts reviewing patentability cases can adopt a cost-benefit analysis on patents sought for grant. Correspondingly, only patents that measure up to an assessment of the marginal social benefit vis-à-vis the marginal social costs of patents will be awarded. While this is theoretically appealing, in reality it is not feasible. This is because, as discussed in Part I and II, such cost-benefit analysis can be easily


eroded by the creativity and deftness of patent applicants and their lawyers. It is also easily undermined by the bounded rationality of the cost-benefit assessor (i.e. the patent office or judge); and political forces, e.g. international trade and the quest for foreign investments, among other factors.

2.4.4 The Need for Remedial Reform

The near impossibility of ensuring that patents are awarded on grounds of efficiency to eschew market failure problem does not mean there is no possibility for redemption. Rather, by following the New Institutional Economics reasoning, solutions can be found in changing the rules of the patent game so that the incentives and ordering of matters by players of the patent market would correspondingly change. This is the reason why Sichelman moots a change to the rules of the patent game hinged on the idea of reforming the patent system from a remedial angle. Sichelman reasons that:

instead of focusing on the substantive rule at issue......policymakers and scholars should also examine ways to adjust the manner of enforcement, judicial procedure, and remedies to achieve effective substantive aims. When the cost of particularized substantive rulemaking is high, as in the case of patent law, particularized enforcement, procedure, and remedies may provide a better route for achieving optimal outcomes. Unfortunately, modifying enforcement approaches has generally been overlooked as a means for compensating for defects in the primary substantive law at issue.

... modifying enforcement in the judicial or executive domains may significantly reduce the costs imposed by the substantive rule. If these modification costs are relatively low, then “measuring” enforcement or remedies to change the effect of a substantive rule may be a superior alternative to modifying the substantive rule directly via legislation or regulation.223

Sichelman’s view is that the patent market can be reordered to mitigate the implications of illiquidity by re-designing the remedial framework of patent law. To appreciate and understand the need for reforming patent law from a remedial angle, the place of legal remedies in the patent market has to be first understood. Thus, the next chapter shifts to an analysis of the significance of remedial aspects of patent law to the patent market.

2.5 Conclusion

The illiquidity of the patent market has been analysed in this chapter. The problem of patent market illiquidity, as explained in this chapter, most often precedes the emergence of opportunism. As can be discerned from the discussion of this chapter the illiquidity problems of the patent system are peculiar and markedly different from those inherent in other kinds of traditional property rights. This is simply due to the manner in which information costs and negotiation and bargaining costs impact on the patent market. It is for this reason that scholars have doubted the possibility of an emergence of a patent market in the neoclassical economic understanding of the concept.

The information costs problems reflected in the indeterminacies of patent scope, validity and notice of existence combined with negotiation and bargaining difficulties make it difficult for contractual negotiations over patents to be smoothly conducted. As argued in this chapter, where contractual agreements cannot be easily reached, the problems of asset specificity would result, which in turn create room for opportunism. While transaction cost factors (information, and negotiation and bargaining costs) are theoretically considered amenable to facilities such as ‘creation of a firm’ and government intervention measures, in practice they are not. This is primarily because of the patent thicket risk. However, the tenet of this thesis is that if incentives for opportunism were removed, regardless of illiquidity problems arising, the patent market would still work efficiently. At this point therefore, it becomes necessary to shift focus to the nature
of patent opportunism and its implications for the patent market, which is the subject of discussion in Chapter 3.
Chapter 3
Legal Remedies and Opportunism in the Patent Market

3.1 Introduction

The preceding chapters have revealed that transaction costs have the capacity to make the patent market unworkable, such as to dampen efficient \textit{ex-ante} bargaining between patentees and users of patented technologies. Further, it has been established that, in this event, a state of patent market failure would likely arise on account of the fact that an attainment of the social welfare objectives of the patent system becomes impeded. It is amidst a state of market failure that room for opportunism is created, which makes it more difficult to attain the welfare objectives of the patent system. As this chapter reveals, opportunism results from the likelihood of undue advantage(s) being obtainable by a party because of the nature of legal remedies applied patents.

Patent opportunism can occur \textit{ex ante}, in the course of bargaining, or \textit{ex-post}, when asset specificity arises.\footnote{See Dan Burk, 'Intellectual Property and the Firm' (2004) 71 \textit{University of Chicago Law Review} 3.} When opportunism occurs \textit{ex ante}, the opportunistic party, who might be the resource owner or intending user,\footnote{See Colleen Chien, 'Holding up and Holding Out' (2014) 21 \textit{Michigan Telecommunications and Technology Law Review} 1.} employs strategic behaviour to procure a bargaining edge so as to make an undue gain.\footnote{Oliver Williamson, 'Opportunism and Its Critics' (1993) 14 \textit{Managerial and Decision Economics} 97.} \textit{Ex-post} opportunism is more common. An unauthorised user of patented technology, when adjudged an infringer, is exposed to \textit{ex-post} opportunism when asset specificity results from transaction costs, particularly information costs, and the
patentee is able to secure undue remedial rewards. Conversely, a patentee is exposed to ex-post opportunism when resources are expended towards securing a patent and likely remedial outcomes of patent enforcement are unlikely to reflect the patentee’s expectations of economic returns. In other words, there is a likelihood that opportunistic behaviour could arise from either the patentee or infringer. Where the patentee engages in opportunism, either ex-ante or ex-post, it is known commonly as ‘patent holdup’ or patentee opportunism. Where, however, the infringer engages in it, it is commonly described as ‘reverse holdup’ or infringer opportunism.

This chapter targets the problem of ‘patent holdup’, as it is the most common form of patent opportunism. It is fed by generous judicial attitudes towards the award of prohibitory remedies in the enforcement of patents, chief among which is the injunction. The argument is that where prohibitory remedies are routinely awarded, a propensity for patentee opportunism will ensue.

As this chapter deals specifically with patentee opportunism, it cannot avoid addressing the arguments put by a number of prominent scholars in the field for the default application of monetary remedies in substitution for prohibitory remedies. The ultimate aim of this chapter is to show the incompleteness of the portrayal of patentee opportunism presented by these scholars. In pursuing this objective, this chapter is divided into four parts. Part I aims at demonstrating that patent remedies shape the patent market by acting as ‘prices’. It includes an analysis of the ex-post patent market, which reveals that where patent bargains fail ex ante, such bargains are likely to occur ex post, in the shadow of litigation. It is especially during this ex post bargaining process that the role of patent remedies

---

comes to the fore. Essentially, remedies act as prices that influence the ordering of the bargaining process. However, their role as prices extends to the *ex-ante* bargaining process and general attitudes towards patents. This is because the economic and legal value of patent remedies enables patentees to make legitimate demands from other negotiating parties.\(^8\)

**Part II** introduces arguments for the default application of monetary remedies in patent law, and the theoretical foundations that inform this call. **Part III** evaluates the strength of these arguments against the alternative of applying prohibitory remedies in the enforcement of patents. It does so by highlighting the social costs likely to result from patentee opportunism. Finally, **Part IV** advocates a more substantive definition of opportunism, which makes it clear that monetary remedies could themselves be a source of patentee opportunism, even absent prohibitory remedies.

### 3.2 Part I: Ex-post Patent Market

As established in the previous chapters, as transaction costs may dampen the prospect of *ex ante* bargains between patentees and users of patented technologies, private deals are frequently concluded *ex post*, after alleged infringement, to counter the lost opportunity or inability to bargain *ex ante*. The US Federal Trade Commission (FTC) has described this *ex post* bargaining process, which is increasing in popularity, as the ‘Evolving IP Marketplace’.\(^9\) An *ex post* patent market is necessitated by the exigencies of patent enforcement litigation or the threat to enforce patent(s).\(^10\) To avoid legal sanctions for infringement, the

---


infringer may do a number of things: agree to pay a mutually determined consideration in return for continued use of the patented invention; contest the patent’s validity; argue non-infringement; counterclaim that the patentee also infringes their own patent; or raise possible defences. It is against the backdrop of these possibilities that ex post bargains are often reached.

Various salient observations have been made concerning patent litigation. It has been described as the ‘Sport of Kings’ to depict the intensity of stakes involved, the height of uncertainty occasioned, and the huge financial costs and lost economic opportunities that can characterise it. Empirical studies show that only a fragment of patents-in-force are actually litigated. In Australia, studies conducted by Weatherall and Jensen and Weatherall and Rotstein reveal that while information on patent infringement and enforcement is patchy, only a few patent owners take steps to enforce patents. Very similar outcomes resulted from empirical studies in the UK by Helmers and McDonagh, and in the USA by Allison et al. A recent study by Weatherall and Dent highlights factors that determine whether a dispute is likely to end in litigation as opposed to settlement. These

---

factors also determine the bargaining positions of patent holders and infringers in relation to one another,\(^22\) and can be summarised as follows:

1) The weightiest consideration for patentees in settlement is the fear that their patents will be invalidated. Others include: high expected cost of litigation; low expectation of litigation success; perceived financial state of an infringer; patent value; general settlement benefits; counter-claim brought by an ‘infringer’; global litigation/settlement strategies; the size of a patentee’s patent portfolio; an ‘infringer’s’ aggressive litigation reputation; and a low desire to deter others.

2) The weightiest consideration for infringers in settlement is the high expected cost of litigation. Again, though, there are other considerations, such as: importance of the technology to an infringer; a low expected chance of success; the general benefits of settlement; low value of infringing products; low chance of a patent being held invalid; the perceived financial state of a patentee; the quantum of damages claimed; global settlement strategies; any offer of advantageous licensing terms; a patentee’s aggressive litigation reputation; the low value of a patent; and overseas court decisions on the relevant patent(s).\(^23\)

Somaya gives another angle to the discussion on the likelihood of litigating, arguing that the strategic stakes of the parties are determined by their patent possessions.\(^24\) For example, where a product commercialised by a patentee relies on complementary technical inputs, and the patentee requires access to some of the patents owned by the infringer, settlement becomes very likely. He calls this


situation ‘mutual holdup threats’. On the other hand, where the patentee’s patent possession is sufficient and there is no need to license-in from other sources, the likelihood of settlement diminishes, as there is a strong incentive to sue. In any event, the preponderance of the studies reveals that a majority of threatened and initiated litigations end up settling, some before trial and some before judgment. As appears from the foregoing, the ex post bargaining process is a significant and viable alternative to the ex ante route because it makes up for the lost opportunity to bargain ex ante.

Settlement, though, has both advantages and drawbacks. La Belle, adapting Fiss’s arguments against legal settlements in general to patent law, maintains that patent infringement settlements could inflate social costs by bringing about peace between the disputing parties at the expense (of justice) to society. In other words, the opportunity to test the validity of patents might be lost and society may bear the costs of invalid patents. She also notes that settlements deny potential judicial precedents. She takes exception to the fact that most patent settlements occur in secret, and decrives the high likelihood that parties to settlements end up dissatisfied.

While these fears are genuine, they lose appeal in situations where the validity of patents has been judicially determined in prior cases or the case at hand. In this event, public interest fears are dispelled, and only the joint interests of the parties appear to matter. For example, Menkel-Meadow has argued that, barring public

\[\text{\footnotesize 25 Ibid.}\]
\[\text{\footnotesize 26 Ibid.}\]
\[\text{\footnotesize 30 Ibid 405-407.}\]
\[\text{\footnotesize 31 Ibid 407-409.}\]
\[\text{\footnotesize 32 Ibid 409-410.}\]

84
policy concerns, parties should be entitled to settle their disputes out of court. Describing protagonists of litigation as labouring from a syndrome she terms ‘litigation romanticism’, she argues that they are blind to the panoply of values that lie in settlement, such as ‘consent, participation, empowerment, dignity, respect, empathy and emotional catharsis, privacy, efficiency, quality solutions, equity and even justice’. It is believed that patent settlements are superior to litigation in most cases. This is essentially because settlements help avoid prohibitive litigation costs and other consequential economic costs. Similar reasoning has been expressed by a number of scholars.

So far, it has been to show how difficulties in concluding ex-ante bargains or transactions between patentees and users of patented inventions do not mean an end to the possibility of bargains or transactions between them. In fact, the parties are often likely to deal ex-post, in the shadow of litigation. This being the case, it becomes necessary to address the principal factor that shapes the ex ante patent market—legal remedies.

3.2.1 Legal Remedies as ‘Prices’ in the Patent Market

Having introduced the ex post patent market, it becomes necessary to focus on the chief condition responsible for its ordering, namely patent law remedies, or the body of rules that fashions or governs the enforcement of patents. As axiomatically expressed by Blackstone:

---

The remedial part of a law is so necessary... that laws must be very vague and imperfect without it. For in vain would rights be declared, in vain directed to the observed, if there were no method of recovering and asserting these rights, when wrongfully withheld or invaded.\textsuperscript{39}

This statement can be best appreciated by viewing enforcement rules, as reasoned by Becker,\textsuperscript{40} as prices that induce the demand and supply of (prohibited) activities.\textsuperscript{41} In other words, the more prohibitive and privately costly the remedies for breach of rights are to a wrongdoer, the more likely he or she will avoid acts that amount to breach.\textsuperscript{42}

The reverse is equally true: the more lenient the sanctions against legal wrongs, the higher the incentive to engage in them, as the private costs of wrongdoing are low. Generally, as persons are rational economic actors, they will only take part in (wrongful) acts to the extent that it benefits them and they can afford to bear the legal consequences.\textsuperscript{43} Accordingly, the nature of remedies that reinforce rights determines the price that the rights owner can demand in settlement arrangements. In other words, remedies determine the bargaining power of both the rights owner and the wrongdoer in private settlement of legal disputes.\textsuperscript{44}

The \textit{ex post} patent market, like all settlement arrangements, is conducted in the shadow of the law. An interesting statement from Mnookin and Kornhauser’s work on settlement in divorce cases captures this point:\textsuperscript{45}

\begin{itemize}
\item \textsuperscript{39}William Blackstone, \textit{Commentaries on the Laws of England: In Four Books} (A. Strahan and W. Woodfall, 11\textsuperscript{th} edition, 1791) 55-56.
\item \textsuperscript{40}Gary Becker, 'Crime and Punishment: An Economic Approach' (1968) 76 \textit{Journal of Political Economy} 169-217.
\item \textsuperscript{44}Jeffrey Rachlinski and Forest Jourden, 'Remedies and the Psychology of Ownership' (1998) 51 \textit{Vanderbilt Law Review} 1541-1580.
\item \textsuperscript{45}Robert Mnookin and Lewis Kornhauser 'Bargaining in the Shadow of the Law: The Case of Divorce' (1979) 88 \textit{The Yale Law Journal} 950-997.
\end{itemize}
Divorcing parents do not bargain over the division of family wealth and custodial prerogatives in vacuum; they bargain in the shadow of the law. The legal rules governing alimony, child support, marital property, and custody give each parent certain claims based on what each would get if the case went to trial. In other words, the outcome that the law will impose if no agreement is reached gives each parent certain bargaining chips ...

Patent settlements are likewise modelled on likely judicial outcomes, as foreshadowed by the design of legal remedies, should disputes end in court. The extent to which remedial rules protect a rights owner determines the degree to which the disputing parties can actualise strategic behaviour so as to advance individual gains in the settlement process. As posited by Cooter et al, what causes failure in settlements is not the excessive optimism of a party as to his or her chances of success in litigation, but the inordinate self-pursuit for advantageous distribution of gains from the settlement. However, prudent parties will try to settle on terms that make the eventual distribution of gains favourable to each side in such a way that leaves each better off than were negotiations to fail. In *ex-post* bargaining it is the possibility of strategic behaviour in the bargaining process that creates room for opportunism on either the patentee or infringer’s side.

As noted above, legal remedies do not merely inform *ex post* bargaining outcomes, but also influence *ex ante* bargains and the general attitude of entities towards legal entitlements. This corroborates Wright’s position that ‘*[t]he law of remedies is, inevitably, a social institution*’ because ‘*[d]ecisions on remedies questions have an inescapable effect upon social order’. The implications of remedies on the decision-making of individuals can be explained from an economics perspective. Mathis expresses the view that rational persons, seeking

---

46 Ibid, 968.
48 Ibid.
51 Ibid.
to advance their interests, do so in the light of two factors: preferences and constraints. 52 ‘Preferences’ can be described as motivations to engage in something, whereas ‘constraints’ are external limitations or incentives. 53 Economic actors, Mathis maintains, pursue their preferences to the extent that external limitations or incentives allow them.54 Thus, legal remedies are akin to the constraints referred to in Mathis’ analysis. Accordingly, it is necessary to understand the nature of legal remedies in patent law to lay a foundation for understanding patent opportunism.

3.2.2 Rationale for and Choice of Legal Remedies in Law

Calabresi and Melamed theorise that regimes of remedies are largely of three kinds: property rules; liability rules; and inalienability rules.55 Property rules are essentially founded on a ‘bargaining’ model in that they protect entitlements whose value are considered by the state to be best left to the subjective evaluation of the owner.56 This is so that an intending user would have to bargain with the owner, without the state intervening to ease bargaining difficulties between the parties.57 This bargaining model is usually aimed at vindicating the exclusive nature of an entitlement.58 Injunctive relief, and punitive or exemplary damages, are the main forms of the bargaining model to secure the exclusivity of an entitlement.59

53 Ibid.
54 Ibid.
Henceforth, prohibitory remedies will be referred to as the ‘bargaining model’ of remedies.

Liability rules are cardinally an administrative model of enforcement in that they apply where the state considers that it is possible to make an objective assessment of the value of an entitlement. This is because where they apply it is considered imperative that such assessment be made so that a third party can interfere with the entitlement with judicial permission. But sanctioning that interference would be conditional upon the third party being able to pay the price of the entitlement as assessed by the State. The State may allow liability rules to direct the enforcement of entitlements for economic efficiency, distributional, or other justice reasons just as the State chooses to bestow an entitlement. Where economic efficiency informs the court’s choice of liability rules, it could be that the courts try to avoid a beneficial outcome going awry because of transaction costs. Or it may be considered that such a beneficial outcome would be costlier if left to the exclusive control of its owner, where applying liability rules would have secured the same outcome at a lower cost. As per distributive considerations, where the State considers that leaving entitlements to the exclusive governance of owners would tilt the balance of gains unduly in favour of the rights owner, liability rules would be applied.

Henceforth, monetary remedies will be referred to as the ‘administrative model’ of remedies.

Generally, in the context of tort obligations and property rights, liability rules target monetary remedies such as compensatory damages, damages in lieu of an injunction and restitutionary damages. Although superfluous to this discussion,

---

61 Ibid, 1106.
62 Ibid, 1107.
63 Ibid, 1106-1107.
inalienability rules, the third class of remedies, apply where the state prohibits transactions over a given entitlement, but then considers itself the best assessor of its objective value in circumstances where that entitlement has been interfered with. Entitlements are usually anchored by inalienability rules for reasons of public policy and morality. For example, where public policy disallows the sale or ex ante negotiations on human body parts or fundamental human rights, the law adopts rules of inalienability.

In reality, entitlements are protected via a combination of remedial rules, that is, a combination of bargaining and administrative rules remedies. This makes it possible for a patentee's right to exclude others from access to his or her patent entitlement, by the application of bargaining remedies, to be converted to monetary compensation where considered necessary.

The essence of Calabresi and Melamed's thesis on legal remedies can be encapsulated in two major propositions:

1) whenever transaction costs are low, it is better to apply the bargaining model of remedies. This is because here parties reach agreement easily and so bargaining efficiency as between the parties is attained at low cost. However, such bargaining efficiency may come with some serious

---

69 Ibid.
implications for distribution of gains from the transaction between the parties, especially for the user of the entitlement. It may come also with the imposition of negative externalities on third parties (e.g. excessive prices of goods or services, or even scarcity of same).

2) whenever transaction costs are high, it is better for both efficiency and distributional reasons to use the administrative model of remedies to protect entitlements. This is because leaving the exchange to the autonomous decisions of the parties is not only likely to result in inordinate private and social costs, but might in fact result in a situation where resources are left to waste. In such circumstances, it is best if the State intervenes to order sale at an administratively or judicially determined price.

3.2.3 The Nature of Patent Law Remedies in Australia, Canada, the UK and the USA

The preceding section provided a brief précis on the (economic) rationale behind the choice of remedies, as a backdrop to addressing patent enforcement policy debate. At this stage it becomes necessary to highlight the nature of legal remedies applied towards patents in the jurisdictions of focus in this thesis (i.e. Australia, Canada, the UK and the USA) and to identify where these remedies fall along the bargaining/administrative rules divide. Patent legislation in Australia, Canada and the UK provides a range of prohibitory and monetary remedies for the protection of property rights, particularly injunctions, damages and disgorgement (otherwise known as an account of profits). In the USA the disgorgement remedy does not exist.

The bargaining model of remedies include preliminary and final injunctions, and in some cases quia timet injunctions, orders for delivery up and destruction, and

---

72 Ibid.
74 Section 122 of the Australian Patents Act 1990 (Cth).
76 Section 61 Patents Act 1977 (UK).
punitive/additional damages, amongst other measures.78 These relate to the proprietary nature of patents because they seek to ensure the patentee’s exclusive enjoyment of the right is protected from the interference of third parties. In the Commonwealth jurisdictions, monetary relief chiefly targets an election between compensatory damages and disgorgement, the election being necessary to avoid double or excessive recovery.79 Compensatory damages aim to place patentees in the position they would have occupied, as far as money can, had the infringement not occurred.80 Disgorgement seeks to confiscate from the infringer all of the gains attributable to the infringement of the patent.81 Thus patentees seeking monetary relief would pursue compensatory damages when losses arising from the infringement exceed the profits earned by the infringer (i.e. patentee’s losses are greater than the infringer’s profits). However, a patentee would consider disgorgement more advantageous if his or her losses are less than the gains or profits derived by the infringer (i.e. patentee’s losses are less than the infringer’s profits).

A further remedy is the monetary remedy known as ‘reasonable royalties’. This remedy is an offshoot of compensatory damages and is usually granted in three main scenarios. One is where the patentee is a producer and is unable to prove that the infringement caused a loss of profits.82 A second is where the patent holder does not exploit the patent through production, and is consequently unable to claim lost profits or demand disgorgement.83 A third scenario is where the court refuses an injunction but grants damages in lieu thereof.84

83 Ibid.
84 Ibid.
These patent remedies can be classed along the bargaining/administrative models divide. It is clear that prohibitory patent relief and punitive damages fall under the bargaining model, as they seek to preserve the exclusiveness of patent rights. Compensatory damages typically fall under the administrative model as a substitutive of patent exclusivity. This monetary remedy tries to mimic the economic state of affairs that would have existed had patent exclusivity been respected. The disgorgement remedy is, however, difficult to classify because it bears characteristics of both the bargaining and administrative models. It borders on the bargaining model in being aimed at ensuring exclusivity in ownership by depriving the infringer of whatever gains were derived from the infringement. But it also has hallmarks of the administrative model as it is neither able to prevent an impending infringement nor stop the continuation of an infringement, but is only capable of confiscating profits from infringement. It is nonetheless taken in this thesis that disgorgement falls under the administrative model of remedies because it involves a monetary remedy awarded post-infringement and does not function to forestall infringement as prohibitory remedies traditionally do.

Reasonable royalties follow the administrative model, representing a court-assessed price for using patents. This can be better understood in respect of the three scenarios identified above to which reasonable royalties apply. In the first, where the patent holder cannot prove lost profits, it is considered that the patent is not the driving factor of market demand; therefore the infringement has caused no loss of profits that patent exclusiveness would have generated. In the second

86 Anthony Kronman, 'Specific Performance' (1978) 45 University of Chicago Law Review 382 ('It is hard to categorize the trust remedy as either a liability rule or a property rule').
90 This is discussed in Chapter 4 that deals exclusively with Compensatory Damages.
scenario the patent holder would not have been able to make the profits made by
the infringer, as he or she lacked the wherewithal (e.g. production plant and
distribution channel) to do so.\textsuperscript{91} Here it can be reasoned the patent holder should
be entitled to a just share of the profits the infringer made from the (mis)use of
the patent. The first two scenarios relate to past infringements, whereas the third
targets the future. It involves a refusal to grant an injunction, and thus permit an
infringement on the condition that the infringer is able to pay the court-
determined royalties.\textsuperscript{92} These three scenarios correspond with the efficiency and
distributional grounds Calabresi and Melamed theorise are triggers for the use of
the administrative model.

\textbf{3.2.4 Encapsulation of Part I}

\textbf{Part I} has shown how the patent market works in the shadow of litigation, despite
its inherent illiquidity. Legal remedies that reinforce patents direct the ordering
of the \textit{ex-post} market. Legal remedies applied towards patents not only shape
bargaining arrangements, but inform how parties would arrange their affairs and
investments \textit{ex ante} in view of what is expectable \textit{ex-post} as a result of likely
outcomes of patent remedies.\textsuperscript{93} Therefore, the functioning of the patent system
and likelihood of attaining its social welfare objectives rest, to a large extent, on
patent remedies. It is for this reason that some scholars and stakeholders have
averted their focus to this area of patent law, with a view to finding solutions to
the dysfunctional state of the patent market and its consequences for the
efficiency of the patent system, particularly opportunism.

\begin{footnotesize}
\textsuperscript{91}Daralyn Durie and Mark Lemley, ‘A Structured Approach to Calculating Reasonable Royalties’
(2010) 14 Lewis and Clark Law Review 626; see also Dan McManus, ‘Incentives Must Change:
Addressing the Unpredictability of Reasonable Royalty Damages’ (2013) 5 Intellectual Property
Brief 6.
\textsuperscript{92} See Stephen Ullmer, ‘Paice Yourselves: A Basic Framework for Ongoing Royalty Determinations
Royalty: What Remedy Should a Patent Holder Receive When a Permanent Injunction is Denied’
\textsuperscript{93} Lucian Bebchuck, ‘Property Rules and Liability Rules: The Ex Ante View of the Cathedral’
\end{footnotesize}
This focus on patent remedies as a feature of the patent market is largely characterised by the call for the adoption of the administrative model of remedies as the standard scheme of remedies. It is to this call that focus now turns.

3.3 Part II: The Case for the Administrative Model of Remedies

Proponents of the administrative model argue that the application of the bargaining model will encourage strategic and inordinate self-seeking behaviour among patentees, and thus help them secure undue rewards from their patent rights in both ex ante and ex post bargaining. The chief social costs of applying the bargaining model to patents are identified as drawbacks in dynamic efficiency and an extortion or undue gains effect. These effects are usually confluent.

Deeporter takes the position that the bargaining model can frustrate the patent marketplace and thus impact negatively on optimal advancement in innovation. This is especially so, he argues, considering the overlap and complementarity of technologies in modern times, which might require the amalgamation of patented technologies to bring about needed products and services.

In similar vein, a party who has already invested in the production of goods that contain patented inputs, or has arranged manufacturing arrangements around patented technologies, might find it very difficult to redirect towards non-infringing alternatives (NIAs). Thus where an injunction is issued, or is likely to issue, against such a party, that party is highly likely to heed the extortionate demands of the patentee to the extent that he or she can afford. The same goes for the fear of punitive damages, and also compensatory damages and reasonable

---

royalties in circumstances of holdup. To properly address the application of liability rules to patents, a deeper inquiry into the theoretical basis of the administrative model is necessary.

3.3.1 Theoretical Foundations of the Administrative Model of Remedies in Patent Law

Insight into the need for the administrative model can be gained through the concept of the Pareto criterion. It has two components: ‘Pareto superiority’ and ‘Pareto optimality’.97 Pareto superiority relates to a situation whereby bargaining parties can exchange values such that the outcomes of an eventual transaction leave transacting parties in a state better than prior to the transaction.98 To the parties, state B (after the transaction) is said to be Pareto superior to state A (before the transaction).99

This can be illustrated using a patent based example. If Firm XYZ, a phone manufacturer, is able to secure a licence to use a patented component essential to the information storage aspect of its phone brand, that licence places Firm XYZ in a position Pareto superior to its pre-licence state. As a result of a reward or consideration earned in return for allowing the licence, it also places the licensor in a Pareto superior position over its pre-licence state—because the licensor earns an income it otherwise not have earned had there been no deal. In sum, both parties gain by transacting rather than not transacting. One party, though, is likely to gain more than the other notwithstanding the mutual gains.100 It could be either party, but in an *ex post* bargaining process where the licensee is caught in asset specificity, it is likely that the licensor will secure greater gain because of the

irreversible costs of the licensee. Thus, Pareto superiority can be occasioned by
distributive disparity between transacting parties; one party being placed at a
distributive advantage over the other.

This leaves Pareto optimality. A state is Pareto optimal if the parties cannot move
from state A to state B without leaving one party worse off.\(^\text{101}\) Again using a patent
example, let it be assumed that Firm XYZ competes with Firm BCD in the mobile
phone manufacturing market. Firm BCD has a patent on the latest and most
efficient technology on phone data storage, which drives market demand for
phones. Let it be further assumed that the patented technology creates a niche
phone market. Firm BCD believes it is only able to meet 80% of that market’s
demand, considering its operational resources and targets. Were Firm BCD to
licence Firm XYZ to exploit the patent, such that XYZ could serve the remaining
20% market demand, Pareto superiority would be attained because both parties
would be better off with such a bargain than without it. But such Pareto
superiority could leave potential economic gains unharnessed as XYZ’s
exploitation of the market beyond a 20% market share might yield additional
favourable gains to both XYZ and BCD, and to consumers. This may in turn be
corrected by extending the licensing contract further, such as to enable XYZ to
control 30% or more of the market. However, any additional licensing beyond the
point originally agreed might abrade BCD’s market share and thus leave BCD
worse off and XYZ better off.

A state before such an extension in licensing would be Pareto optimal. Similarly,
Pareto optimality is preserved where BCD refuses to license XYZ in any way
whatsoever, so that only BCD is able to exploit the patent exclusively. This is
because things are left as they were \textit{ab initio}, with no party being left better or
worse off than previously the case. In favouring maintenance of the \textit{status quo},
Pareto optimality suffers the drawback of being intolerant of change\(^\text{102}\) (known as

---

\(^{101}\) See Allan Feldman, ‘Bilateral Trading Processes, Pairwise Optimality, and Pareto Optimality’

\(^{102}\) James Buchanan, ‘The Relevance of Pareto Optimality’ (1962) 6 Journal of Conflict Resolution
341-354; see also James Buchanan and Warren Samuels, ‘On Some Fundamental Issues in
Political Economy: An Exchange of Correspondence’ (1975) 9 Journal of Economic Issues 15.
the initial endowment problem). This can result in allocative inefficiency, which is the inefficient allocation of resources needed for production; and dynamic inefficiency, which essentially relates to retardation in innovation or inventive outcomes.

The Pareto criterion, being founded on autonomy, assesses economic welfare from the standpoint of parties involved in a given transaction. Veljanovski asserts that the Pareto optimality is based on three pillars:

1) individuals can best decide what is good for them (i.e. consent or unanimity);

2) the welfare of all members of society is of primary importance without discrimination; and

3) (therefore) any improvement in the conditions of one party, which leaves another worse off, fails to improve the sum of welfare of members of a society.

Thus, the Pareto criterion would rather preserve status quo than let one member of society or a party to a transaction be worse off. Posner describes the Pareto optimality as hanging on Kantian foundations, as it is rigidly focused on autonomy and fairness to all parties. However, like many other scholars, Posner

---


criticises it for being static, as it does not inquire into the implications of a set of arrangements on third parties or connected matters. Thus Posner writes:

If A sells a tomato to B for $2 and no one else is affected by the transaction, we can be sure that the utility to A of $2 is greater than the utility of the tomato to A, and vice versa for B, even though we do not know how much A’s and B’s utility has been increased by the transaction. But because the crucial assumption in this example, the absence of third-party effects, is not satisfied with regard to classes of transactions, the Pareto-superiority criterion is useless for most policy questions. For example, if the question is not whether, given a free market in tomatoes, A’s sale to B is a Pareto-superior change, but whether a free market in tomatoes is Pareto superior to a market in which there is a ceiling on the price of tomatoes, the concept of Pareto superiority is unhelpful. The price ceiling will result in a lower market price, a lower quantity produced, lower rents to land specialized to the growing of tomatoes, and other differences from the results of a free market in tomatoes. It would be impossible to identify, let alone negotiate for the consent of, everyone affected by a move from a price-regulated to a free tomato-market, so the criterion of Pareto superiority cannot be satisfied.\textsuperscript{110}

Posner is suggesting that the Pareto criterion ignores ‘third party’ effects (i.e. dynamic implications) by focusing inordinately on the transaction between parties A and B, and the mutual gains to both parties.\textsuperscript{111} However, taking into account the implications of the transaction on third parties may reveal that the transaction is socially costly, and is apt to dampen the social wealth (potentials) of a society, especially when there are cheaper alternative routes to achieving a potential exchange between A and B.\textsuperscript{112}

\textsuperscript{111} Ibid.
\textsuperscript{112} Ibid.
Expressed another way, the distributive inequality (likely to result from Pareto superiority) and the inefficiency effects (likely to arise from Pareto optimality) could have adverse effects on other matters of equal or greater social importance. A patent-based example can be used to illustrate this point. If A licensed a patent to B at exorbitant royalty rates, although there is Pareto superiority, funds and resources that B could have directed towards other socially beneficial outcomes, such as research and development, would go to A. This would be costly to B and to members of society. If A refused to license B, even though Pareto optimality is attained, inefficiencies in allocative and dynamic sense on a social scale might result. This would also be privately costly to B, in that B would be unable to access resources needed for B's advancement, as well as being socially costly in allocative and dynamic efficiency terms to third parties.

Posner considers the Kaldor-Hicks criterion of efficiency to be superior to the Pareto criterion.113 The Kaldor-Hicks criterion states that movement from a state A to a state B is superior, even when one party is made worse off by that transition, so long as the party benefited is able to compensate the party made worse off.114 In other words, instead of focusing on the avoidance of one party being made worse off (as Pareto optimality affixes upon), the Kaldor-Hicks criterion is concerned about advancing social wealth115 that might be lost or diminished were transactions left to the parties' autonomous decisions.116 The Kaldor-Hicks criterion is known as ‘potential’ Pareto superiority because it only assumes the possibility of the gainers compensating the losers, even though there might be no

actual compensation or insufficient compensation. The possibility of compensation, it is said, ultimately makes all parties better off and thus helps attain Pareto superiority effects. In essence Kaldor-Hicks criterion hinges on utilitarianism.

The administrative model of remedies abuts Kaldor-Hicks efficiency reasoning. Posner explains that the courts stand in good stead for the parties, and have the capacity to mimic the market and assess a socially optimal price. He reasons as follows:

I believe that in many cases a court can make a reasonably accurate guess as to the allocation of resources that would maximize wealth. Since, however, the determination of wealth made by a court is less accurate than that made by a market, a hypothetical-market approach should be reserved for cases where transaction costs preclude use of an actual market to allocate resources efficiently.

Hypothetical-market analysis plays an important role in the economic analysis of common law. Much of that law seems designed, consciously or not, to allocate resources as actual markets would, in circumstances where the costs of the market transactions are so high that the market is not a feasible method of allocation.

As can be deciphered from Posner’s words, liability rules are interventionist curial measures in cases where transactions are difficult to achieve or would have inefficiency implications. Proponents of the application of liability rules to

121 See Ian Ayres and Eric Talley, 'Distinguishing between Consensual and Nonconsensual Advantages of Liability Rules' (1995) 105 Yale Law Journal, 236 at footnote 3; see also Keith...
patents, such as Posner, consider patents to be economic rights meant for the sole purpose of incentivising innovation and not necessarily about enriching inventors. Therefore, they believe that patents should receive legal protection only to the extent that this advances social wealth.

3.3.2 Encapsulation of Part II

As can be gleaned from the foregoing theoretical analysis, the key concern that has given force to the call for the default application of the administrative model of remedies is access to patents. As maintained by proponents of this approach, difficulty in accessing patents is a problem created or exacerbated by the routine or default application of the bargaining model of remedies to patents. This problem has the tendency to result in both allocative inefficiency and dynamic inefficiency. It also results in distributive disparities between patentees and users of patented inventions, usually in the favour of the former.

This is why its proponents argue that problems in accessing patents can be corrected by a substitution of the bargaining model of remedies with the administrative model. It is their argument that since the administrative model enables better access to patents, allocative and dynamic inefficiencies, and distributive disparities can be reduced.

Considering that the theoretical analysis undertaken thus far does not present a vivid picture of the practical problems that can result from the bargaining model, it is necessary to further highlight the social costs of applying that model as identified by proponents for the administrative model. This is with a view to justifying the theoretical foundations of the administrative model presented.

3.4 Part III: The Social Costs of the Bargaining Model as Identified by Proponents of the Administrative Model

As already identified, the case for the default application of the administrative model to patent enforcement is informed by an attempt to avert patent opportunism, as it replaces patent exclusivity with judicially assessed payments. To reiterate, opportunism is the quintessence of market failure, and it commonly manifests in the form of 'holdup'. Also as noted earlier, holdup is likely to arise whenever a party, due to transaction costs especially information costs, is caught in asset specificity and becomes vulnerable to opportunism by a right or resource owner. Some advocates for the application of the administrative model do not consider this traditional description of holdup as exhaustive. Farrell et al, for example, explain that 'pure economics is largely unaffected by whether or not there is guile involved' (as a result of information costs) but whether the patent holder fails 'to license in a reasonable fashion'. In other words, 'opportunism or holdup arises when a gap between economic commitments and subsequent commercial negotiations enables one party to capture part of the fruits of another's investment'. This holdup results in allocative inefficiency (also known as static inefficiency) and dynamic inefficiency. Cotter confirms this, reasoning that a holdup occurs where the patent holder causes:

a) 'static deadweight losses that are not justified by likely increases in dynamic efficiency', that is, the patent holder is able to cause the infringer...
to be at a distributional disadvantage, which misaligns with the purposes of patent protection.

b) ‘dynamic efficiency losses due to reduction in incentive ... to engage in follow-up innovation’.\textsuperscript{128}

Lemley and Weiser state that patent holdup relates to the patentee’s ability to extract economic reward, bearing not on the technological cum economic value of the patent, but on the unfortunate circumstance of a party having already incurred costs in its exploitation, or in dire need of gaining access to exploit it.\textsuperscript{129} Several scholars who favour applying the administrative model to patents support this position.\textsuperscript{130} From the varying scholarly perceptions of holdup, it becomes clear that patent opportunism can arise \textit{ex-post} due to asset specificity (i.e. classic holdup) and \textit{ex-ante}, due to the refusal to deal except on the condition of undue gains appear in sight (hold-out). Both are considered as alternative forms of economic holdup. This reinforces the view expressed earlier that, although this chapter is directed largely at the \textit{ex post} bargaining process, it is not limited to that process, but generally concerns patent bargaining in the shadow of litigation. For this reason, the significance of the administrative model to both stages of bargaining merits discrete treatment.

3.4.1 Classic Holdup

Classically, ‘holdup’ stems primarily from an infringer’s asset specificity.\textsuperscript{131} The infringer incurs costs or invests irreversibly towards the exploitation of a patent, due to information costs (i.e. difficulties in determining the validity, scope and existence of patents). Thereupon, the patentee is able to take advantage of the

\begin{flushleft}
\textsuperscript{129} Mark Lemley and Phil Weiser, ‘Should Property or Liability Rules Govern Information?’ (2007) 85 \textit{Texas Law Review} 787.
\textsuperscript{131} Mark Lemley, ‘Ten Things to Do about Patent Holdup of Standards (and One Not To)’ (2007) 48 \textit{Boston College Law} 149.
\end{flushleft}
misfortune of the infringer. Standard Essential Patents (SEPs) are the stock-in-trade of the analysis of a classic patent holdup situation.\footnote{Joseph Miller, ‘Standard Setting, Patents, and Access Lock-In: Rand Licensing and the Theory of the Firm’ (2007) 40 Indiana Law Review 351; see also Michael Lindsay and Robert Skitol, ‘New Dimension to the Patent Holdup Saga’ (2013) 28 Antitrust 34-41.} As defined by Yeh, an industry standard ‘is a set of technical specifications that provides a common design for a product or process’.\footnote{Brian Yeh, ‘Availability of Injunctive Relief for Standard Essential Patent Holders’ (2012) Congressional Research Service, 7-5700, available at https://www.law.berkeley.edu/files/CRS_SEP_Report_9-2012.pdf (last accessed 02/05/2016).} Standards enable interoperability or working compatibility among complementary technologies (e.g. software, accessories or devices) such that the need to switch from one technology to another or to have disparate goods may be obviated where they can be fused.\footnote{Research in Motion Inc v Motorola, Inc., 644 F SUPP. 2d 788, 790-91 (N.D. Tex. 2008); see also Apple, Inc. v. Motorola Mobility, Inc., 2011 WL 7324582.} A patent becomes an SEP when it forms part of the several fundamental technologies assembled to form a given standard.\footnote{Josh Lerner and Jean Tirole, ‘Standard-Essential Patents’ (2013) National Bureau of Economic Research, Working Paper 19664, available at http://www.nber.org/papers/w19664 (last accessed 02/05/2016).} Once a standard is adopted and becomes settled industrial practice in a given industry, the patent’s value heightens beyond its usual technological-cum-economic worth had it not been adopted.\footnote{Valerio Torti, ‘IPRs, Competition and Standard Setting: In Search for a Model to Address Hold-Up’ (2012) 33 European Competition Law Review 387.}

Where the value of a patent is circumstantially heightened and it is asserted against a third party or parties who have incurred irreversible costs on it, it becomes a trump card in the hands of a patent holder in the bargaining process. This is largely for two connected ‘lock-in’ reasons: network effects and switching costs. ‘Network effects’ (also known as network externalities) relate to the user base or population of a particular product. According to Bernieri, ‘[n]etwork externalities arise in markets where the use of a product by one
consumer increases the value that other users obtain from the product’.\textsuperscript{138} She provides an example of network externalities:

The illustrative example is communication networks, in which the user can establish contact with other users of the network and adding an additional user increases the value that others might derive. Externalities can occur directly, as in the first case described, or indirectly, when the utility of the users of a network, for instance, an operating system, increases because the developers of application programs will produce more software for the operating systems that are most vastly used.\textsuperscript{139}

Where there is settled use of a technology or an assemblage of technologies, it becomes difficult for users to migrate or divert to a new one.\textsuperscript{140}

‘Switching costs’ refer to the costs of a manufacturer to switch from one technology or an assemblage of technologies, usually patented, to alternative (non-infringing) ones. Switching costs can be inordinate and with significant inefficiency implications.\textsuperscript{141} Magliocca provides a vivid similitude of the implications of switching costs for manufacturers: ‘When an integration occurs and the item is being made, the manufacturer cannot remove the given patent from this bundle without redesigning the entire device—it is like pulling a thread from a tapestry’.\textsuperscript{142} To carry this analysis on classic holdup further, it is important to provide practical illustrations; one derived from economic analysis and the other derived from case law.

\textsuperscript{139} Ibid.
i. Lemley and Shapiro’s Economic Analysis of Classic Holdup

Lemley and Shapiro provide the leading economic analysis of patent holdup.\textsuperscript{143} Simplified, they identify six variables as determinative of the holdup bargaining outcomes:

\begin{itemize}
  \item \( V \): relating to the value of the patent, judging by its incremental technical and economic value over other available technologies;
  \item \( M \): representing margin earned per unit product. It can be computed as difference between the price of the product sold and its marginal cost (i.e. incremental cost of making an extra unit of production). Thus, if a phone is sold for $300 and the marginal cost of producing the phone is $50, then the margin earned is $250. The higher the marginal cost, the lower the margin earned per unit;
  \item \( \theta \): depicting the strength of the patent in terms of the probability of it being found valid and infringed or invalid and/or not infringed;
  \item \( C \): standing for an infringing manufacturer’s switching cost (from the infringing technology to a non-infringing alternative);
  \item \( L \): signalling the number of sales that would be lost by an infringing manufacturer should an injunction be granted to forestall further sales; and
  \item \( B \): symbolising the bargaining skills of both the patent holder and the infringer.\textsuperscript{144}
\end{itemize}

Lemley and Shapiro maintain that, upon designating these variables, the fair and standard royalty should be based on \( B \times V \) (bargaining strength on both sides multiplied by the value of the patent).\textsuperscript{145} In other words, parties agree to a royalty depending on the incremental value of the patent over other non-infringing technologies.


\textsuperscript{144} Ibid, 1995-2000.

\textsuperscript{145} Ibid.
alternatives in accordance with their respective bargaining strengths. The likelihood of the patent being found valid and infringed would also factor into the adoption of a royalty rate.

They posit that should the infringer choose to litigate without redesigning the infringing product or production process — ‘the litigate strategy’ — the holdup power of the patentee lies in the adverse consequences to the infringer of having to redesign to avoid infringement in the event that the patent is found valid and infringed. Thus the patentee would be in a position to take advantage of two things: the costs of the infringer in having to redesign the infringing matter; and (should an injunction be issued to halt further infringement) the lost sales during the course of redesigning. To avoid losing not just the value attributable to the infringing aspects of the product or process, but the entirety of the product or process, the infringer may agree to a royalty rate that exceeds the value of the patented aspects of the product or process, but not a sum that exceeds the profitable earnings on the non-infringing aspects.

However, were the infringer to litigate, and to seek alongside to redesign — the ‘redesign and litigate strategy’ — the holdup effects would be substantially the same, but different in nature. Thus, where the patent is valid and infringed, and the cost of redesigning is somewhat low, the infringer would still stand to lose from being excluded from the market pending the completion of a non-infringing redesign. This is where the holdup would lie. Where, conversely, the cost of redesigning is high, then both the costs of redesigning and being excluded from the market factor into the holdup. The infringer would therefore be willing to pay a sum higher than the market value of the patent, but not one exceeding the cost

146 Ibid.
147 Ibid, 2001
150 Ibid, 2002
of being excluded from the market (in the first case) or the costs of being excluded from the market and to redesign (in the second case).152

The ‘rip off’ or overcharge value of such a patent magnifies in correspondence to the weakness of the patent’s strength. Accordingly, the likelihood of the patent being found invalid and/or non-infringed would also account for the extent to which the infringer is being ‘extorted’. Where the patent has no incremental value over other non-infringing alternatives, the holdup effect lies in the costs avoided in having to redesign already produced items (i.e. extra fixed costs) and being excluded from the market pending that redesign.153 This is so with respect to both the litigate and the redesign and litigate strategies.

Lemley and Shapiro surmise that holdup effects remain possible even were an injunction out of the picture, say, where the infringement suit is pursued after patent expiry or where an injunction would not issue for equitable or efficiency reasons. In these circumstances, they argue, reasonable royalties are likely to be awarded on account of what would have been the patentee’s bargaining position had an injunction been granted, explaining the point as follows:

The consequence of this circularity is that reasonable royalties are elevated above the benchmark level, and the problems of holdup identified earlier “infect” the court-awarded level of reasonable royalties. Since negotiated royalties reflect a premium based on holdup, so will the reasonable royalties awarded by the court. And this in turn gives patent holders more negotiating power in a self-reinforcing manner, which ultimately magnifies the effects of holdup on negotiated royalty rates.154

The purpose of presenting Lemley and Shapiro’s economic analysis of patent holdup is to illustrate how classic holdup occurs. The case law is now turned to in order to illustrate this point further.

152 Ibid.
153 Ibid.
ii. Case Law Instances of Holdup

Proponents of the administrative model have a number of real-life examples of holdup situations to buttress their economic assumptions or theory. One example is the infamous *RIM v NTP* case, where the defendant felt pressured, by the fear of an injunction being issued against it, to enter into a $612.5 million settlement with the patentee. Eventually it transpired that the patent upon which the settlement was entered was invalid, but this realisation only came to light after the payment had been made.

Another case of significance that highlights the dangers of indiscriminate application of the bargaining model is *Ebay v MercExchange*. In this case, which was ultimately decided by the US Supreme Court, Justice Kennedy warned against the likelihood of owners of patents with insignificant value using the threat of injunctions to gain undue negotiation or settlement powers to extract excessive licensing fees.

*Rambus* is another well-known case of patent holdup. In this case the patentee earned close to $142.5 million licensing fees from users of its patented technologies because the users had been locked into the use of the patented technologies. However, the users had been locked into the application of the patented technologies because of the patentee's failure to disclose patents over the technologies at the time of forming a technological standard.

3.4.2 Patent Holdout

Holdouts most likely arise where alternative non-infringing technologies available or known to third parties are not as viable as the patented ones in issue.

---

155 *NTP, Inc. v. Research in Motion, Ltd.*, No. 3:01CV767 (E.D. Va. Aug. 5, 2003); see also 418 F.3d 1282 (2005).
Holdouts are more palpable in cases of patent thickets, particularly where there are technologically overlapping or complementary patents, such that products or services cannot be provided completely or satisfactorily without access to vital patents.\textsuperscript{160}

A holdout situation may be fuelled and encouraged by the application of the bargaining model to patents, particularly the use of injunctions.\textsuperscript{161} A third party may fear that, should he or she infringe the patent without a licence, the issuance of an injunction to enforce that patent could exclude it from the market until a redesign is successfully pursued.\textsuperscript{162} This could bring about both Cournot complementarity and royalty stacking problems (discussed in Chapter 2), which involves owners of patents strategically seeking more than other owners of technological inputs in terms of the price for use of their patents.\textsuperscript{163} The almost inevitable consequence is that producers pass these extortionate prices on to consumers, thus resulting in double marginalisation.\textsuperscript{164} It might also result in consumption of good and services being too low, making investments made towards commercialisation difficult to recoup.\textsuperscript{165}

Lemley and Shapiro maintain that \textit{ex ante} negotiations assist a third party little in circumstances where the patent is invalid. If the third party has an alternative technology to rely on or the patent is iron-clad in terms of validity and infringement, they envisage that a fair and standard royalty rate based on $B \times V$ (bargaining skills and the incremental value of the patent) can be reached.\textsuperscript{166} But where this is not so, there is a likelihood of overcharge or ‘rip off’ because the third party pays the price of a valid patent even though it is likely invalid.\textsuperscript{167} This

\textsuperscript{160}Ibid.
\textsuperscript{163}Ibid, 2013
\textsuperscript{164}Ibid
\textsuperscript{166}Ibid, 2002-2005.
\textsuperscript{167}Ibid.
patentee opportunism is not circumscribed to *ex-post* situations after asset specificity has occurred.

Armstrong et al argue that holdout effects are real, and particularly visible in the Smartphone market, as owners of patents on components are able to extortionately tax manufacturers.\textsuperscript{168} Since smartphones are amalgamated devices, in the sense that they are products of the convergence of difference components, manufacturers are exposed to the problems that come with patent thickets.\textsuperscript{169} For this reason, patent holders are able to charge royalties on the basis of a percentage of a phone’s market sales price rather than on the value of the component to the entire phone.\textsuperscript{170} Armstrong et al posit, on the assumption that an average smartphone sells for $400, that on account of how royalties are being charged, about $120 of that price is attributable to patent royalty overcharges. This estimation does not take into account the economic value or costs of cross-licenses agreed for the purposes of enabling the production of the phones.\textsuperscript{171}

Lemley and Shapiro’s study on royalty stacking aligns with Armstrong et al’s, but predates it. Lemley and Shapiro argue that, based on their own studies of 3G Cellular Technology, the surcharge value per phone due to royalties on the 3G Cellular Technology accounts for about 30% of the price of the phone.\textsuperscript{172} They surmise that the plethora of patents that make up the IEEE 802.11 collection of standards that facilitates wireless local area networking, also known as Wi-Fi, possibly runs into thousands of patents.\textsuperscript{173} Reference is also made to other technologies, such as DVD media technologies, which are awash with patents and assailed by stacking problems.\textsuperscript{174}

\begin{flushright}
\footnotesize

\textsuperscript{169} Ibid.

\textsuperscript{170} Ibid.

\textsuperscript{171} Ibid.

\textsuperscript{172} Mark Lemley and Carl Shapiro, ‘Patent Holdup and Royalty Stacking’ (2007) n143, 2025.

\textsuperscript{173} Ibid, 2026-2027.

\textsuperscript{174} Ibid.
\end{flushright}
3.4.3 Encapsulation of Part III

The negative effects of patent holdup, both (classic) holdup and holdout, can result in both static (as between the patentee and infringer) and dynamic (third party effects) terms. In static terms, patent holdup can place infringers at a distributive disadvantage, as the patent holder is able to extract extortionate reward from infringers to the extent that they can pay. It provides an incentive to own patents with a view to monetising them. Patent holdup brings adverse dynamic or third party effects such as: abandonment of projects; disproportionately high prices of goods and services; impediments to open innovation; discouraging independent inventions; duplicative research due to redesigning necessities; suboptimal exploitation of inventions; and the wasteful deployment of human and material resources towards the defence of patent claim assertions. The distributive disadvantage an infringer might suffer, though a static matter, is not without dynamic implications. Thus an infringer might forfeit opportunities for deploying resources towards other innovative ends, which may in turn generate a society loss.

However, while the case advanced by proponents of the administrative model provides clues on the likely social costs of routine application of the bargaining model, it is an incomplete account of patent opportunism. It is incomplete because the administrative model of patent remedies, as it exists, is a product of libertarian philosophy. Even though monetary remedies, which are the province of the

administrative model, may not directly retard access to patents, as do prohibitory remedies, they can have the effect of making access to patents costly. This is because the prevailing scheme of monetary remedies deployed towards patents is fashioned with a view towards fostering exclusivity in the commercialisation of patented goods.

3.5 Part IV: The Deficiencies of the Case for the Administrative Model

As the foregoing discussion reveals, the case for the default application of the administrative model appears forceful. Yet contrary to the arguments of proponents for the administrative model, patentee opportunism is inherent in that model too. The non- attribution of patentee opportunism to the administrative model lies in the flawed standpoint from which opportunism is viewed. Proponents of the administrative model centre their arguments on the relationship between the economic conditions of users of patented technologies and excessive rewards likely to accrue to patentees. However, as established in Chapter 1, pursuant to Smith’s definition of opportunism, provided there is a chasm between patent law policy and patent monetary remedies, opportunism is inherent.\(^{182}\) In order to elucidate this argument further, it is important to refer to the relationship between the social welfare objectives of the patent system and enforcement policy.

The patent system is, again as explained in Chapter 1, conceived and operated on the basis that society must forbear static efficiency, otherwise known as ‘free-market’ practices, to secure dynamic efficiency outcomes through the inventive efforts of inventors and their sponsors.\(^{183}\) Society’s forbearance to encourage inventive outcomes is represented in the legal remedies applied to enforcing

\(^{182}\) See Part I of Chapter 1.

\(^{183}\) Ibid.
patents. This confirms the argument made in Part II of this chapter that legal remedies shape the incentives and constraints of entities in relation to legal entitlements, in this case patents. Society’s forbearance in protecting patent entitlements must be at least commensurate with society’s gains—i.e. marginal social costs equal marginal social benefits (MSC=MSB). Otherwise, the marginal social costs of protecting and enforcing patents would exceed the marginal social benefits of inventions procured, which would make the patent system inefficient. In other words, the utilitarian aims of the patent system would be defeated. For this reason, patent law remedies must not overreach the purposes of patent law policy.

As Opderbeck has argued, there is an ideological gap between patent law policy and enforcement policies as represented in legal remedies, especially monetary remedies. This gap is the real cause of patent opportunism, and not merely the possibility of patent holders securing robust rewards on account of the likelihood of a prohibitory remedy being issued. This gap between patent law policy and enforcement policy is also the reason why Sichelman reasons that the prevailing scheme of monetary remedies applied towards patent enforcement provides excessive incentives to incentivise inventive activities. In furthering Opderbeck and Sichelman’s submission, focus turns in the next part of this thesis to an assessment of remedies under the administrative model, with a view to showing how they foment patent opportunism. In the three chapters that follow, the three monetary remedies of compensatory damages, reasonable royalties and disgorgement will be critically discussed to demonstrate how patent opportunism is aggravated upon their deployment.

3.6 Conclusion

This chapter discussed patent opportunism as a function of deficient conception and deployment of legal remedies towards the enforcement of patents. In pursuing this theme, it focused on two major issues: the role of patent remedies in ordering the patent market, and the call for the default application of the administrative model of remedies in patent law. As regards the role of patent law remedies in directing the ordering of the patent market, it has shown that the more prohibitive are patent remedies, the less likely infringement will be encouraged. Conversely, the more lenient are patent remedies, the greater the motive for infringement. This chapter likens legal remedies and the approach to enforcing them to market prices that inform attitudes within a market. This reasoning translates to the patent marketplace. Patent remedies, depending on how they are applied, have effects on the ordering on the patent market. They can make up for the failure of *ex ante* opportunities, or even render the patent marketplace inefficient in circumstances where they are not optimal.

As regards the case for the default application of the administrative model of remedies to patent law, this call rests on the argument that the bargaining model of remedies impedes access to patents, which in turn brings about sub-optimal social outcomes. Advocates of the administrative model cite allocative inefficiency, the likelihood of extortion and discouragement of innovation as outcomes of the application of the bargaining model. They argue that were monetary remedies the standard remedial facilities available to patent holders, the problem of access to patents would be easily solved and the sub-optimal implications of difficulties in accessing patents would evaporate. This chapter, though, reveals that this default application of monetary remedies may not be an immediate panacea to the problem of patent opportunism until those monetary remedies are of such a nature that they can be reconciled with the utilitarian objectives of patent law. It therefore provides a prelude to a discussion of how the present regime of patent monetary remedies enables opportunism in the ensuing chapters.
Chapter 4
Compensatory Damages and the Patent Market

4.1 Introduction

Chapter 3 established that patent opportunism is a function of the improper application of legal remedies to patent law, and that monetary remedies can be a particular source of the problem. The aim of the current chapter is to critically evaluate the effects of the monetary remedy of compensatory damages, with a view to showing the likely effects of compensatory damages on opportunism and the patent market. This chapter argues that compensatory damages are unsuitable to the patent system and the patent market for two primary reasons:

a) compensatory damages provide legal protection that overreaches the purposes of the patent system, and as such inherently create room for opportunism; and
b) by simply adjusting the rules on the computation of damages, the inherence of opportunism facilitated by the remedy remains uncorrected.

The chapter concludes, accordingly, that compensatory damages should not be applied in patent law cases. This chapter is divided into two parts. Part I addresses the first premise by emphasising the utilitarian foundation of the patent system. It illustrates that a gap exists between the utilitarian nature of the patent system and rules for the enforcement of patents creates room for opportunism. It concludes that the problem with the application of compensatory damages to patent law is that it causes focus to be shifted away from the incremental value of patents such that excessive focus is placed on compensating the patentee for losses resulting from infringement. Part II, in dealing with the second premise above, is the lynchpin of this chapter. It chiefly addresses the primary heads of
compensation recognised in patent law, with the aim of establishing that the quantification of compensatory damages is highly likely to be opportunistic because it is assessed using highly uncertain measures. This is pursued by parsing on the doctrinal chasms between Commonwealth jurisdictions and the US, and the novel nuances in Canadian judicial practice on the subject.

4.2 Part I: The Ideological Gap Between Compensatory Damages and the Purpose of the Patent System — the First Premise

Overcompensation resulting from the award of general damages in patent law has long been a major source of concern and discussion among interested entities in the intellectual property (IP) sphere. While some empirical studies have decried damages as being truly excessive,¹ others have sought to show that damages computations are not excessive as purported, although a few outliers may be.² The bulk of these empirical studies have focused largely on reasonable royalties, with only anecdotal reference to compensatory damages. However, as correctly posited by Opderbeck, these empirical studies have been incapable of ascertaining whether or not damages awards are truly excessive.³ As Opderbeck submits, this can only be determined on the basis of whether the award of monetary reparations conforms to the ideological and theoretic foundations of the patent system.⁴ It should also be noted that while the focus of these commentaries has

³ David Opderbeck, 'Patent Damages Reform and the Shape of Patent Law’ (2009) 89 Boston University Law Review 128, 137 (“The studies that have been conducted to day, as well as the original study presented here, are inconclusive concerning whether damages in patent cases are becoming systematically excessive”).
⁴ Ibid, 137-138.
largely been overcompensation, under-compensation may also be a source of concern.

Both over and under-compensation fall squarely in the realm of patent opportunism. To demonstrate the inevitability of opportunism in the application of compensatory damages to patent law, it becomes imperative to explore the gap between the theoretical and ideological foundation of patent policy and the remedy. As discussed in Chapter 1, a simple economic ‘fear’ serves as the *raison d’être* for the patent system. Essentially, without a legal regime that assists in privatising inventive ideas so that inventors and their sponsors can recoup their investments towards inventive engagements, the incentive to invest in such ends will be low. This fear is founded on the belief that third parties would freely and inexpensively utilise the inventive ideas developed at considerable cost by inventors and their sponsors, thus making it impossible for them (the latter) to recover their costs as they compete freely with the former. This is generally known as ‘free-riding’. Free riding is not necessarily a problem, as it can result in significant positive externalities by providing inexpensive access to inventive ideas. However, it does create a serious problem if it reduces the incentive to pursue such inventive activities.

As also explained in Chapter 1, where the marginal social costs to society (in terms of static efficiency and some dynamic efficiency losses) exceed the marginal social benefits derived in enforcing patents, the patent system’s value diminishes or declines. Thus the patent system will not qualify as optimal where sacrifices

---


made towards the encouragement of inventive activities surpass the gains to society from inventive activities. 10 This is why the continued use of compensatory damages in patent law would continue to result in suboptimal outcomes. And this is chiefly because that remedy ignores the incremental value of patents to society, which is the essence of the patent system, in the pursuit of protecting the patentee’s exclusivity in the commercialisation of goods that incorporate their inventions.

This is clearly an institutional defect, which was identified by a practical and pragmatic English judge of his time,11 Justice Arthur Kekewich, in his first instance determination in Clement Talbot Ltd v Wilson and Another.12 This judgment was given in a compensatory damages decision, which was overturned on appeal. In this case the infringer had imported a car that incorporated infringing accessories. The patentee’s claim was that but for the infringing nature of the car they would have captured that car sale, and as such demanded as compensation the profit they lost. In response to the patentee’s arguments Justice Kekewich had this to say:

They are not the Patentees of the car; they are Patentees only of the accessories, and if I were to accede to their argument I should be giving them, by way of damages, profit, which it seems to me they have not earned by reason of the sale of the accessories, but by reason of the sale of the car. For that I can see no foundation.13

Apparently on account of the utilitarian nature of patent law, His Honour went on to reason that in patent law, the loss suffered by the patentee is measurable by the gain derived by the infringer from infringement. Thus, he went on to say: “I think I must ascertain, in order to determine the damages, what is the difference between the value of such a car as the Defendants have bought with the

12 [1907] RPC 511.
13 Ibid, 515.
accessories and the value of such a car without the accessories.” 14 In essence, His Honour considered that the only fair and equitable reparation for a patentee is a sum commensurate with the value gained by the infringer from the infringed patent as the purpose of patent law is simply to promote inventive activities, commercialization being a secondary consideration.

This point is extended by Lemley, who argues forcefully that the patent system functions to enable inventors and their sponsors to recover their fixed costs and make reasonable profits to encourage them to further invent; its function is not to enrich inventors and their sponsors.15 Inventors should be able to charge a price for the use of their patented inventions, this way they can secure reasonable profits from users of their inventions to cover their costs.16

Sichelman corroborates this position on two major bases.17 The first is that patents are awarded for the utilitarian purpose of optimally encouraging innovation, not to protect property rights in inventions in the libertarian sense. This view reflects the position in major patent law jurisdictions. For example in the Canadian case of Teva Canada Limited v Pfizer,18 the Supreme Court of Canada identified the rationale of the patent system to be the encouragement of the disclosure of inventive ideas.19 In other words, the patent system operates on the basis of a quid pro quo basis rationale that in return for the disclosure of inventive ideas inventors are granted exclusive rights over their inventions. For this reason, patentees are only exclusively entitled to receive rewards that their inventions can fetch either by their stand-alone value or their value as manifested when combined with other non-infringing components.

14 Ibid, 516.
16 Ibid.
18 [2012] 3 SCR 625.
19 Ibid, paragraphs 31-35.
The second reason is linked to the first, that patents should be reinforced with legal remedies that are sufficiently necessary to securing inventions. Sichelman accordingly argues that where the design of patent remedies is such that they are modelled after a libertarian system of protecting property rights, such as tort law, society expends or sacrifices more than is needed to optimally incentivise invention. Applying a libertarian philosophy to patent protection (which may be suitable in the context of traditional property rights) has the implication of deflecting focus from the inventive concept of a patent, which should be the chief object of patent protection, to the effects of infringing goods or services on the patentee’s trade. In the joint determination of two Australian cases, JT International SA v Commonwealth of Australia and British American Tobacco Australasia Limited v The Commonwealth, the instrumentalist nature of patent rights was confirmed by Australia’s apex court, the High Court. In doing so, the court described the nature of IP rights with reference to an excerpt from Cornish et al that:

the terms of the section make it plain that an act of economic policy was intended: the objectives were the encouragement of industry, employment and growth, rather than justice to the ‘inventor’ for his intellectual percipience.

The foregoing analysis is compounded by the fact that empirical economic studies have proven unable to confirm that patents do in fact incentivise inventors or their sponsors to engage in inventive activities or to embark upon commercialisation, as generally touted. The exception is pharmaceutical inventions, which are

---

21 Ibid.
22 See Bernard Chao, ‘Causation and Harm in a Multicomponent World’ (2016) 164 University of Pennsylvania Review Online 61, 67.
23 [2012] HCA 42
considered to be the ‘poster child’ of the patent system for commercialisation purposes.  

This is because the inordinate costs of marketing and meeting regulatory requirements that characterise pharmaceutical inventions give impetus to the need to jealously protect them from infringement.

The unsuitability of compensatory damages to patent law is not, however, circumscribed to the ideological gap between the purpose of the remedy and patent law, as identified above. Its unsuitability also stems from the manner in which the restorative objectives of compensatory damages are attained through the ascertainment of ‘loss’ that informs their quantification.

**4.3 Part II: The Lottery Effects of the Methods of Computing Patent Damages—The Second Premise**

With the ideological gap between compensatory damages and patent law policy inherently creates room for opportunism identified, it is important to show how the application of the remedy, particularly its causation component, exacerbates the propensity for opportunism. The prevailing legal treatment of patent infringement as a statutory tort in common law jurisdictions must first be noted as a prelude to discussing the place of causation in the application of the legal remedy.

Although the modern recognition of compensatory damages as a remedy is largely a product of patent legislation, the rules governing the computation of the remedy are largely fashioned by the courts. Legal methods that govern the computation of patent damages are structurally similar to those that apply to tort law. This is because in both contexts it is first required that causation, the nexus between the injury suffered and the wrong done (the ‘but for’ assessment), be

---


28 Ibid.

29 *Fabio Perini SPA v LPC Group Plc* [2012] RPC 885, 904.

established before moving on to the quantification of loss.\textsuperscript{31} In other words, once causation is established focus is directed to determining the quantum of loss resulting from the wrongdoing.

As explained by Justice Jacob in the first instance determination of \textit{Gerber v Lectra},\textsuperscript{32} causation seeks to determine what happened in the past—i.e. whether the infringement caused the losses the patentee claimed to have suffered.\textsuperscript{33} In proving causation, anything that is ‘more probable than not’ is considered certain.\textsuperscript{34} Once causation is established focus shifts to measuring the loss suffered.\textsuperscript{35} Justice Jacobs explained this as relating to an assessment of futuristic events—i.e. what would have happened in the future if past event had not occurred.\textsuperscript{36}

Customarily claims for patent damages are stated in form of the loss of a chance to make profits as a result of the infringement.\textsuperscript{37} As explained by Harvey, in his classic textbook \textit{McGregor on Damages}, when a case for damages is founded on the loss of a chance, ‘(c)ausation is then established by showing that the claimant has lost the chance and showing this on the balance of probabilities’.\textsuperscript{38} He goes on to identify the steps that guide the court in addressing a loss of a chance claim as follows:

This then makes for three stages in the enquiry: first, it must be ascertained whether loss of a chance is recognised as a head of damage or loss in itself; secondly, it must be shown that on the balance of probabilities the claimant has lost the particular chance; thirdly, the lost chance must be quantified by resort to percentages and proportions.\textsuperscript{39}


\textsuperscript{32} [1995] RPC 383.

\textsuperscript{33} Ibid, 395.

\textsuperscript{34} Ibid.

\textsuperscript{35} Ibid.

\textsuperscript{36} Ibid.

\textsuperscript{37} \textit{Gerber v Lectra} [1997] RPC 443, 459.

\textsuperscript{38} Harvey McGregor, \textit{McGregor on Damages} (Sweet and Maxwell, 19th edition, 2014)

\textsuperscript{39} Ibid, 374.
A translocation of Harvey’s analysis to the patent context means that the patentee must show that the loss for which compensation is sought is recognised as compensable in patent law; that on the balance of probabilities, the infringement caused the loss; and finally, a plausible measure of loss resulting from the infringement is presented. It is the last two components that particularly bring about opportunistic outcomes in patent damages computation. The first simply relates to the recognised heads of compensation. Therefore, that component must be addressed before switching to the last two components and how they impact on the computation of damages.

4.3.1 What Are the Recognised Heads of Compensation in Patent Law?

The heads of compensation recognised in patent law generally fall under two broad categories: the primary and secondary heads of compensation. The primary heads of compensation accommodates ‘immediate’ economic losses to patentees resulting from the inability of the patentee to exclusively capture the market of goods incorporating the patented invention(s) infringed. The primary heads of compensation are usually of two types: profits not earned due to lost sales, and reduced profits due to price erosion. 40 In establishing lost sales, it is generally expected that the patentee proves: first, that the patent in issue is exploited through sales, and not by licensing; 41 secondly, that there was market demand for goods incorporating the infringed patent; 42 and thirdly, that the patentee had productive capacity to meet market needs. 43 In the USA, 44 and now in Canada, 45 however, there is an additional factor; and this factor distinguishes these

---

42 Pneumatic Tyre Company Ltd v The Puncture Proof Pneumatic Tyre Company [1898] RPC 405; see also Boyd v The Tootal Broadhurst Lee Company Ltd [1894] RPC 175.
44 Panduit Corp. v. Stahlin Bros. Fibre Works, Inc. 575 F.2d 1152 (1978)
45 On account of Apotex v Merck 2015 FCA 171 and Pfizer v Teva Canada Limited 2016 FCA 161
jurisdictions from Commonwealth jurisdictions. It is whether there were non-infringing alternatives open to the infringer during the period of infringement.

The other primary head of compensation—price erosion—although treated as a different head from lost profits, inextricably connects with lost sales.46 As Werden et al. state, infringing competition can compel patentees into price reduction to attract customers.47 It is their position that to simply account for lost sales that the patentee suffered without having regard to two other components could mean the patentee is short-changed. These components are:

   a) the prices at which the patentee would have sold the goods for which sales were lost; and

   b) the reduced prices at which the patentee sold his products during the period of the infringing competition.

With regard to the first point, let it be assumed that the infringer sold infringing items at $5 apiece when the patentee would have sold same for $7 had there been no infringing competition. To compensate the patentee with the infringer's price gain of $5 ignores a shortage of value to the tune of $2. With regard to the second point, if the patentee had been forced, due to infringement, to sell, say, 20 items at $6 when he or she could have sold at $7, there would be a shortage in value to the tune of $1 per item sold. Thus, the patentee would have lost $20 due to price erosion.

In Wellman, Seaver and Head v Burstinghaus and Co,48 a competitive bid by the infringer for a contract compelled the patentee to lower his contract price. However, upon securing the contract, the patentee sued the infringer for the difference between the original contract price he had offered and that which the

48 [1911] R.P.C 326
infringing bid had caused him to accept. In the Canadian case of Allied Signal v Du Pont Inc\(^{49}\) the Canadian Federal Court advanced a more nuanced view. It held that the patentee could also claim for higher prices above the historical price that he or she could have sold for, but for the infringement. It is also important to note that price erosion has the effect of causing the patentee’s production costs to be wasted or unexpectedly increased.\(^{50}\)

Secondary heads of compensation, on the other hand, concern the extent to which the infringer can be held responsible for the losses suffered by the patentee in pursuance of commercialisation undertakings.\(^{51}\) Losses under this category of compensation are redeemable under any of the following heads:

1) lost sales on convoyed goods;
2) post-infringement lost sales; and
3) lost sales due to the infringer's accelerated re-entry into the market.\(^{52}\)

This class of compensable losses relate to the spillover effects of the infringement to the patentee’s business pursuits. They are, however, products of judicial creation arguably resulting from the overzealousness of judges in protecting inventors from commercialisation losses.\(^{53}\) This thesis does not pursue an explication and analysis of these secondary heads of compensation. This is because of reason of space and because the primary heads of compensation serve as sufficient basis with which to illustrate the difficulty of applying compensatory damages to patent law.

4.3.2 The Difficulties in Computing Patent Damages

\(^{49}\) 78 CPR (3d) 129, paragraph 217 ("Where competition by the infringer forces the patentee to reduce the selling price of its patented product, the patentee is entitled to the profit it lost both on the sales it actually made, and the sales that it would have made, at the selling price it would have maintained but for the presence of the infringing product")


\(^{52}\) Ibid.

\(^{53}\) Ibid.
It becomes pertinent to discuss how uncertainties of causation and ascertainment of loss affect the computation of damages. It is these uncertainties that result in the opportunistic effects of compensatory damages in patent law. These components borrowed from tort law are difficult to apply in patent law because treat an infringed patent as though it were solely responsible for the entire functioning and nature of products or commercial matters incorporating patented technology. Yet technical knowledge is cumulative, and so novel and inventive knowledge builds upon a pre-existing body of information. Moreover, patented inventions generally work in tandem with non-patented or public domain elements to produce goods and services. Thus, as explained by Justice Learned Hand:

It is of course impossible to imagine an invention for a machine or composition, or process, which is a complete innovation, emerging, full-grown, like Athene, from its parent’s head. It would be easy then to say that profits were to be attributed wholly to the invention. Such inventions are however mythical. All have a background in the past, and are additions to the existing stock of knowledge which infringing articles embody with the invention. ...54

This shows that patented inventions advance upon pre-existing ones and that patented inventions incorporated into finished goods along with other non-infringing component. Therefore, it is an entirely unworkable process to try connecting infringing sales to the loss of patentees.

Judges acknowledge that the components of causation and ascertainment of loss pose difficulties in patent law. In Betts v De Vitre55 Wood VC described the speculative and judicially problematic nature of ascertaining causation thus:

...damages of this description, namely, damages for the infringement of a patent where there is no licence granted at any time for the use of the patent, can only be ascertained on those very vague and guess-like data upon which it appears juries have been obliged to act in ascertaining what the actual loss has been that has occurred to a patentee by the user by some wrongdoer of his patent right.56

---

54 Cincinnati Car Co v New York Rapid Transit Corporation, 66 F.2d 592, 593 (1933).
55 [1865] 11 Jur NS 9 (Ch D) 10.
56 Ibid, 290.
Justice Jacob acknowledged the speculative nature of damages computation in the first instance determination of *Gerber v Lectra*,

Although courts warn against speculation, and urge a weighing of facts on a balance of reasonably foreseeable probabilities as customarily applied in civil litigation, the delineation between speculating and drawing inference from foreseeable probabilities remains thin.

The guess-like nature of determining lost profits is why Lord Shaw in *Watson, Laidlow and Co*

For the reason that patent infringement is a species of economic tort, determining causation and the ascertainment of loss is always contingent on the prevailing market conditions that characterize patented goods. In *Allied Signal Inc v Du Pont Inc* the Canadian Federal Court enumerated of such factors to be as follows:

(a) Presence of competing products in the market; (b) Advantages of the patented product over competing products; (c) Advantages of the infringing product over the patented product; (d) Market position of the patentee; (e) Market position of the infringer; (f) Market share of the patentee before and after the infringing product entered the market; (g) Size of the market before and after the infringing product entered the market; and (h) Capacity of the patentee to produce additional products.

Factors (a)-(g), as would be shown, are often shaped by one essential factor—substitutability. This relates to the degree or extent to which other goods or matters available to users can substitute patented goods. Factor (h), although significant, is not addressed in this discussion, as it is not often a source of

---

57 [1995] RPC 413.
59 [1914] RPC 118.
controversy. The substitutability of patented goods to users falls under three broad classes, namely: zero, imperfect and perfect substitutability. Before shifting to how demand elasticity shapes causation determination it is important to introduce the concept of price elasticity of demand.

Perfect substitutability exists where an increase in the price of a patented good would bring about a change (usually a decrease) in the quantity of it demanded. For example, if a piece Item A (patented), sold in August for $5 becomes $8 in September, some buyers would be unable to afford Item A at the new price and so would shift to other cheaper (perhaps technically inferior) economic alternatives to Item A. As is shown later in this article, where there a state of perfect substitutability it is generally difficult to determine both causation and the measure of resulting loss to the patentee. However, imperfect substitutability of patented goods may arise where a change in price can still result in a degree of change in demand even though it might not be significant.62 For this reason, where imperfect substitutability arises there is also likely to be difficulty with regard to the establishment of causation and loss just as with perfect substitutability cases.

However, it is possible that certain patented goods can have zero degree of substitutability. This can arise where due to technological superiority goods or because users find themselves (inextricably) locked into the consumption or use of the goods.63 In such situations, an increase in price would bring about little or no change in the quantity demanded, or cause no shift to other alternatives.64 Where there is zero substitutability, causation and quantum of loss is usually easy to determine. This is because in such situations sales that an infringer made on such goods would otherwise have gone to the patentee. However, notwithstanding zero substitutability an infringer may still secure patronage of buyers because of consumer loyalty to the infringer. An example is evident in *Watson Laidlow v Potts, Cassel and Williams*65 where Lord Shaw acknowledged

62 Ibid.
63 Ibid.
65 [1914] RPC 104
that certain customers would have stayed loyal to the infringer even had they not infringed; as such, the patentee could not have secured their patronage. At this juncture, it becomes necessary to address the consequence of substitutability on the assessment of causation and quantification of loss.

4.3.2.1 Proof of causation and loss where substitutability to user(s) is zero.

The traditions of Commonwealth jurisdictions, the US and Canada on patent damages converge when substitutability to user(s) is zero. This is because in such a market state, the patentee functionally has a monopoly. As such, it will be a straightforward matter to whatever sales the infringer made would have necessarily gone to the patentee, so long as the patentee had the productive capacity to produce as much as the infringer did. Hence, the loss suffered by the patentee is generally attributable to the volume of infringing activities.

A ‘two-supplier’ market often characterises such a market state, as it only consists of the patentee and infringer. In such situation, both causation and loss are easily established or implicit. The recent English case of AP Racing Ltd v Alcon Components Ltd presents a good example. In this case the English High Court found that although there were other alternative sources of calipers (Brembo and PFC), sold even at cheaper prices, the buyers of the infringing calipers (Joe Gibbs) were singularly interested in buying calipers incorporating the infringed technology due to its superior performance enabled by a structural optimisation design process. For this reason, they were willing to pay a price higher than other alternative sources would have demanded. Thus, the court held that sales that had the infringers not sold the calipers the patentees would certainly have done so.

A ‘two-supplier’ market will not, however, be considered to exist where relevant customers would have sought other market alternatives if they considered the patentee’s to be too expensive—i.e. imperfect substitutability. Thus, in *Fabio Perini SPA v LPC Group plc*,[70] the English High Court ruled that the patentee (Perini) would only entitled to damages for the full contractual gains lost if they could prove that they would have secured the contracts but for the infringement. The court found that although the patented technology was most suitable to the contractual specification, but there was an alternative technology (Gambini) which the customer (LPC, also an infringer) could apply.[71] Hence Norris J made the observation that ‘[i]n the world of (what would have been) Perini was not the sole company offering converting lines to LPC, and LPC would not have been compelled to accept what Perini offered’. An expression of similar reasoning is seen in *Coflexip v Stolt Offshore MS Ltd*,[73] also a contract-based case. In rejecting the patentee’s claim that they would have earned the contracts lost, Justice Jacob (as he then was) remarked that:

> the claimants are not in a position to show that the use of the apparatus or process of the invention was considered crucial or indeed even material for either the defendants or their customers. So this is not a case.....where there were really only two machines in the market — plaintiffs’ and the defendants’ — and it was the machines themselves which the customer wanted.[74]

Thus, where the market situation is one of imperfect substitutability, because a two-supplier market situation does not exist, causation and loss become onerous to prove.

A ‘two-supplier’ market may, however, take a different shape. Although there may be several alternatives available to relevant customers, the circumstances may place the patentee’s goods or services in a ‘niche market’ (a ‘mini-market’ or ‘submarket’ as the American jurisprudence calls it).[75] Consider, for example, a

---

[70] [2012] RPC 885.
[71] Ibid, 198.
[72] Ibid.
[73] [2002] EWHC 1686 (Ch).
[74] Ibid.
[75] See *Yarway Corp. v Eur-Control USA, Inc.*, 775 F.2d 268, 276 (Fed. Cir. 1985); see also John Skenyon and Frank Porcelli, ‘Patent Damages’ (1988) 70 *Journal of the Patent and Trademark Office*
market for fruit juice-based sweeteners where sweeteners sold in liquid form. If, however, patented sweeteners sold in tablet form render them a species of fruit juice-based sweeteners, this may create a sub-market. In such a niche market situation, using the words of the US Federal Circuit in *Kaufman Co. v Lantech*,76 ‘it is reasonable to infer that the infringement probably caused the loss of profits’.77 Some other Federal Circuit cases, such as *Electro Scientific Industries, Inc. v General Scanning*,78 and *Standard Havens Products, Inc. v Gencor Industries, Inc*,79 have adopted this reasoning. In sum, once again, it is apt to say that the Commonwealth and US approaches would converge in situations of zero substitutability.

As already hinted, there is the consideration that where there is zero substitutability causation and quantification of loss is easy as lost profits are ascribable to the infringement. The problem with such outcome is likely one of false attribution the implication of which is that the patentee is overcompensated. This is so for two reasons. The first is that as protection centres on the patentee’s goods, and ignoring the incremental nature of patents, there would be an undue imputation of patents. This is due to disregarding the roles of pre-existing stock of knowledge that the patent’s inventive concept builds upon and other non-infringing elements incorporated in the patented goods.

The second reason is essentially economic. It is that the patentee due to zero substitutability, the patentee can sell above market price to a considerable number of customers, and a small share of the market served by the infringer, even at lower prices, would not diminish the patentee’s economic rents.80 In other words, regardless of the infringement the patentee would have been able to recoup its marginal costs and capture a reasonable measure of social surplus contributed through its invention.


77 Ibid.
The implication of compensatory damages, where there is zero substitutability, on the patent licensing market and in litigation settlements is an endowment of the patentee with undue bargaining powers such that he can demand royalties up to the licensee’s ability to recover their marginal costs. The follow-on consequence of this is a serious likelihood of abrasion to consumer welfare and this is particularly so for two possible reasons. One is that licensees would have no choice but to pass on licensing costs to consumers through high prices if they choose to license from the patentee. The other is that if they choose not to license, then they avoid the market to which the patented goods relate, which leaves the patentee in a monopolistic position such that the patentee is able to charge supracompetitive prices.

4.3.2.2 Proof of causation and loss in cases of perfect substitutability

Where there is perfect substitutability and, as already identified, also where there is imperfect substitutability, causation and measurement of loss cease to be simple matters and the judicial attitudes of the Commonwealth and US diverge. For this reason, the analysis here applies equally to cases of imperfect substitutability.

In cases of perfect (and imperfect) substitutability courts engage in elaborate counterfactual analysis of what events would have ensued had infringers not unlawfully competed as they did.\(^\text{81}\) This analysis is largely founded on a reconstruction of what the patentee’s economic fortunes would have been had no infringement occurred.\(^\text{82}\) As foreshadowed, any such reconstruction of the market is fraught with robust imagination and practical difficulties. In *British United Shoe Machinery Company v Fussel Sons*\(^\text{83}\), on account of user(s)’ substitutability, Neville J acknowledged that mathematical certainty in the computation of damages is impossible because of the need for a reconstruction exercise, reasoning that ‘when


\(^{83}\) [1910] RPC 205.
it comes to a consideration of these damages, it is a matter of very great difficulty’.84 In the end his Lordship awarded a discretionary sum, reasoning that ‘it is impossible to arrive strictly at what the damages are, and therefore, that the best thing I can do it to assess them as well as I can’.85

Similarly, in *British Thomson-Houston v Goodman*,86 although the defendant had infringed the patented lamps, the English High Court was unable to determine the extent of damage in terms of lost sales the infringement had actually cost the patentee. Upon describing, as unfortunate, the patentee’s inability to lost profits, Russell J said ‘I am left in the dark as to evidence which entitles me to say that they have been deprived of any particular amount in respect of the sale by the infringer of these infringing lamps’.87 His Lordship then reversed the damages computation of £500 of the Master, and in its place awarded the patentee six guineas.

As the ascertainment of causation precedes the determination of loss, the ascertainment of causation in such market states is first addressed. For the purposes of establishing causation, the link or nexus between the infringement and the patentee’s loss, courts customarily rely on what is commonly known as ‘probabilistic causation’. A judicial approach to determining causation is to be probabilistic when it relies on a balance of probabilities to reach a judgment on the cause of economic injury to the patentee.88 Probabilistic causation is a pragmatic approach to establishing causation such that once it is shown that the wrongdoer caused injury by a given probability, then causation is considered established by that proven probability.89 Customarily, in applying probabilistic causation, it is usually expected that the claimant should prove causation more likely than not by more than fifty percent (50%) probability.90 Thus, for example,
in *Fabio Perini SPA v LPC Group PLC*, Norris J required the patentee, Fabio Perini, to show evidence that is more than 50% probable that but for the infringement the loss of profits would not have occurred. 91

Upon determining causation by showing sufficient probability of connection between the infringement and the ensuing loss, the patentee is required to present an analysis of loss suffered. Again, in perfect substitutability situations this is usually uncertain. The courts determine resulting loss based on the market conditions, particularly on account of consumers’ attitudes towards the patented goods. Thus, as reasoned in a recent English case, *SDL Hair Ltd v Next Row Ltd*:92

Where the quantification of the claimant’s loss depends on future uncertain events, such questions are decided not on the balance of probability but on the court’s assessment, often expressed in percentage terms, of the loss eventuating. This may depend in part on the hypothetical acts of a third party...93

In practice, however, there are two major competing models for determining causation and loss in perfect (and imperfect) substitutability situations: the Commonwealth model, chiefly used in Commonwealth jurisdictions; and the ‘Non-Infringing Alternatives consideration’ model, commonly used in the USA and recently adopted in Canada by the Federal Court of Appeal in *Apotex v Merck*94.

4.3.3 The Commonwealth Model

The House of Lords in *United Horse Nail Company v Stewart*95 categorically established the Commonwealth model by enunciating a rule that excludes consideration of non-infringing alternatives (NIAs) the infringer could have applied in reconstructing the counterfactual state of affairs.96 Their Lordships

---

91 [2012] RPC 885, 905.
93 Ibid, paragraph 31 (9).
94 2015 FCA 171.
95 [1887] RPC 260.
96 See, the Australian case of *Advanced Builders v Ramset Fastners* [2001] FCA 1098; see also *Boyd v Tootal Broadhurst Lee Co Ltd* [1894] RPC 175.
ruled that the availability of NIAs would not alter the court’s inference that the infringement caused lost sales to the patentee in circumstances where there was proven market demand for goods or services incorporating the patented invention.97

The implication of the exclusion of NIAs from the reconstruction exercise means that the proof causation would generally be a relatively easy matter for the patentee under the Commonwealth model. This is because all that is expected of the patentee is to show that it is more probable than not that the infringement caused the loss. Once that requirement is satisfied causation is established, then focus shift to quantification of loss. Quantification of loss would usually be difficult, however. In dealing with the determination of loss, the primary curial step adopted is statistical probabilities, often stated in percentage(s).

The market factors identified in Allied Signal Inc v Du Pont Inc,98 would usually shape the court’s assessment on the quantum of loss. Again, it is reiterated that factor (h) is excluded from discussion because it is usually uncontroversial. In addition, as already stated factors (a) to (g) are shaped by the factor of substitutability. However, in quantifying loss in cases of perfect substitutability and (imperfect substitutability) two routes emerge under the Commonwealth model to quantify loss. One is based on the statistical probabilities that the patentee would have secured a given portion or volume of sales secured by the infringer—factors (a) to (e), while the other is to quantify loss in proportion to the market share of the patentee—factors (f) and (g). This classification is not definitive as factor (e)—market position—could influence market share.

Via the first route, the court extrapolates the patentee’s loss from the market realities that preceded the infringement, or market realities it believes would have existed had there been no infringement,99 as illustrated in Gerber v Lectra.100 There the defendant’s counsel argued that the just way to quantify loss would be

97 [1887] RPC 130, 268.
98 78 CPR (3d) 129.
to proceed as in tort law whereby each wrong sale of the infringing machines is
treated as a tort. Thereupon, an assessment of the probability that each sale would
have been secured by the patentee had there been no infringement, is made.
Justice Jacob J dismissed this argument, saying:

Theoretically, one should look at the evidence concerning each sale; form
a view as to the probability of that going to Gerber, and a view as to the lost
profit on each such sale. Then there should be an overall summation of the
products of each probability and its associated lost profit. In practice that
would be a pointless exercise given the impossibility of coming up with
anything like precision, for each probability or each particular loss of
profit.101

Jacob J ultimately favoured reaching quantification of loss based on the general
impression of things, as he perceived them.102 He took certain market factors
bordering largely on those of factors (a) to (f) into account in approximating
‘collective probabilities’,103 and upon this basis ruled that the patentee would only
have been able to secure sales of 15 out of the 25 machines that the infringer
produced. This amounted to 60% of the sales made by the infringer. The Court of
Appeal affirmed this finding, saying: ‘He simply awarded 60 per cent of the total
sum claimed as loss of profit on the 25 machines. One can infer that he must have
found the figures for loss of profit for each machine proved’.104

In Fabio Perini SPA,105 in seeking to quantify the injury suffered by the patentee in
losing the contracts in issue to the infringer, Norris J considered varying market
factors. Balancing these competing considerations, he ‘put the chance of Perini’s
successfully obtaining the contracts at 65%’,106 and awarded the patentee

101 Ibid, 407.
102 Ibid, 408.
103 Ibid, 414.
damages on that basis. Similarly, in *Ultraframe (UK) Ltd v Eurocell Building Plastics Ltd*¹⁰⁷ Kitchin J approached causation on the balance of probabilities, saying:

To my mind all of these points illustrate the inherently difficult and uncertain nature of the exercise both experts were seeking to perform. At the end of the day I have to make an assessment based upon my impressions of the evidence as a whole. Doing the best I can on the materials before me I have reached the conclusion that a reasonable split of the sales of Pinnacle 500 between each of the three categories falls somewhere between the two positions taken by the experts and is as follows:

i) Category i): 56,700 (42% of 135,000)
ii) Category ii): 10,800 (8% of 135,000)
iii) Category iii): 67,700 (50% of 135,000)¹⁰⁸

As noted above, the second route to quantifying the patentee’s loss under the Commonwealth approach is through a market-share analysis—factors (f) and (g). The court assesses the probability or chance that the patentee could have secured the infringer’s sales in proportion to the patentee’s market share. In the USA, where market share analysis is also popularly applied, it is used to overcome evidentiary uncertainties only where infringers cannot show that there were NIAs available—i.e. where NIA considerations are excluded from causal determination.

*State Industries Inc. v Mor-Flo Industries*,¹⁰⁹ a US case, provides a good example of market share analysis. In this case, the market in question involved the water heating industry. The patentee’s product, based on a patented method, was adjudged to have 40 per cent share of the market. The infringer rightly contended that the market was competitive and that there were alternative products available for consumers to purchase and for the infringer to have applied (NIA argument). The court rejected the infringer’s argument for NIA considerations because, even though there were NIAs, the patentee’s products were technically superior to other NIAs and the patented feature was the basis of purchase from the infringer. However, because of the uncertainty as to the measure of lost sales

¹⁰⁷ [2006] EWHC 1344 (Pat).
suffered by the patentee, the court simply awarded the patentee 40 per cent of the sales obtained by the infringer (i.e. 40 per cent of the infringer’s sales in correspondence with 40 per cent of the patentee’s market share).

A similar approach featured in the Canadian case of Jay-Lor v Penta Farm Systems,¹¹⁰ where the Federal Court ruled that the patentee’s ability to secure over 800 sales during the infringing period was sufficient proof that it had considerable market position. Therefore, the patentee’s market share could be a basis upon which damages computation. For this reason, the court proceeded to determine what proportion of the 337 infringing sales the patentee would have captured by reference to the patentee’s market share.¹¹¹

Yet this market-share based assessment of probability may be unfair to the infringer, as it is simplistic and presumptive. It may also have unfair implications for the patentee in certain circumstances. An example is evident in the Canadian case of Allied Signal v Du Pont Inc.¹¹² The Federal Court found that, although there were other competitors in the market, all those who purchased from the infringer had previously been customers of the patentee, and would not have ceased purchasing from the patentee had the infringer not entered the market. This, the court observed, made it ‘necessary to look at the question on a customer-by-customer basis, instead of on a wider, ‘market’ basis under the market-share theory’.¹¹³ This was because the patentee had a 19 per cent share of the market, and to assess the probability that it would have secured the sales of its longstanding customers based on that 19 per cent share would produce an outcome unfair to the patentee. The court thus approached it because of lost sales on a customer-by-customer analysis.

Although the market share and ‘probability of sale’ routes differ, both are based on market conditions and bear the same flaws when applied towards assessing damages. Thus, O’Brien’s criticisms directed at the market share route apply to

¹¹⁰ 2007 FC 358.
¹¹¹ 2007 FC 358, paragraphs 209-225.
¹¹² 78 CPR (3d) 129.
¹¹³ 78 CPR (3d) 129, paragraph 36.
both. O’Brien maintains that applying market conditions to determine lost sales does not answer the question: to what extent did the infringement caused lost sales? He cites illustrations that can be adapted for the purposes of this discussion. For example, if a patentee had a market share of 30 per cent (or a probability of securing 30 per cent) of the infringer’s sales, to compute damages based on this 30 per cent market share (or probability) might be too simplistic. He reasons that reality might be over the course of time the patented goods attract higher prices. This might be because patentee's market share or probability of capturing sales would have increased by, say, 10 per cent.

Conversely, it might be that in the course of time another entity could have discovered and supplied NIAs that would have reduced the patentee’s market share or probability of securing a given sales volume by 10 per cent (subtracted from 30 per cent of the pre-infringement market condition equals 20 per cent). To compute damages based on market share or the probability of capturing sales could in the first situation undercompensate the patentee, but in the second situation have the opposite effect. As Werden maintains, a probabilistic assessment of loss in an elastic market is not without speculation, because it requires an assumption about the shape of the demand curve. He adds:

There will always be a margin for error because of the statistical uncertainty in estimating demand elasticities and because simulation depends on the assumptions that are never precisely right. But the tendency of the courts to resolve uncertainty against the infringer should assure that simulated damages are not too speculative to be accepted.

The Commonwealth model, as explained, has two principal features: first, it ignores NIAs, and secondly, it adopts a probabilistic causation and quantifies loss using statistical probabilities. In measuring loss, it takes the probability of something likely to happen, based on experience, as representing what actually

115 Ibid.
116 Ibid.
118 Ibid.
happened on a given occasion. Its upshot is that the quantification of losses most likely to founded on the chance that a thing could have happened, not on the factual finding that it did happen.

The application of this simplistic strategy for determining loss in tort law is one of the major reasons why some scholars argue that tort law damages computations are a lottery;\textsuperscript{119} it is worse with compensatory damages in patent law. If the extent of the economic injury to the patentee cannot be properly determined, to found the pursuit of adequate compensation upon statistical probabilities would inevitably have lottery effects. This can favour the patentee or instead, the infringer, thus being a catalyst for opportunism, as it enables either the patentee or the infringer, depending on the circumstances, to gain against the expectations of the patent system.

The likelihood of the lottery effect swinging in the patentee's favour — patentee opportunism — is more common. Opportunism is also very possible where the patentee is a well-established and efficient manufacturer. The only significant hurdle for patentees is to show that the patent reasonably accounts for market demand in relation to goods to which it is applied — in other words, that the patent is not a mere functional appendage.\textsuperscript{120} Upon proving this, and productive capacity is demonstrated, patentees simply need to show their products have gained traction in the market — a relatively easy task for a manufacturer with an established market position or brand. In sum, the effect of applying the Commonwealth model could mean other market factors (e.g., the infringer's marketing efforts and conditions of substitutability) prove less influential in the ascertainment of damages. This would mean that mounting a case of lost profits becomes easier. The consequence such on licensing and settlement outcomes is that the patentee's bargaining position would unduly stronger than the licensee's, which would mean a likelihood of higher licensing costs. Such a state of affairs has


\textsuperscript{120} A patented process or method that reduces the cost of production could also account for market demand because it could reduce the selling price of goods and services way below levels that other competitors are willing to sell.
the propensity to impose significant social costs in terms of decimating consumer welfare because licensees would pass on licensing costs to consumers in form of high prices.

Infringers too could gain some lottery effects—infringer opportunism—because of statistical probabilities used in ascertaining loss. This is especially so where an infringer is more efficient and/or well established and possesses greater means than the patentee. Greater efficiency in terms of productive means might enable the infringer to defeat a patentee’s claims that the patentee would have been able to satisfy market demand. In addition, an infringer with a well-established market brand and network can also easily defeat a patentee’s claim that they (the patentee) would have secured a given proportion of lost patronage. The infringer can attain this by countering the patentee’s use of statistical probabilities to establish loss, simply by showing that they enjoyed settled customer loyalty and market recognition. The likely outgrowth of this on licensing and litigation settlement is that the infringer would earn a stronger bargaining position and can thus diminish the chances of the patentee privately enjoying a reasonable share of the social surplus contributed through their inventions. This situation also has a likelihood of social costs represented in terms of significant reduction in the incentives of inventors and their sponsors to pursue inventive engagements.

4.3.4 The 'Non-Infringing Alternative Considerations’ Model

The ‘NIA Considerations’ model is an improvement upon the Commonwealth model, which only pays regard to demand-side factors, not to supply side-factors. It allows the court, in the counterfactual reconstruction of the market, for the purposes of causal determination, to have regard to NIAs available to the infringer during the course of infringement before assessing the statistical probabilities of how the infringement caused lost profits. Thus statistical probabilities for determining loss (as in the Commonwealth model) only come to the fore when the infringer cannot make a convincing case for NIA considerations. What this essentially means is that, under this model, causation is stricter than under the Commonwealth model, where probabilistic causation is founded on a balance of probabilities, excluding from consideration NIAs. Under this model, however, the
patentee must show that the infringer had no choice but to infringe. This is because the patentee is required to show that there were no NIAs available to the infringer.

If, however, the infringer can plausibly make a case for likely NIA, then as laid down in Panduit Corp v Stahlin Bros. Fibre Works, the USA locus classicus on patent damages, causation is considered disproved and there would be no need to proceed upon a quantification of loss. This would result in an award of reasonable royalties in lieu of compensatory damages. This is why much of the debate relating to this model largely relates to causation and not the quantification of loss. The reverse is the case with the Commonwealth model where less concentration is placed on causation, with more focus placed on quantification.

This model, as already hinted above, is rejected in most Commonwealth jurisdictions, but has settled use and application in the USA and has seen recent adoption in Canada with the Apotex decision. The Federal Circuit in Grain Processing Corp v American Maize-Products encapsulated the significance of assessing NIAs in determining causation, with the court remarking that:

... a fair and accurate reconstruction of the “but for” market also must take into account, where relevant, alternative actions the infringer foreseeably would have undertaken had he not infringed. Without the infringing product, a rational would-be infringer is likely to offer an acceptable noninfringing alternative, if available, to compete with the patent owner rather than leave the market altogether. The competitor in the “but for” marketplace is hardly likely to surrender its complete market share when faced with a patent, if it can compete in some other lawful manner.

Schlicher identifies the singular essence of the ‘NIA considerations’ model by pointing out that ‘in order to gauge the derived demand for the invention, it is necessary to assess the availability of substitute inventions’. This is because

---

121 575 F.2d 1152 (6th Cir. 1978).
123 185 F.3d 1341 - Court of Appeals, Federal Circuit 1999.
124 Ibid.
doing so helps ‘inquire about the nature and value of the product that infringer
could have made had it not infringed’.\textsuperscript{126} Ignoring an assessment of options
available to the infringer is liable, he argues, to over-compensate a patentee for
sales the infringer could truly have made absent an infringement, especially where
the technological value of the invention is marginal.\textsuperscript{127} The Canadian Federal
Court of Appeal in Apotex v Merck, holding that, shared this view: ‘Thus, the
American jurisprudence is clearly to the effect that the “but for” causation requires
consideration of non-infringing alternatives. Otherwise, patentees may be
overcompensated’.\textsuperscript{128}

There is a slight, but nonetheless important, doctrinal gulf between the Canadian
and US conception of the ‘NIA considerations’ model. In Canada, as ruled by the
Court of Appeal in Apotex v Merck, NIAs are only taken into account when four
conditions are met:

a) the supposed NIA is a true substitute and therefore a substantive
alternative;

b) the supposed NIA bears economic viability just as the patented invention
(i.e. would have competed well against the patented goods in the market);

c) during the period of infringement could the infringer have sold the NIA?;
and finally;

d) would the NIA have been sold by the infringer?\textsuperscript{129}

The difference between the Canadian and US conceptions borders on the third and
fourth conditions. In Canada, the supposed NIA must be available on the market
contemporaneously with the infringing products. The court in Apotex v Merck
reasoned that since Apotex (the infringer) lacked NIAs immediately available on
the market at the time of infringement that could replace the infringing products
if the infringement halted. For this reason, the court reasoned that there was no
proof that Apotex would have competed against the patentee with the supposed

\begin{flushright}
126 Ibid.
127 Ibid.
128 2015 FCA 171, paragraph 57.
129 Ibid, paragraph 73.
\end{flushright}
NIA. This was so even though Apotex had historically traded in the supposed NIA before eventually stopping it before infringement. The effect of this condition not being satisfied was the reason the court excluded the NIA supposedly claimed to be available to the infringer. For this reason, the Court of Appeal upheld the computation of damages that had been determined by the court of first instance based on probabilistic causation, which it described as the ‘common sense’ approach.

In the USA, conversely, all that infringers must show to be entitled to NIA considerations is that they could have competed using an NIA against the patentee. As reasoned by Judge Easterbrook in the first instance determination of Grain Processing Corp v American Maize-Products: ‘All that matters is that a product missing from the market can strongly affect, if not determine, the price a patent holder can obtain, and therefore the profit lost by infringement’. The Federal Circuit, on appeal, adopted this reasoning. In that case, although the infringer had an NIA process that could enable it to produce low-dextrose malto-dextrins, to a 2.3 per cent reduction in production cost compared to the NIA, it opted to apply the infringing process. The court reasoned that this did not mean the infringer could not have competed with the patentee in the production of low-dextrose malto-dextrins, especially as customers agreed that the outcomes of the NIA and the patented process were virtually identical. As the infringer could have competed using the NIA, the infringement could not be blamed for the lost sales the patentee claimed it could have captured had there been no infringement.

Based on the economic reasoning of William Baulmol et al and George Stigler, Judge Easterbrook opined that a potential competitor could nonetheless have an effect on price outcomes in the market. While Grain Processing Corp was not the first case on NIAs, it stretched the rules on NIAs.

---

130 979 F. Supp. 1233 (N.D. Ind. 1997).
131 185 F. 3d 1341 - Court of Appeals, Federal Circuit 1999.
135 A long line of cases such as: Central Soya Co. v Geo. A. Hormel & Co (723 F. 2d 1573 - Court of Appeals, Federal Circuit 1983); Pall Corp v Micron Separations (792 F. Supp. 1298 - Dist. Court, D.
Prior to *Grain Processing Corp* it was supposed that NIAs ought to have been available in the marketplace, and that it was not enough that they were readily available to be deployed to the market.\(^{136}\) *Grain Processing Corp* showed that NIAs that were complete in technical conception during the period of infringement ought to be treated as available even if not yet offered on the market.

With the nature of the ‘NIA considerations’ model identified and the difference in the conception of the model in the Canada and the USA highlighted, it becomes apt to demonstrate the second premise of this thesis—that uncertainty in both causal and loss assessment is inherent, and that this creates room for opportunism. To do this, the first and second conditions of the NIA model, as expressed by the Canadian court in *Apotex v Merck*, will be elided. This is because they are common to the Canadian and US conceptions. The third and fourth conditions are then addressed together, as it is on these conditions that the Canadian and US conceptions differ.

### i. Identifying the NIA

Identifying an NIA that would compete favourably in the market with patented alternatives is an exercise characterised by uncertainties. The chief source of uncertainty is whether the NIA must be a technical or economic alternative. In the Canadian case of *Apotex*, this issue was not explored as the court focused on the question of whether the supposed NIA was available (i.e. the *could/would* question).

In the USA, however, it is an issue explored to a considerable degree by the courts. Some US courts, including in *TWM Manufacturing Co v Dura Corp*\(^ {137}\) and *Radio Steel & Manufacturing Co v MTD Products Inc*,\(^ {138}\) have held that an NIA qualifies for consideration only if it has all the essential technical advantages that the infringed patent bears (termed ‘technical alternatives’). Yet others accept the sufficiency of


\(^{137}\) 789 F. 2d 895 - Court of Appeals, Federal Circuit 1986.

\(^{138}\) 788 F. 2d 1554 - Court of Appeals, Federal Circuit 1986.
the infringer having NIAs that could be sold in competition with the patented invention, without necessarily bearing the technical advantages possessed by it (termed ‘economic substitutes’). 139

As Skenyon et al argue, given that price is an essential factor that governs realities of the marketplace, it is only right to allow a consideration of economic alternatives, which, though lacking the technical advantages of the infringed patent, are apt to form part of the competitive market situation. 140 O’Brien, who deplores a consideration of economic substitutes as likely to cause the patentee to be undercompensated, holds a contrary view. 141 This is because ‘[a] patent may provide significant economic advantage in the marketplace even when there are substitutes’. 142 In economic reality, the view of Skenyon et al wins the day because it is the competitiveness of an NIA based on price, as against the patented product, that should matter for the purposes of compensatory damages. Requiring the NIA to be a technical substitute overlooks the fact that price is a chief determinant of market competition trends. To ignore economic substitutes on the market is apt to create a windfall for the patentee.

ii. Availability of NIA (the could/would question)

Having addressed the definitional problem associated with the model, it is pertinent to shift to the question of availability (i.e. the could/would question). The Canadian conception, which considers it imperative that the infringer would have sold the supposed NIA, is first addressed. The simple submission on the Canadian conception of the model is that, in effect, it pays lip service to the problem of overcompensation and is not capable of solving that problem as it avows. This is because requiring the infringer to be in immediate possession and market supply of NIAs detracts from the essence of the ‘NIA considerations’ model—a supply-
side consideration—which is basically to assess the market value of the patented product over other competing products in a relevant market.

Therefore, placing such a condition on infringers who raise an NIA defence becomes pointless on account of the essence of the defence which is to test if there are non-infringing technical substitutes which infringers could have competed lawfully in the market against the patentee with. It would be unreasonable to expect an infringer to possess NIA substitutes available in the market contemporaneously the infringing one sold. In such a situation there would be little point infringing when the infringing party already has NIAs already in the market. It is reiterated that the clear purpose of the ‘NIA considerations’ model is aimed at determining the incremental market value of a patented invention.

This point is supported by the reasoning of the Scottish Court of Session in United Horse-Nail Co v Stewart,\textsuperscript{143} which was overturned by the House of Lords for incorporating an NIA analysis. In this case, the patentee’s claim was that the infringement had caused it lost profits, which the court of first instance had assessed at the value of £530. The infringer contested this assessment, arguing that it could have produced the infringing articles without infringing the patent. The Court of Session accepted that the infringement only reduced the infringers’ costs of production, and therefore reversed the £530 assessment, reducing it to £50, reasoning that:

\begin{quote}
In such a case the question to be determined was, what advantage did the infringer derive from using the invention over what he had in using other processes then open to the public, which would have enabled him to obtain an equally beneficial result.\textsuperscript{144}
\end{quote}

On appeal, the House of Lords reversed, and the initial judgment of £530 damages reinstated. The House of Lords disagreed with the reasoning of the Court of Session because it turned a compensatory damages inquiry into what their Lordships considered akin to a disgorgement or an account of profits assessment.

\textsuperscript{143} (1887) RPC 130.

\textsuperscript{144} Ibid.
The US conception, characterised by the question ‘could the infringer have adopted and competed with an NIA’, is equally problematic because it is likely to throw up questions that could present informational uncertainties, such as “why was any such alternative not adopted and what would the market response have been if the NIA had been adopted”? This conception diverts inquiry from ‘what injury the patentee suffered’ to what ‘kind of injury the patentee could have suffered’. Moreover, it enables infringers to treat the patented inventions as ‘real options’. The ‘real options’ argument is that by allowing an account of NIAs infringers are able to shield themselves from damages liability by claiming that they could have achieved the same outcome without infringing. This creates the impression that, even without infringing, the infringer would have competed with the patentee favourably and caused the same losses by the infringement. The consequence is that an infringer can gain opportunistically from infringement.

To prove the point that ‘what the infringer could have done’ may not impact on the competitive advantage of a patented product, Hausman et al. present a plausible argument on how the incremental value of a patent may inform competitive outcomes and effectively bar an infringer from the market. They argue that if the technical advantage of a patent lies in reducing the cost of production (as in Grain Processing, where the infringed patent reduced the cost of the product over the NIA by 2.3 per cent), the infringement would certainly be the cause of lost profits. This is so even if the infringer would still have competed using the NIA. The reason, they explain, is that the 2.3 per cent difference in the cost of production would have affected the price of the eventual goods produced,

---


thereby causing the infringer (had it not infringed) to sell at a higher price than
the patentee.\textsuperscript{149} This difference in price, however marginal, would likely have
affected demand for, and competitiveness of, both goods in the market.

Counter to this is the decision of Judge Easterbrook in his (first) District Court
judgment of \textit{Grain Processing}. His Honour reasoned that the infringer was already
enjoying sufficient profit margins, such that if they switched to the NIA and
incurred the incremental production cost of 2.3 per cent (which it avoided by
infringement), it would still be able to compete favourably with the patented
goods by selling at the same price as it did while infringing.\textsuperscript{150} Hausman et al.'s
argument may be appealing in circumstances where the cost saving enabled by
the patented invention would have facilitated a competitive edge for the patentee.
Even if that were the case, the appeal of their argument starts to crumble when
the \textit{raison d'être} of the patent system is borne in mind. The patent system is aimed
at ensuring the promotion of inventive activities and investment towards
inventive ends by enabling inventors and their sponsors to capture a measure of
the social surplus bestowed on society through their inventions. This enables
them to recoup their marginal costs. Therefore, it is only just and optimal that a
remedy that enables the capture of a reasonable fraction of the social surplus
enabled by an invention is applied to patent law.

\subsection*{4.3.5 The US Conception of ‘NIA Considerations’ Model Is Best, But It
Does Not Obviate Systemic Opportunism}

The foregoing analysis surrounding the computation of patent damages reveals
that the only dependable touchstone for determining causation is the US
conception of the ‘NIA considerations’ model. This reflects the essence of the US
conception of ‘NIA considerations’ model as being focused on the incremental
market value enabled by a patent's inventive concept. It is superior to the
Commonwealth model because that model simply focuses on market factors

\textsuperscript{149} Ibid.
\textsuperscript{150} \textit{Grain Processing Corp. v American Maize-Products}, 893 F. Supp. 1386 - Dist. Court, ND Indiana
1995.
without emphasis on the competitive edge bestowed by a patent’s inventive concept. The US conception is also preferable to its Canadian counterpart, which is in substance little different to the Commonwealth model it departed from. As explained, this is because the Canadian approach hinges on market factors but not the incremental market value of an inventive concept.

The superiority the US conception of the ‘NIA considerations’ model bears over these other approaches is that it enquires into how the inventive concept of patents endows goods and services with a competitive edge, by comparing other alternatives the infringer could have competed with. By this standard, it is not enough that there is a decline in demand resulting in loss of profits to the patentee by reason of infringement; it must be shown that the infringer had no choice but to infringe.

However, the chief flaw of the ‘NIA considerations’ model is that rather than being used to simply determine the incremental value or profits derived from the inventive concept of an infringed patent, it is used to determine the degree of competitiveness of patented goods and services in a given market. Thus, where the infringer cannot raise an NIA defence and the market is elastic or relatively inelastic then causation is considered proven. The courts then rely on statistical probabilities to ascertain economic loss that results in problems of patent opportunism. By so doing, the same pitfalls that inhere in the Commonwealth model in perfect (and imperfect) substitutability would come to play. These are that: a) patentees can present a case of specious loss profit in proportion to their strength of their market position—patentee opportunism and b) infringers with sufficient market position can equally defeat patentees’ claims of lost profits in proportion to the strength of their (i.e infringer’s) market position—infringer opportunism. Both likely outcomes would determine the nature of both the patentee and licensee’s bargaining position in licensing and settlement arrangements.

With the identified pitfalls of the ‘NIA considerations’ model discussed above, it becomes pertinent to address the price erosion head of compensation and how it advances the second premise of this thesis.
4.3.6 Price Erosion and the Second Premise

Price erosion is equally problematic in its determination and can equally have opportunistic effects. In theory, the price erosion compensatory head might appear to be easier to determine than lost sales. This is because it is easy to compare the patentee’s pre-infringement with prices the patentee is able to charge during the period of infringement. Yet the real reason why prices dropped may be unrelated infringement. Glick accurately posits that a proper ascertainment of price erosion requires paying close attention to substitutability.\textsuperscript{151}

Epstein believes that to ignore substitutability may culminate in the patentee being overcompensated through the price erosion head in circumstances where buyers would have avoided the goods or shifted away to other alternatives.\textsuperscript{152} He gives the following example:

Suppose, for example, that in a market for widgets a patent holder sold 800 units and an infringer, the sole competitor, sold 200 units and that the actual price was $10. The patent holder therefore had revenue of $8,000, the infringer had revenue of $2,000, and the total market size was $10,000. Assume the infringement caused 10\% price erosion (i.e., in the but-for market the patent holder would have charged $11) and that there was 40\% profit margin on each actual sale. An erroneous damages calculation that ignored elasticity might proceed as follows: 10\% price erosion times $10,000 total market revenue ($1,000) plus the product of 40\% profit margin times $2,000 “lost” sales ($800) equals $1,800 total lost profits. The problem is, due to elasticity, the higher but-for price would not allow the patent holder to sell the 1,000 widgets that underlie this calculation.\textsuperscript{153}


\textsuperscript{153} Ibid, 15.
Another factor likely to cause overcompensation, in this context, is the condition of ‘quantity accretion’. This relates to an increase in sale accrues to a producer as result of reduction in prices, such that that producer is able to sell more than would have otherwise have been the case if higher prices had been maintained. Thus, when a patentee receives compensation for the measure of loss owing to price erosion without regard to the gains to the patentee from ‘quantity accretion’, then the patentee would naturally be overcompensated.154

As can be gleaned from case law analysis, courts are not blind to market realities. For example, in *Ultraframe (UK) Ltd v Eurocell Building Plastics Ltd*155 the English High Court refused a full award of damages under this head of compensation because, for a given period, significant market factors other than the infringement caused the patentee to reduce prices.156 This notwithstanding, judges are prone to award speculative sums based on their overall impression, using statistical probability, if they perceive that there has been price erosion due to infringement. This was clearly the case in *Gerber v Lectra*,157 where Justice Robin Jacob said: ‘[t]aking the best estimate I can, I think a figure of 90% is too high. More realistic I think is 60%. This must then be further discounted by 25% of itself to allow for the other’. The simple implication of this indeterminacy is overcompensation, the quintessence of patentee opportunism.

### 4.4 Conclusion

The leitmotif of this chapter is to demonstrate the dissonance and divergence between the purposes of the patent system, which is to promote inventive engagements, and compensatory damages, which are best suited to the protection of libertarian entitlements. The chapter argues that the patent system and compensatory damages are at cross-purposes, so that the two regimes cannot be harmoniously melded without the latter defeating the utilitarian objectives of the

---

former. For this reason, the submission of this article is that compensatory damages should be abolished from patent law. This submission is made chiefly because of the deleterious abrasion to social welfare which the mismatch poses. By avoiding this mismatch, the patent market would function more optimally; and certainly, there would be mitigation in the prevailing inefficiencies besetting the patent system, particularly the undue bargaining powers vested in patent holders.

To reiterate, the first premise shows that compensatory damages over-protect the commercial interests of patentees, and this has the propensity to go beyond what is necessary to ensure the promotion of inventive engagements. Therefore, the application of compensatory damages to patent law marks a doctrinal cleavage between the remedy and the utilitarian foundation of the patent system. The upshot of which is that the marginal social costs of operating the patent system likely surpass its marginal social benefits.

As to the second premise, case law analysis demonstrates that the quantification of compensatory damages is wholly probabilistic and speculative. The process is innately fraught with factual ascertainment difficulties, which the courts seek to address by applying statistical probability to resolve factual uncertainties and decide on damages computation. Its implication is to occasion a likelihood of unfairness to either infringers or patentees. But there can be no denying of the fact that the propensity of unfairness to infringers is greater. Even if the factual indeterminacies that surround causal ascertainment under direct causation can be mitigated (e.g. by adopting the US conception of the ‘NIA considerations’ model), opportunism remains inevitable.
Chapter 5
Reasonable Royalties and the Patent Market

5.1 Introduction

Reasonable royalties are a form of patent law remedy awarded alternatively (in circumstances where the patentee cannot prove lost sales) or cumulatively (in circumstances where the patentee is partly able to prove lost sales) to compensatory damages, which have been dealt with in the preceding chapter. Three rationales have been identified for the award of reasonable royalties and they are:

a) the need to prevent the injustice resulting from under-compensation in the award of nominal damages;

b) the need to eliminate incentives to infringe; and

c) stimulating incentives to innovate.¹

There are, however, two kinds of reasonable royalties. One is termed “post-infringement royalties”, usually awarded upon the cessation of the infringement.² This form of reasonable royalties is awarded either on those infringing sales for which the patentee is unable to prove lost profits, or in circumstances where the patentee did not practise the invention as a manufacturer or producer. Some writers argue that an award of both post-infringement royalties (on those sales on which the patentee could not prove lost sales on) and lost profits damages (on those sales proved to be lost) inherently overcompensate patentees.³ They

---

maintain that once patentees receive compensatory damages for lost sales, their real economic injury is compensated, and to further award reasonable royalties for those sales not proven to be lost amounts to an unnecessary windfall to the patentee. The other kind of royalty is awarded in lieu of an injunction, and is also known as an “on-going royalty”. It is awarded where courts, generally for equitable reasons, temporarily approve further infringement by suspending an injunction, or permanently do so by refusing an injunction.

The significance of the reasonable royalty remedy to the regime of patent law is very perceptible. The prevailing judicial approach on the application of the remedy shapes trends in the patent marketplace, particularly with respect to privately determined royalties. How royalties are determined is usually left to the mutual agreement of the parties to a licensing arrangement. However, parties to a licensing agreement are wont to disagree on computation of royalties. This is attributable to impinging patent market factors as discussed in Chapter 2, including information costs peculiar to patents (often resulting in the inherent illiquidity of patents as rights), and transaction costs.

Judicial attitudes, as reflected in case law, on the computation of reasonable royalties can have an influential impact on how parties privately determine royalties. In other words, judicial precedents on reasonable royalties furnish private parties with a default benchmark on how to determine royalties for the purposes of licensing. Should private bargaining fail and infringement occur, litigated matters would result in royalties usually being judicially determined. It is also important to reiterate that the patent market might be ex-ante or ex-post:

---

8 Ibid.
this has been discussed expansively in Chapter 3. Where it is ex-ante, the parties’ licence agreement is usually reached devoid of legal dispute. However, where it is reached ex-post it will often be in the shadow of litigation or threatened litigation.

Judicial determination of reasonable royalties is assailed with difficulties, not just in the parameters or methods of computation, but also in the cardinal rhetoric that patents are property rights—a philosophy that guides the courts in the determination process. This premise seeks to ensure the exclusivity of patents and protect them from incursions by third parties. In pursuing this objective under the aegis of the property model, courts pursue a compensatory agenda, enabled by the ‘compensatory model’ of computation. This model bears a propensity to inadvertently inflate the value of patented technologies for the purposes of reasonable royalty determination, which has a tendency to defeat the social welfare objectives of the patent system.

In adhering to this rhetoric in the computation of reasonable royalties, the incremental nature of knowledge upon which the patent system is founded is overshadowed. Further, the information costs peculiarities of patent rights are also often ignored, or paid lip service. It is submitted in this chapter that this simply facilitates opportunism. Opportunism, as repeatedly stressed in this thesis, arises when entities engage in behaviour that is lawful but nonetheless enables them to unfairly secure private gains. The corollary implication of this is that they impose inordinate (dynamic) social costs on others.

The chief argument of this chapter is that the alternative ‘restitutionary model’ of computation is best used in determining reasonable royalties in order to avoid opportunism. This argument rests on two major strands. The first is that the illiquidity of patents is often ignored in the ascertainment of reasonable royalties. As discussed in Chapter 2, the illiquidity of patents inheres in the information costs peculiar to patents: these relate to the ascertainment of their validity, scope

---

and existence. Thus, an infringer might infringe due to doubts about the validity of patents. They might infringe because they honestly believe that their activities are outside the scope of relevant patents. Infringement might occur where an infringer is ignorant about patents. The costs of acquiring information about patents may also be inordinate or inefficient to incur. Therefore, it is submitted in this chapter that to compute patent reasonable royalties on a compensatory rationale based on ‘property rhetoric’, creates an avenue for opportunism. By ignoring non-infringing alternatives that the infringer would have resorted to, but for information costs, as is customarily the case in Commonwealth jurisdictions, patent holders are apt to be opportunistically rewarded beyond the value of their patent to the infringer.

The second major strand of this chapter dwells on the truism that knowledge is cumulative and incremental in nature and as such patented inventions are simply additions to the body or stock of pre-existing knowledge. To treat a patented invention as though all its integers were entirely ascribable to the inventor’s efforts or ingenuousness, as the property rhetoric often erroneously urges in the determination of reasonable royalties, is at odds with the spirit of the patent system. This is because the patent system functions largely on the acknowledgement of the incremental nature of knowledge as is clearly reflected in the inventive step/non-obviousness requirement for the award of a patent. To fail to take into account the incremental nature of a patent right in the computation of reasonable royalties is apt to create avenues for windfall earnings (i.e. opportunism) to patentees in the computation of reasonable royalties.

This chapter is divided into three parts. Part I introduces the judicial and academic debate on whether reasonable royalties are of a compensatory or restitutionary nature. Part II dwells on the compensatory trend in reasonable royalty computation, and the potential implications this can have on the ordering of the patent market given resulting opportunism. Part III deals with the

restitutionary approach and its likely effects on the ordering of the patent market. This Part also considers the role of opportunism in applying this alternative approach. Here the superiority of the restitutionary model over the compensatory model in patent contexts is accentuated. Specifically, it is submitted that the restitutionary model should be upheld and the compensatory model dispensed with.

5.2 Part I: The Competing Juridical Bases for Reasonable Royalties

In the USA reasonable royalties are a product of statutory creation. The provisions of §284 of the US Patent Act state that reasonable royalties may be awarded to patentees to compensate for infringement where compensatory damages are not applicable. This has been taken to mean that reasonable royalties are intended to compensate patentees for the lost opportunity to bargain with infringers.12 This view has, however, been contested by Judge Richard Posner in Apple v Motorola13, and also by some academics.14

13 No. 1:11-cv-08540, 2012 WL 1959560.
15 See W & J Wass Limited v Stoke-on-Trent City Council [1988] 1 WLR 1406; see also John Glover, ‘Restitutionary Principles in Tort: Wrongful User of Property and the Exemplary Measure of
The text on this page and pages 162-201 have been removed for copyright or proprietary reasons. It has been published as: Abdussalam, M., Nielsen, J., Nicol, D., 2016. The superiority of the restitutionary model in computation of reasonable royalties for patent infringement: a comparative discussion, Intellectual property quarterly, 2016:4, 1-27
Having discussed the reasonable royalties remedy it becomes necessary to address the disgorgement remedy, which is often election in alternative to damages, with a view to discussing the place and implications of that remedy on the patent market. It is to the disgorgement remedy that focus shifts in the next chapter.
Chapter 6

Disgorgement and the Patent Market

6.1 Introduction

In this chapter the disgorgement remedy, otherwise known as an account of profits, is addressed. This is undertaken with the intention of highlighting its implications for the patent market and advocating an optimal and suitable standard for its application in patent law contexts. In Chapter 5, it was proposed that reasonable royalties are best conceived and applied as a gain-based remedy in the patent law context. In furtherance of the chief argument of this thesis that gain-based remedies are most suitable for patent enforcement, it is argued in this chapter the disgorgement remedy is the most ideal remedy for this objective.

This chapter is divided into three major parts. In Part I the nature of disgorgement as a remedy, is discussed. The equitable origins of disgorgement, and its constituent components, namely the quantum of disgorgement (i.e. ascertainment of gross revenue derived from infringement) and deductible expenses, are identified with a view to setting the stage for discussing the nuances of the remedy in patent law contexts. Part II deals with the ‘quantum of disgorgement’ component. Part III discusses the ‘deduction of expenses’ component.

The chief argument of this chapter is that the prevailing conception and application of the remedy in the patent law context deviates from the nature and peculiarities of the patent system and market. This chapter proposes instead that the remedy should be applied in the light of the cumulative and dynamic nature of technical knowledge. It is submitted that this is attainable when the remedy is conceived and applied to determine the differential profits derived from infringement as a proxy of the patent's marginal or incremental value, over technical alternatives available before it was granted.
It is also important to state upfront that this chapter draws largely on Commonwealth jurisprudence and academic literature on the disgorgement remedy in respect of patent law and other regimes of intellectual property (IP) rights. Rules relating to other regimes are relevant to the discussion because these other IP rights are essentially utilitarian in nature. Most prominent of the Commonwealth cases discussed in this chapter is the UK case of *Celanese International Corp v BP Chemicals*\(^1\). The judgment in this case delivered by Justice Laddie appears, to date, to be the most elaborate judicial disquisition on the disgorgement remedy in the patent law context. It has also remained relevant to the application of the remedy in other IP contexts. This case has shaped, and continues to shape judicial outcomes in Commonwealth jurisdictions despite its numerous flaws, identified during the course of this chapter.

This chapter also draws on relevant case law from United States (US) jurisprudence and academic literature on disgorgement. US case law and academic commentary on the application of the remedy is underpinned by a profundity of accounting and economic understanding. It is important to state that this chapter also draws on US material because of the close doctrinal contiguity across the common law world in the application of the disgorgement remedy. The body of US case law discussed in this chapter includes older cases applying the remedy to patents, decided before the coming into force of the *US Patent Act* of 1948 that abolished the application of the remedy to patents. The discussion also incorporates US cases on the application of the remedy to other IP rights, as the remedy has been recognised as applicable to other forms of IP such as design rights, trademarks, copyright and trade secrets. It is also important to state that while the regime of patent law is distinct from those of other species of IP rights, the aim of the remedy across IP rights is singularly to disgorge ‘profits’ derived by an infringer. Therefore there is a cross-fertilization of ideas around IP rights in the application of the remedy; although there might be nuances where the peculiarities of a given IP right so require.

\[^1\] [1999] RPC 203.
6.2 Part I: The Nature of the Remedy

As explained in Chapter 3, a party seeking reparations under the disgorgement remedy must elect the remedy over damages. Doctrinally, disgorgement and damages are considered to be mutually exclusive remedies. The former is intended to disgorge the unlawful gains derived by the infringer, while the latter is aimed at compensating for losses suffered by the claimant.\(^2\) Requiring the claimant to choose between the two remedies helps to avoid dual reparation or recovery as it prevents the claimant from being compensated for losses and also being rewarded through a monetary measure of gains derived by the infringer.\(^3\) The account of profits or disgorgement remedy pivots on a deterrence objective through its confiscatory \textit{modus operandi}: it enables the proprietor of an infringed entitlement to deprive a wrongdoer of gains derived from an interference with that entitlement.\(^4\) It is an equitable remedy that imposes on the wrongdoer a duty to keep all the gains derived from the infringing wrong for the proprietor of the entitlement as though the infringer were a (constructive) trustee.\(^5\) However, the infringer's mandate to account for profits derived from the wrongful act is personal, not proprietary.\(^6\) The implication of this is that the infringer is treated as a debtor, whose state of insolvency does not vest the claimant with the status


\(^3\) Ibid.


\(^5\) Per Justice Lloyd-Jacob, Electrolux v Electrix Ltd and Another [1953] RPC 158, 159: “The principle upon which the court grants an account of profits, as I have always understood is to be, is this, that where one party owes a duty to another, the person to whom that duty is owned is entitled to recover from the other party every benefit which that other party has received by virtue of his fiduciary position....”.

of a secured creditor over the assets of the infringer. Because the infringer is not a substantive fiduciary to the claimant, the infringer’s liability to account to the claimant does not attach to the infringer’s assets.

As with other equitable remedies it is discretionary: whether and how it should be awarded will be contingent upon the equitable considerations on the occasion. Thus, factors such as undue delay, laches and acquiescence, and knowledge borne by the wrongdoer may bear upon how and whether the remedy will be awarded. However, while these procedural aspects of the remedy are important as they can have significant implications on the outcome of the remedy, they are not particularly relevant to the discussions of this chapter. Instead, this chapter aims at addressing the substantive components of the account of profits remedy in the patent context—definition of the term ‘profits’, ascertainment of causation and deductible expenses—and the likely implications of these constituent parts for the ordering of the patent market. In that they determine what the patentee is likely to get, and of course, what the infringer is likely to lose in the pursuit of recovery under the remedy. The sum of these factors can shape the ordering of the patent market.

Having identified the nature of the remedy and its components it becomes necessary to address the first component of the remedy, namely the quantum of disgorgement.

8 Ibid.
9 Novoship (UK) Ltd and others v Mikhaylyuk and others [2015] QB 499, 535 (“We consider that where a claim for an account of profits is made against one who is not a fiduciary, and does not owe fiduciary duties then, as Lord Nicholls said in the Blake case [2001] 1 AC 268, the court has a discretion to grant or withhold the remedy”).
6.3 Part II: Quantum of Disgorgement

This Part targets the ‘quantum of disgorgement’ component of the remedy. This component has its own elements, namely the definition of the term ‘profits’ and the ascertainment of quantum (or causation). In addressing these elements, this part of the chapter adopts four major sections. **Section I** addresses the term ‘profits’, the definition of which is judicially disputed and lacking in uniformity. **Section II** introduces the quantification of the quantum of disgorgement and the pith of competing approaches for its ascertainment, these being the apportionment and the differential or incremental approaches. **Section III** critically discusses the features and practical implications of these competing approaches. Finally, **Section IV** makes a case for the differential approach in patent law contexts.

6.3.1 Definition of Profits

Some of the text on this page and pages 208-234 have been removed for copyright or proprietary reasons. Along with text on pages 254-255, it has been published as: Abdussalam, M., Nielsen, J., 2017. Rethinking the standard for ascertaining the quantum of disgorgement in patent law contexts, Journal of business law, 5, 413-435.
6.4 Part III: Deduction of Expenses

As already explained in the first section of this chapter the application of the disgorgement remedy implicates two constituent components: gross revenue/benefits and a determination of legitimately deductible expenses. So far the gross revenue/benefits component of the remedy has been expounded upon. It becomes necessary at this juncture to address the deduction of expenses component. The rationale behind allowing the infringer's deduction of expenses is the equitable precept of counter-restitution that ensures that a claimant does not gain an undue windfall from the efforts or resources of an infringer.128 It hinges largely on the reasoning that if the claimant had sought to pursue the profits so earned by the infringer, the claimant would have had to deploy his or her own resources towards that end.129

As Roach posits, if the infringer's expenses are not deducted, it could naturally result in an overstatement of the gains truly derived by the infringer.130 This is because the infringer might have necessarily incurred costs that supported the profits to be disgorged to the claimant, and to ignore these costs would amount to unjust enrichment on the claimant's part.131 However, while it is important to ensure that the claimant is not unjustly enriched in this regard, it is equally important to avoid allowing the infringer to make an inflated case of expenses to be deducted.132 For this reason, it is considered important, for the purposes of this remedy, that a clear line be drawn between expenses that can be equitably or fairly deducted to avoid the claimant being unjustly enriched, and those that should be ignored and as such not deducted.133 Therefore, it is incumbent upon the infringer

131 Ibid, 511-514.
132 Ibid.
to make a case for deduction; otherwise it will be resolved in favour of the
claimant.  

This was clearly expressed in *Diversified Products Corp. v Tye-Sil Corp*,
where the court stated, among other things, that:

In establishing an infringer’s profits, the plaintiff is required to prove only
the defendant’s sales; the burden then shifts to the defendant to prove the
elements of cost to be deducted from the sales in arriving at profit. Any
doubts as to the computation of costs or profits is to be resolved in favour
of the plaintiff.

6.4.1 Competing Methods of Deducting Expenses

The question of what expenses may be equitable and fair for the infringer to
deduct is not without debate. As reasoned by the US Court of Appeals in *Walker v Forbes*,
the disgorgement remedy is characterised by ‘fiendish difficulties’ due
to the ‘numerosity of variables’ that occasion gains and losses. Hence, there are
three competing approaches for computing the expenses of the infringer:

a) the full absorption approach;

b) the proportional absorption approach; and

c) the incremental approach

The basic feature of the deduction debate pivots upon how fixed costs are to be
treated, while variable costs are treated in a common fashion. Variable costs (also
known as direct costs) are costs that increase with output. Taking an agricultural
undertaking as an example, variable costs include expenses on inputs such as
seeds, fuel, water, electrical power, herbicides, and labour. The law treats these
costs as expenses that would not have been incurred but for the infringing

---

136 Ibid, 390.
137 28 F.3d 409 (4th Cir. 1994).
138 Ibid, 28 F.3d 412.
activity. However, fixed costs (also known as fixed overheads or simply overheads) are costs that do not vary with the level of output. These are costs that would usually be incurred by the infringer anyway, regardless of whether or not the infringing activity took place. Again, using an agricultural business example, fixed costs could include the costs of renting or buying land, tractors, machinery, other farming implements and fertilizer.

Finally, there are also intermediary costs (or step fixed costs) that are fixed costs by nature but are treated as though they are variable costs because they are only acquired as a result of the infringing activity. For example if a set of fertilizer products (i.e. fixed costs) is purchased specifically to replenish the soil before the cultivation of infringing seeds, such fertilizers are likely to be treated as a variable fixed cost because they comprise a fixed cost that was incurred specifically for the purposes of the infringement. However, what are fixed or variable costs will sometimes differ depending on business or economic dispensations. Thus what are variable costs to one business undertaking or in a given situation might be fixed costs to another.

6.4.1.1 Full Absorption Approach

The full absorption approach is judicially the least popular of the three costing approaches. This approach, as described by Margolis, allows for a deduction of fully allocated costs of the infringing activity as it enables a deduction of ‘all of the directs costs of the activity, plus any other variable costs, plus (and this is key) an appropriate allocation of fixed costs of the enterprise….’. For example a hypothetical farmer already established with land, farming implements, crop processing structures and support facilities (i.e. fixed costs) might choose to

140 Ibid.
141 Ibid.
142 Ibid.
143 Ibid.
146 Ibid, 1531.
cultivate infringing seeds, deploying these fixed assets towards this end. The full absorption approach will naturally allow a deduction of all the direct costs incurred for the purposes of the infringing cultivation, and then it will further allow a deduction of fixed costs. Where these fixed costs are assets, the value of depreciation of each multiplied by the duration of use will be deducted.\textsuperscript{147}

This can be illustrated using an example. Suppose a fixed factor such as a plough costing $500, with an expected life-in-use of 20 years, was used by the infringer towards the infringing cultivation for 2 years. To arrive at the deductible cost we divide the costs of that plough by its expected life, and then multiply the result by the years of use (i.e. $500/20\text{yrs} \times 2\text{yrs} = $50).\textsuperscript{148} Moreover, where the expenses are variable fixed costs or are non-depreciable fixed assets such as land, interest payments and intangible assets, they will be allocated to the infringing activity according to the duration or periods of use (i.e. the cost(s) divided by duration of use).\textsuperscript{149} However, in circumstances where the farmer, in addition to the infringing crop cultivates two or more non-infringing crops, then there will be an apportionment of the fixed costs among the ‘profit pot’ or profit base of each cultivated matter undertaken in accordance with the level or degree to which the fixed costs have been applied to each.\textsuperscript{150} For this purpose, it is important that the duration of time over which the infringing activity lasted be taken into account so that the extent, in terms of duration, to which the costs are applied can be ascertained.\textsuperscript{151}

Margolis explains the rationale for this deduction approach to be that ‘all costs must be accounted for somewhere’.\textsuperscript{152} He goes further to say that if only variable costs were deducted and fixed costs were disallowed for deduction, we could end up with a false impression that a given business undertaking is making profits

\textsuperscript{147} Ibid.
\textsuperscript{148} See, C. Steven Bradford, Basic Accounting Principles for Lawyers n17, 49-54.
\textsuperscript{149} Ibid.
\textsuperscript{151} Ibid.
\textsuperscript{152} Ibid, 1536; See also \textit{Neal v. Thomas Organ Co} 241 F. Supp. 1020, 1022 (C.D. Cal. 1965)
when in fact it is making lesser profits on account of fixed costs already incurred. Thus Margolis says:

After all, if each of the production of a multi-product firm shows positive margins on variable costs, but the sum of all those margins is inadequate to cover fixed costs, the firm as a whole will be unprofitable, notwithstanding the consistently encouraging news that we get by computing profits on variable costs one product at a time. 153

Laddie J reasoned similarly in *Celanese International Corp v BP Chemicals*154 in his application of the full absorption approach. He however founded his application of this method on the basis of ‘economies of scale’ or the need for a business entity to maximize its profits and make up for all expenses incurred.

### 6.4.1.2 The Proportional Absorption Approach

This is a commonly used approach that competes with the incremental approach (discussed below) and is in fact sometimes conceptually confused with that other costing approach in judicial circles. However, it is an approach that is characterized by niceties. This deduction method allows only a deduction of direct or variable costs, plus those fixed costs that can be proven to have supported or bear a nexus with the infringing undertaking.155 Proponents of this approach hold that the law must treat the infringer just as any businessperson should be treated.156 This approach is founded on the rationale of profit maximization or economies of scale just as the full absorption approach but differs from it in that it requires a nexus or connection between the profits earned and the fixed costs to be deducted. The following dictum expressed by McHugh J in *Dart v Décor* encapsulates the conceptual foundation of the costing approach:

153 Ibid, 1555.
156 McHugh J in *Dart v Décor* making reference to the words of LockHart J’s statement in *Odeon Theatres v Jones* ((1973) Ch 288, at pp.294, 299, 305) that“.....The courts are influenced strongly by the views adopted by professional accountancy bodies and men of business and the evidence of accountants is given great weight by the courts.”
Yet no business can be profitable if its revenue fails to recoup its general overhead as well as the direct cost of selling its products. That being so, no product can generate a profit unless its selling price recoups both the direct costs of its production and distribution and its proportional cost of the general overhead.\textsuperscript{157}

In following this approach, deductible overhead is limited to those items that ‘actually assisted in the production of the infringing profits’.\textsuperscript{158} In other words there must be a nexus between the overhead and the infringing activity. The foundation of this costing method, as identified in \textit{Kamar International Inc v Russ Berrie and Company},\textsuperscript{159} is simply that the claimant who seeks disgorgement must ‘pay’ rents for enjoying the fixed costs that supported the profit earned by the infringer. Thus in \textit{Kamar International Inc} the US Court of Appeal of the 9th Circuit reasoned that if overheads were not deducted the claimant would get a windfall. To this effect the court said:

\begin{quote}
If the copyright owner currently uses his fixed overhead to capacity, he would obtain by lawsuit net profits greater than he could have earned. Not only will the profits not cost him an increase in his own overhead; he will actually receive a premium representing the disallowance of the infringer’s overhead\textsuperscript{160}
\end{quote}

\textit{Dart v Décor} appears to be the first Commonwealth case to adopt this method of deduction, thus departing from the incremental method that is traditionally used. In this case the Australian High Court rejected the incremental costing method. The majority in the High Court considered that it would be unfair if the infringer were denied the deduction of both its opportunity cost \textit{simpliciter} (i.e. the reward the infringer would have gained if an alternative economic end had been pursued) and ‘a deduction for the cost of the overheads which sustained the capacity that would have been utilized by an alternative product and that was in fact utilized by

\footnotesize
\begin{itemize}
\item \textsuperscript{157} (1993) 179 CLR 101, paragraph 10.
\item \textsuperscript{158} Ibid, paragraph 36; See also \textit{Alfred Bell and Co. v. Catalda Fine Arts} ((87) (1949) 86 F Supp, at p.415.)
\item \textsuperscript{159} 752 F.2d 1326 (1984).
\item \textsuperscript{160} Ibid.
\end{itemize}
the infringing product’. The court reasoned that to deny both would leave the infringer to be in a worse off condition than if they had not infringed the patent in issue. The court however adopted the latter position (i.e. allowing a proportion of fixed costs), dispensing with the need to consider the infringer’s opportunity costs. It is however important to analyze both heads of deductions identified by the court, even though the court preferred one over the other.

The possibility of deducting the opportunity cost simpliciter of the infringer will be first addressed. This view is founded on the need to take account of alternative economic gains the infringer would have derived from other activities instead of pursuing the infringing one. However, as explained by Blayney and Wyburn, to allow a deduction of opportunity cost simpliciter could result in unfair outcomes to a claimant. They give two hypothetical situations to buttress this claim. One is a situation where the profit to be disgorged is $40 and the profits the infringer could have gained from an alternative outcome (the opportunity cost simpliciter) is $40. The implication of deducting the opportunity cost simpliciter in this situation will be that nothing will be given to the claimant. The other situation is where the opportunity cost simpliciter is $30 and the profit to be disgorged is $40. Deducting this opportunity cost means $10 will be awarded to the claimant. Unassailably, both outcomes would turn out to be unfair to the claimant because it overly subsidizes the infringer; this appears to be why the court jettisoned it.

The other possible basis for deduction is the proportion of overheads (from among all the infringer’s overheads) that supported the infringing activity. This was allowed by the High Court in *Dart v Décor* over the infringer’s opportunity cost simpliciter. Its origins, however, lie in US jurisprudence. In *Sheldon v Metro-Goldwyn Pictures Corporation*, a US judgment that influenced the court in *Dart v Décor*, Justice Hand held that there was no need for the infringer to show that the fixed costs for which deduction was claimed arose because of the

---

163 Ibid.
164 106 F. 2d 45 - Circuit Court of Appeals, 2nd Circuit 1939.
infringement. It is simply enough that those costs supported the infringement. His Honour said further that “Overhead” which does not assist in the production of the infringement should not be credited to the infringer. This position was followed in other popular cases such as *Wikie v Santly Bros* and *Warner Bros v Gay Toy* that clearly distinguished the *Sheldon* position on deduction of expenses from the full absorption method. A more recent classic US authority for this approach is *Schnadig Corporation v Gaines Manufacturing Company*. This case dismissed the incremental method as narrow for not including every element of expense that went into the manufacture and sale of the infringing matter.

It is however important to note that the decision in *Dart* has two differing bases for allowing deduction based on ascribable overheads. The first is that espoused by the majority, that the infringer can deduct the supporting overheads only if it can be shown that that portion of fixed costs had an alternative application it could be put to. In other words, if the infringer cannot show that there was an alternative business end to which that portion of overheads could be applied, then it means that the overheads costs would have nonetheless been incurred. Therefore, in the court’s reasoning, it should not be allowed for deduction. Further, the majority considered that Windeyer J in *Colbeam Palmer* did not allow a deduction of overheads because the infringing matters constituted a sideline. Hence, the infringing activity took a portion of the fixed costs that was in ‘excess’, for which there was no alternative use and would therefore have been incurred regardless of whether or not the infringing items had been sold.

This reasoning appears to have been accepted in the UK in recent cases such as *Hollister Inc v Medik Ostomy Supplies* and *Nigel Woolley v Up Global Sourcing*.

---

165 It is important to note that Sheldon was itself built upon the reasoning of the *Levin Bros v Davis Mfg. Co.*
166 106 F.2d 45 - Circuit Court of Appeals, 2nd Circuit 1939.
167 139 F.2d 264 (2d Cir. 1943).
169 620 F.2d 1166 - Court of Appeals, 6th Circuit 1980.
However, in the recently decided case of Design & Display Limited v Ooo Abbott, the UK Court of Appeal reasoned that this rationale for deduction does not end with the proof or otherwise of surplus capacity. The court reasoned that regard must alternatively be had to whether or not the infringer had a foregone non-infringing undertaking it could have applied the overheads to.

The second foundation, however, is that espoused by McHugh J in Dart v Decor, that it is of no relevance whether or not there was an alternative end that would have taken that portion of fixed cost, all that matters is that it supported the infringing profits:

However, the argument that overhead is a necessary element of the production of any good and the concept of opportunity cost are as applicable to “side line” activities as to other activities. If the infringer can prove that its overhead assisted the production or sale of the sideline product and can provide a fair and reasonable method of allocation, it is difficult to see why a proportion of overhead should not be allowed.

McHugh J expressed disagreement with the majority’s position on Colbeam Palmer as regards the sideline reasoning. He reasoned that it was of no consequence whether or not the infringing matter was a sideline product; deductions can be made on a portion of fixed costs that supported the infringing outcome so long as a nexus can be shown between both. McHugh J’s position appears to have been embraced in Canada as can be gleaned from the Monsanto v Rivett decision in which Zinn J rejected the plaintiff’s argument that only expenses that relate solely and directly to the infringement—incremental costs—are deductible. He reasoned that so long as expenses, whether fixed/capital or incremental, supported the infringement, they could be legitimately deducted from the infringer’s gross profits. However, he categorically rejected the full absorption

174 ([2016] EWCA Civ 95).
175 (1993) 179 CLR 101, paragraph 42.
176 Ibid, paragraph 42.
177 2009 FC 317.
approach. On appeal to the Canadian Court of Appeal Zinn J's reasoning was upheld.

6.4.1.3 Incremental Costing Approach

The incremental deduction approach is generally considered the traditional method of deducting costs. It allows only a deduction of variable costs that were incurred necessarily because of the infringing activity. It can be illustrated using an example. Suppose a farmer expends 50 cents on each of all the 1000 infringing seeds that he cultivates, and those costs ($500) represent the only variable costs so expended towards the infringing end. Let it be further supposed that when he reaps his cultivation he offers his cultivation for sale at $1 per harvested crop. However, he is only able to sell 600 units of the harvested crops out of the 1000 before an injunction is served against him. Thus he earns gross profits of $600 from the sale of the infringing farm produce. Applying the incremental costs requires a deduction of those variable costs that supported the earning of the $600 and this would be (50 cent x 600 infringing seeds – $600) $300.

In Teledyne Industries v Lido Industries Products the Canadian Federal Court, in applying this approach, clearly rejected the deduction of any overhead as the court reasoned that 'the principles of direct cost accounting do apply, and since these items represent portions of fixed expenses, they should not have been deducted from the revenue'. According to this approach, the only instance in which fixed costs (and in fact only a proportion of them) can be deducted is when they are incurred fundamentally for the purposes of the infringement. Thus in

179 Ross Grantham and Charles Rickett, Enrichment and Restitution in New Zealand (Hart Publishing, 1st edition, 2000) 466 ("...the concept of profits entails that deductions must be made from the total gross receipts in respect of the immediate costs of generating the total receipts.")
180 1982 CLB 7720.
181 Ibid, paragraph 47.
Gibson J, in the Canadian Federal Court described it as ‘contrary to the basic principles of equity to allow the infringer to deduct, as opposed to the increase of fixed expenses attributable to the infringing operation...’\textsuperscript{184} In the Australian case of \textit{Leplastrier v Armstrong Holland},\textsuperscript{185} Harvey CJ expressed the rule governing a deduction of such kinds of overheads in the incremental approach. He said:

If for instance, for the purpose of manufacturing these machines the defendant found it necessary to install a particular piece of machinery which was useful for making these machines and for nothing else, then it might be that depreciation of this machinery would be a proper item to allow as part of his costs of manufacturing the machines; if his machinery is used partly for the purpose of other machines it may be proper to allow him such depreciation for wear and tear on the value of his machinery as may be properly allocated to the work which has been done on the infringing machines as compared with the work done on the other machines.\textsuperscript{186}

The foundation of the incremental method rests entirely in the economic reasoning that all fixed costs are sunk and do not come to play in future decision making.\textsuperscript{187} Thus, when an economic actor considers engaging in an activity, he or she simply considers the direct costs that it will require and the profits likely to result therefrom in comparison with the direct costs and gains of other alternative undertakings.\textsuperscript{188} Judge Posner apparently adopted this method in \textit{Taylor v Meirick}\textsuperscript{189} when he said only costs necessarily incurred to generate the profit should be allowed for deduction. His Honour further said that costs that would have been incurred anyway, not being necessitated by the infringement, should

\begin{footnotesize}
\begin{footnotes}{183} 2001 FCT 889.
\footnotemark{184} Ibid, paragraph 115.
\footnotemark{185} (1926) 26 SR (NSW) 585.
\footnotemark{186} Ibid, 593.
\footnotemark{189} 712 F 2d 1112 (7th Cir. 1983).
\end{footnotes}
\end{footnotesize}
not be deducted. Another angle to view the incremental approach from is the deterrence goal of the disgorgement remedy. Roach has argued that allowing a deduction for overheads negates and weakens the deterrence value of account of profits, which is to make infringements of rights unattractive. He argues that disallowing a deduction of any kind of overheads, except those necessarily incurred for the infringement, is not punitive.

It appears that in the UK the incremental method still holds its traditional position of authority. Unfortunately, this is not entirely certain. In *Hollister Inc v Medik Ostomy Supplies*, the Court of Appeal clearly rejected Laddie J’s full absorption approach. This is because Laddie J did not require the defendant to show a connection between the overheads and the infringement. Upon rejecting Laddie J’s approach, the Court expressed acceptance of the view expressed by the Majority in *Dart v Décor*, without appearing to accept that of McHugh J. But then the acceptance of this majority view appears limited only to the consideration of certain factors such as sideline undertakings, excess capacity or the opportunity cost of that portion of overheads that sustained the infringement. This is evident as the court considered that the incremental approach is the dominant position in the UK and that is what it applied. However, the following quotation from the court’s judgement belies the view of the court that the incremental approach, in its veritable sense, was adopted:

> The defendant must show that the relevant overheads are properly attributable to that activity. All will depend upon the facts and circumstances of the case. For example, it may be relevant to consider whether a defendant has surplus capacity, whether the infringing activity was an additional line to an established business and whether the defendant’s overheads have been increased as a result of the infringing

\[\text{\textsuperscript{190}}\text{Ibid.}\]
\[\text{\textsuperscript{192}}\text{Ibid.}\]
\[\text{\textsuperscript{193}}\text{[2013] Bus LR 428.}\]
activity or whether its overheads would have been lower had it not engaged in that activity.\textsuperscript{194}

It becomes important to address the two factors (i.e. necessarily increased overheads and surplus capacity) identified as relevant by the judge with a view to showing that the incremental approach was adopted. As regards necessarily increased overheads, the court held that it was incumbent upon the infringer to show that overheads increased or would have been reduced but for the infringement.\textsuperscript{195} It appears that in the court’s consideration, overheads are attributable to the infringement when they were only incurred because of it.\textsuperscript{196} This is the quintessential nature of the incremental costing approach—only necessarily increased overhead costs can be deducted.

However, the second and alternative factor for deduction of overheads relates to surplus capacity (i.e. whether or not a given proportion of fixed costs had alternative uses they could have been put to). The court reasoned that the infringer was neither running to capacity, nor had alternative ends that the overheads could have been applied to, therefore its overheads were already incurred.\textsuperscript{197} However this position runs into difficulty if one asks: what would the court have done if the infringer had an alternative? The court gave no hint on this, as Kitchin LJ said: ‘We have heard little or no argument on the question of opportunity costs and they formed no part of the case advanced by either side so I need express no final conclusion upon them’.\textsuperscript{198} If, however, there was an opportunity cost or an alternative forgone by the infringer, then it appears that the court would have followed the reasoning of the majority in \textit{Dart} and allowed a deduction based on that ‘proportion of those general overheads which would have sustained that opportunity’.\textsuperscript{199}

\textsuperscript{194} Ibid, paragraph 85.
\textsuperscript{195} Ibid
\textsuperscript{196} Ibid
\textsuperscript{197} Ibid
\textsuperscript{198} Ibid, paragraph 86.
\textsuperscript{199} Ibid, page 452, paragraph 82.
This conclusion finds corroboration in a recently decided UK trademark case, *Nigel Woolley v Up Global Sourcing Ltd* 200 where the court interpreted *Hollister Inc* to mean that:

…. in any case where a defendant seeks to deduct an element of general overheads it will be for it to prove its business was running to capacity or that but for the infringement it would have sold other products or that its overheads would have been lower if it had not infringed. 201

Justice Richard Arnold, in extra-curial activity, describes the *Hollister Inc* reasoning as furthering the incremental cost approach in a book review. 202 In fact, the *Hollister Inc* reasoning is not a traditional application of the incremental cost approach as it allows for deductions of a proportion of overheads for which the defendant can show opportunity costs. A strict incremental cost approach would only be concerned about a deduction of variable costs and variable overheads, or variable fixed costs that were necessarily incurred for the purposes of the infringement. An incremental approach, in its true sense, is simply concerned with necessarily increased costs that supported the infringing activity.

6.4.2 Comment: The Superiority of Incremental Costing

It is submitted that as regards the infringer’s deductible expenses, it would be best to allow infringers to deduct only direct costs and incremental fixed costs solely referable to the infringing activity. One major reason is that the other competing costing approaches are apt to further attenuate whatever is left of the deterrence value of the disgorgement remedy upon following a strict causation standard to determine the quantum of disgorgement. 203 Unless the utility or value of a patent is singularly responsible for the profits earned by an infringer, which is rarely ever

200 *[2014] EWHC 493 (Ch).*
201 Ibid, paragraph 34.
202 Richard Arnold, ‘A Welcome Addition to the Literature on the Economics of Copyright’ (2015) 10 *Journal of Intellectual Property Law and Practice* 271-272 (“It is therefore unfortunate that Margolis’s analysis…was not cited to the Court of Appeal of England when it adopted the incremental costs rule in *Hollister Inc v Medik Ostomy Supplies Ltd*…..”).
the case, profits earned from infringement are usually attributable to a variety of joint factors including the patent right.\textsuperscript{204} This informs the need for scrutiny into the degree to which the infringement enabled the profit earned, as highlighted above. The necessity for this scrutiny is largely due to the utilitarian nature of patents: as explained by the Australian High Court in \textit{Warman International v Dwyer},\textsuperscript{205} this is to prevent a windfall to, or unjust enrichment on the part of, the IP owner.\textsuperscript{206}

This is in contradistinction with the concept of general property or private rights of libertarian foundation to which equity takes an expansive view in the attribution of profits. In this other context equity assumes, for deterrence purposes, that profits earned from a breach of an entitlement are ascribable to that breach, unless the wrongdoer can show demonstrably that the profits are attributable to innocent third parties or to lawful events.\textsuperscript{207} Simply stated, patent law is utilitarian in nature and the application of disgorgement to reinforcing patents reflects this such that pursuit of deterrence is modest. To therefore allow the infringer to deduct overheads would dampen or erode whatever vestige of deterrence is left of the remedy for patent protection.

Margolis defends an allocation of overheads, particularly along the proportional costing line, on the grounds that such allocation is economically sensible. According to him this approach to deduction of expenses is particularly pertinent because the average infringer firm is usually a multi-product one, which pursues a line of products supported by its fixed costs.\textsuperscript{208} Therefore, Margolis makes a case for deduction of overheads for the following reasons:

\begin{footnotesize}
\begin{enumerate}
\item \textsuperscript{204} See, Celanese International Corp v BP Chemicals Ltd [1999] RPC 203; see also, Potton v Yorkclose Ltd [1990] FSR 11; see also, Lubrizol Corp v Imperial Oil Ltd 1996 CLB 7481.
\item \textsuperscript{205} (1995) 182 CLR 544.
\item \textsuperscript{206} Ibid, paragraphs 23-33.
\item \textsuperscript{207} Mark Leeming, 'When Should a Plaintiff Take an Account of Profits?' (1996) 7 \textit{Australian Intellectual Property Journal} 127.
\end{enumerate}
\end{footnotesize}
a) an allocation of overheads has regard to the infringer’s opportunity costs in terms of alternative activities that would have engaged the (proportion of) fixed costs209; and

b) even if profit margins on variable costs appear positive, so long as supporting fixed costs are ignored, the real costs of production would not have been fully taken on board.210

According to Margolis a deduction of (a proportion of) overheads approximates the ‘incremental opportunity cost’ rule, which he argues is the touchstone for determining profits.211 As he posits, in economics all costs are opportunity costs, as such, applying the ‘incremental opportunity cost’ rule:

the profits of an action are the incremental revenues and other benefits, less the incremental opportunity costs. This measure of cost includes all of the explicit cost increases associated with factors that are variable, plus the opportunity costs of using fixed facilities for the infringing activity.212

Margolis argues that just as any rational profit-maximizing firm would not operate with excess capacity neither would the average infringer firm.213 Therefore, he reasons that to ignore the alternative ends to which the fixed costs of the infringer would have been applied is economically indefensible and can have punitive effects on the infringer.214 However, he holds the view that in circumstances where the infringer is a wilful or habitual infringer then it is only fair to hold that the infringer had excess capacity or no alternative end to apply the fixed costs to.215

It is submitted that to have regard to the alternative uses of overheads, as Margolis suggests, amounts to a deviation from the objective of the disgorgement remedy. The confiscatory concern of the account of profit remedy, which is well

209 Ibid, 1545.
210 Ibid, 1533.
212 Ibid, 253.
213 Ibid, 252.
214 Ibid, 250.
215 Ibid, 255.
acknowledged by Margolis himself, is aimed at ascertaining the gains derived by
the infringer from the infringement by subtracting expenses as expressed in
accounting terms, not in economic terms. That being said, the fact that the
average infringer firm is likely to be a multi-product firm, should not warrant the
deduction of overheads. This is because the remedy targets the 'profit pot' that is
implicated by the infringing activity, separately from other (non-infringing)
earning sources or engagements of the infringer.

It appears that Margolis reads too much into the analogy of the infringer with a
constructive trustee or agent of the IP owner. That analogy, however, should only
be one of convenience, and as such not be carried too far in ways that make it
appear as though the infringer acted as proxy for the patentee. This would have
the effect of overly diluting the modest deterrence objectives of the remedy in the
patent context and presenting an impression of the infringer representing the
interests of the patent owner.

Another factor that bestows superiority on the incremental costing method over
other competing costing approaches is that those other costing methods have the
inherent effect of subsidizing the infringer in terms of the costs of production. This
is based on the fact that an account of general overheads, whether in full or in
proportion to infringing activity, would give an impression of larger production
costs. If overheads were not ignored, they would enable an understatement of
profits earned from infringement and as such reduce the sum to be disgorged.

For this reason, an account of fixed costs, in any fashion, can be dismissed as
unsuitable to the remedy in the patent context. This position is justified by the fact
that to have regard to overheads is apt to render the disgorgement remedy weak
or lenient, the effect of which would breed infringer opportunism, and thus free-

---

216 Ibid, 254. (“The normal meaning of ‘profits’, unqualified, in the legal context is accounting
profits, so a reasonable understanding of the statute is that the owner of infringed property
would receive the infringer’s normal accounting profits.”)
there is only one profits ‘pot’. If different plaintiffs seek accounts in respect of different
infringing activities of a defendant within a single business, the totality of the profits ordered to
be paid should not exceed the total profits made by the defendant in that business.”
219 See, Robert Price, ‘Monetary Remedies Under the United States Copyright Code’ (1958-59) 27
Fordham Law Review 555, 566.
riding. The resulting implication of this could amount to defeating the purposes of the patent system, which is to encourage the pursuit of inventive activities.

Certainty is another major reason why incremental costing is superior. As shown above it is always difficult to demonstrate the nexus between fixed costs and profits gained; this is not that fixed costs do not support profits earned, but the degree to which they do so is always open to debate or unascertainable. To allow infringers to deduct fixed costs that supposedly supported the profits earned would open the door to strategic inflation of costs that could have the effect of diminishing the profits to be disgorged to the claimant. To avoid the adverse implications of this to the claimant it is best to adhere to that which is more certain: allowing only costs that are solely referable to, or necessarily incurred because of the infringing engagement.

Blayney and Wyburn correctly argue that an allocation of overheads can be arbitrary and incorrigible, and as such not lent to defensible or verifiable justification. They build this position on the arguments of Thomas. Thomas made a strong case against allocation of overheads on the basis of the reasoning that it is irredeemable because it is not based on a verifiable reference or nexus with market value and that is also arbitrary because there is no defensible theoretical justification for it. The proportional absorption costing approach is even more peculiar: as it may have regard to the opportunity costs or alternative activities that the fixed costs may be applied to, it calls for an inquiry into hypothetical events that did not in fact happen by assessing the opportunity costs of the infringer’s overheads.

The vitality of the ‘certainty’ head of argument in favour of the incremental approach becomes more perceptible in cases where the infringer is a multi-

product firm. Thus if the infringer were a multi-product firm with its varying activities (including the infringing one(s)), supported by the fixed costs of the firm which would have been incurred anyway, it appears only fair to allow a deduction of those direct costs that were necessarily incurred towards executing the infringing one(s). This is simply because an enquiry into the nexus between the overhead costs and infringing activities, if McHugh J’s reasoning in *Dart* were to be followed, is likely to throw up insoluble questions that can only be answered with a thorough collection of information of how such fixed costs connect with the infringing activities—a task which is virtually impossible of attainment.223

If, however, the reasoning of the majority in *Dart*, involving an assessment of the overhead’s opportunity costs, were to be applied, the nexus question would still remain undetermined. This is because such an assessment involves a counterfactual enquiry.224 This would raise speculative questions such as: what activity would have taken up that proportion of fixed costs; would the proportion of fixed costs have remained unused; or how would it have been re-allocated among other pre-existing (non-infringing) activities?225 Therefore, to avoid the information costs that can arise from enquiring upon the connection borne between overheads and the infringing activity it is only sensible and apt to have regard to the direct costs which are usually more easily ascertainable. In such a firm the direct or incremental costs of each separate activity pursued by the firm are more discernible and calculable than an allocation of fixed costs around them.

It is incontrovertible, however, that in circumstances where the infringer is very efficient, especially in managing or reducing costs, the claimant is likely to be handsomely rewarded in pursuit of the disgorgement remedy.226 However, to this position the reasoning of Zinn J in *Monsanto v Rivett* is a good response. His Honour reasoned that should there be any such windfall to the claimant as a result of an account of profits, ‘it does not lie in his mouth (i.e. the infringer) to argue that

225 Ibid, 845.
The text on this page and some of the text on the following page has been removed for copyright or proprietary reasons. Along with text on pages 207-234, it has been published as: Abdussalam, M., Nielsen, J, 2017. Rethinking the standard for ascertaining the quantum of disgorgement in patent law contexts, Journal of business law, 5, 413-435.
As regards the second set of issues, concerning deduction of expenses, it is submitted that only incremental expenses necessarily incurred expenses should be deductible (i.e. variable costs or variable fixed costs incurred solely for the purposes of the infringement). The rationale for supporting an application of the incremental costing approach is that it the only way opportunism from an infringer can be well curtailed; the deterrence objectives of the remedy can be well preserved; and certainty in standards of deductible expenses can be assured.

Having discussed the proper standard for applying the disgorgement remedy in patent law context, it becomes necessary to harness the discussions on compensatory damages in Chapter 4 and those on reasonable royalties in Chapter 5 with a view to proposing how the scheme of patent law remedies should be. For
this reason, focus shifts to Chapter 7 where an optimal scheme of remedies is proposed based on the piths of Chapters 4, 5 and 6.
Chapter 7
The Case for the Exclusive Use of Gain-Based Remedies in Patent Law

7.1 Introduction

This chapter marks the culmination of this thesis. The purpose of this chapter is to draw together the foregoing discussions from preceding chapters. In doing so, this chapter pursues a singular objective—to make a case for the exclusive use of gain-based monetary remedies in patent law. In pursuing this objective, this chapter builds upon a central argument of this thesis, which is that the patent system and compensatory remedies are at cross-purposes. Accordingly, this chapter advocates a scheme of monetary remedies, which are gain-based as being best suited to the utilitarian nature of the patent system. Thus, disgorgement and the restitutionary model of reasonable royalties are submitted as suited to ordering the patent market and reinforcing patents.

In pursuit of this objective, this chapter is divided into three parts. Part I of this chapter conveys the recommendations of this thesis on remedial reform. It highlights the scheme of remedies, of a gain-based nature, considered most suitable to the purposes of patent law. In this Part, a general justification for the exclusive application of gain-based remedies to patent law is provided. In Part II specific justification for the scheme of monetary remedies proposed in this thesis is provided. In providing this justification, an examination of the substance and viability of the recommendations made by Ted Sichelman1 and those of David Opderbeck2 is made. The gist of the recommendations of both scholars on monetary remedies reform is identified, and thereupon, the difficulties of their recommendations are explored. In Part III the modality for the application of the  

scheme of remedies proposed in this thesis is highlighted. This is then followed by a general conclusion to the thesis.

7.2 The Proposed Scheme of Gain-Based Remedies

In furtherance of the conclusions of Chapters 4, 5 and 6 of this thesis, the following scheme of monetary remedies is proposed as being most suitable to reordering the patent market:

1. an abolition of compensatory remedies (i.e. lost profits and the compensatory model of reasonable royalties);
2. an adoption of disgorgement of an infringer’s profits as the primary monetary remedy for patent law;
3. in the alternative, where the quantum of profits to be disgorged is difficult to ascertain or the patentee considers that the remedy provides unsatisfactory reparation, application of the restitutionary model of reasonable royalties; and
4. additionally, in cases where an infringer willfully infringes patents, the court may, as it thinks fit, with due regard to the utilitarian purposes of the patent system, award exemplary or punitive sums against the infringer in the application of the disgorgement or restitutionary model of reasonable royalties.

As is apparent from the scheme of monetary remedies proposed, it is submitted that disgorgement and the restitutionary model of reasonable royalties are the only monetary remedies suitable to patent law enforcement, with the former being treated as prime and the latter as second choice.

As can be gleaned from the discussions in Chapters 3, 5 and 6, these gain-based remedies are of resititutionary origin and nature. Scholars such as Lemley have always opposed the application of gain-based or restitutionary remedies to patent law.³ For example, he describes disgorgement as having strong deterrent effects against infringement. Based

on its punitive implications in the patent context, he argues that such remedies are aimed at promoting the exclusivity of ownership without regard to the incremental value of patented inventions. In addition, Lemley considers disgorgement to be inapposite to modern patent law, and in fact a drawback to it, because the majority of patent infringement suits are against entities who are independent inventors.

It would be highly erroneous to assert that Lemley is ignorant about how disgorgement and other remedies of restitutionary nature function. In fact, he has demonstrated an intricate understanding of the disgorgement remedy. This notwithstanding, Lemley does not, with respect, appear to appreciate how the disgorgement remedy can be applied to patents in such a way that fears of punitive effects and unfairness to independent inventors can be allayed. As correctly expressed by Opderbeck, what the disgorgement remedy aims at is simply a return to the patentee of the economic rents obtained by the infringer, without more. In other words, the remedy helps return to the patentee a reasonable measure of social surplus which they ought to enjoy from the sum of social benefits they have bestowed upon society through their inventions. Accordingly, independent inventors only give back the gains derived from the incremental value of the patent. Also, as Opderbeck argues in favour of the disgorgement remedy, it enables infringers to recoup their marginal costs, thus the likelihood of punitive implications is averted.

Valuable insight into the pertinence of gain-based remedies to patent law can be gleaned from the view of Kahn that:

(e)ach novel element arises inevitably from the past and itself sets up a complex interplay of causes and effects which in turn induce still further change. These novel elements are what we call inventions. They are, of course, created by individuals; but these individuals merely make explicit what was already implicit

5 Ibid.
8 Ibid, 179.
in the technological organism which conditions their thought and effort and within which they must work. Strictly speaking, no individual makes an invention, in the usual connotation of the term. For the object which, for linguistic convenience, we call an automobile, a telephone, as if it were an entity, is, as a matter of fact, the aggregate of an almost infinite number of individual units of invention, each of them the contribution of a separate person. It is little short of absurdity to call any one of the interrelated units the invention, and its “creator” the inventor.9

In essence, the superiority of both gain-based remedies is based on the fact that they function optimally in advancing the much needed riveting of patent law enforcement strategy to the utilitarian foundation of the patent system. As already established in chapters 1, 2 and 3 of this thesis, the biggest problem confronting the patent system is the misalignment between patent law policy and patent law remedies (i.e. the measures for the enforcement of patents). However, with gain-based remedies such misalignment can be easily reversed or corrected. This is simply because gain-based remedies are market-based by reason of the fact that their outcomes are contingent upon the interaction between the incremental value of that patented invention and market conditions. They enable patentees to secure, as reparations, not more than the economic value their inventions are able to command in the market. This reflects the superiority of gain-based remedies over compensatory remedies to patent law. Having stated the case for gain-based remedies, it becomes necessary to justify the scheme of remedies proposed.

7.3 Justifying the Scheme of Monetary Remedies Proposed in this Thesis

Providing justification for the scheme of remedies proposed in this thesis can only be duly appreciated with reference to academic works at the forefront in specifically addressing the gap between monetary remedies and the utilitarian nature of patent law. These are

the academic works of Sichelman and Opderbeck, identified earlier. These works suggest different schemes of monetary remedies that the respective authors, based on independent considerations, believe to be suitable to the enforcement of patents. In the discussion that follows, Sichelman’s recommendations are analysed but respectfully rejected on the basis that they are largely difficult to apply. Opderbeck’s recommendations are preferable for the simple reason that he considers disgorgement to be the most suitable monetary remedy. However, based on the analysis in Chapters 1 to 6 and the ensuing discussion, this thesis advances upon the recommendations made by Opderbeck, particularly on account of his conception of the disgorgement remedy and his rejection of reasonable royalties.

7.3.1 Ted Sichelman’s Postulations on Mooring Patent Remedies to the Utilitarian Foundation of the Patent System

Sichelman propounds a radical suggestion for remedial reform in that he calls for an abolition of private law remedies in patent law. Although Sichelman’s recommendations rightly emphasize the regulatory cum utilitarian nature of the patent system, and his paper plausibly exposes the root causes of patent opportunism, his suggestions on remedial regime change are difficult to apply. Sichelman’s recommendations suggest that in the award of monetary remedies to patentees, the courts should have due regard to the \textit{ex ante} incentives of inventors to invent.\textsuperscript{10} In other words, Sichelman proposes that patent law monetary remedies should be designed in such a way that they correspond to the degree of inducement necessary to incentivize inventive activities or engagements. According to Sichelman, monetary awards can be increased or decreased in proportion with the inducement conditions that characterize the conception of a patented invention being enforced.\textsuperscript{11} Another element that Sichelman considers important to influencing the adjustment, upwards or downwards, of remedial outcomes, is the social benefits of the patented invention.\textsuperscript{12}

\textsuperscript{11} Ibid.
\textsuperscript{12} Ibid.
Sichelman enumerates a list of factors that should guide courts in assessing or gauging inducement conditions, and thus informing the courts on how to increase or decrease monetary awards:

(1) the novel aspect of the invention is but a mere component of a complex invention and switching costs are high; (2) full compensatory damages substantially exceed or fall below the amount needed to incentivize research, development, and commercialization costs and risks of the invention, including opportunity costs; and (3) infringement is especially difficult to detect.13

He further identifies additional factors that should be influential in determining monetary awards:

... total R&D costs, testing costs (including regulatory approval), commercialization costs, increased profits versus baseline profits, comparable royalties, the value of other patented components, invention risk, commercialization risk, competition risk (including the risk of free riding), available profit from use or licensing, potential R&D duplication costs, potential consumer deadweight losses, the social benefit of the invention relative to noninfringing alternatives, the ease of detecting infringement, other barriers to entry (including other forms of IP), and the like in order to appropriately gauge remedies for infringement. 14

7.3.1.1 Critique of Sichelman’s Recommendations

The chief factor that embodies the difficulty of Sichelman’s thesis is his argument for adjusting monetary awards, upwards or downwards, to correspond with the levels of inducement and the social benefits provided by a patented invention. An adoption of Sichelman’s thesis would require the courts to possess information on the quantum inducement that occasioned the conception of an invention, and also a quantification of its social benefits value. 15 Certainly, information of this nature, which Sichelman suggests, is incapable of being extrapolated. As regards the inducement angle of

13 Ibid.
14 Ibid.
Sichelman’s thesis, Lemley reasons that even if such information were capable of being deduced, the courts would spread the supposed estimated sum essential to inducement across all possible suits that relate to the enforcement a given patent:

Suppose that a court could, with perfect knowledge, decide that $3 million was sufficient reward to encourage a particular type or quality of invention. Should the court simply award $3 million in the case before it? Presumably not, because the patentee can sue many different defendants, not just one. The court would want the total reward to the patentee to be $3 million, but that should include past suits, currently pending suits, any potential future suits, and whatever confidential licensing agreements were signed outside of litigation.16

Lemley asserts that if such an approach were taken in determining the award of monetary remedies, infringers would act strategically by waiting for the ‘inducement sum’ to be exhaustively apportioned among previous infringers successfully sued, and then embark upon infringement.17 This would certainly create a public goods problem, as one set of infringers would shoulder the burden of compensating patentees, while other subsequent infringers could infringe without suffering private costs. This would not only distort the redistribution of wealth, but it would also negatively affect incentives, especially because it would inordinately and indiscriminately deter access to inventive technologies.18

Lemley also addresses the social welfare component of Sichelman’s thesis on remedial reform. He reasons the award of monetary remedies in correspondence with the social value of an invention is likely to adversely affect third parties who might themselves be inventors.19 He submits that it would be best to let market-based remedies, which determine legal reparations based on market trends or forces, reinforce patents, rather than allowing arbitrarily computed remedies founded on social welfare judgements of patents.20

16 Ibid.
17 Ibid.
18 Ibid.
19 Ibid.
20 Ibid.
Cotter equally expresses disapproval of Sichelman’s recommendations, based on arguments essentially identical with those of Lemley. However, Cotter makes an incremental point: a remedial regime of this nature, which involves adjusting reparations higher or lower in correspondence with the quantum of inducement and social welfare, will result in significant diversion and use of political capital.\(^{21}\) It also bears the likelihood of rendering patent law more complex than it already is.

The rebuttals made by both Lemley and Cotter in response to Sichelman’s recommendations on remedial reform seem valid, largely because they do seem to be unworkable in practice. Notwithstanding the unworkability of his recommendations, the brilliance of his argument that patent remedies must complement patent policy’s utilitarian nature is acknowledged.

### 7.3.2 Juxtaposition with Opderbeck’s Recommendations

Having discussed the gist and vulnerabilities of Sichelman’s recommendations, it becomes necessary to shift focus to addressing Opderbeck’s recommendations. Opderbeck calls for abolition of both compensatory damages and reasonable royalties, while making a case for the sole application of disgorgement to patent infringement.\(^{22}\) In consideration of the fact that Opderbeck’s recommendations support the sole application of the disgorgement remedy, which is only a part of the gain-based remedies proposed in this thesis, it is necessary to address his recommendations in juxtaposition with the scheme of remedies proposed by this thesis.

It is important to state upfront how Opderbeck’s case for disgorgement is slightly, but significantly, different from that upon which this thesis is based. Opderbeck founds his justification for disgorgement on the interaction between the incremental utility of a patented invention and price elasticity.\(^{23}\) On the other hand, the justification for the disgorgement remedy in this thesis is that disgorgement enables a restitution of the incremental gains, monetary and non-monetary, derived by an infringer over non-

---


\(^{23}\) Ibid, 170-187.
infringing alternatives (NIAs) available before the grant of the infringed patent, as discussed in Chapter 6 of this thesis.

The component common to both bases of justification is the incremental value of a patented invention. The incremental utility of a patent can be assessed on the basis of the economic concept known as ‘marginal utility’. Going by this concept, as more of a class of co-equivalent or substitutable inventions is well supplied, any additional supply of such inventions would have no marginal value over pre-existing or already available ones.  

However, where a new technical solution or effect is an improvement to existing ones, then its value is to be measured against their utility. This assertion finds support in the words of Machlup that:

For example, if an invention permits an annual net saving of $1 million worth of labor and material, and if there are uses for the released labor and material, one is safe in estimating that the invention has a social value of $1 million per year.  

Having identified what the incremental value of a patented invention comprises, it becomes necessary to address Opderbeck’s justification for disgorgement as tied to price elasticity. As already discussed in Chapter 4, price elasticity relates to the responsiveness of consumers’ demand to changes in prices of goods and services. This responsiveness is shaped by the inverse relationship between price and demand, such that the higher a price goes, the lower demand drops. There is, however, an exception in responsiveness to price when goods are inelastic such that a change in price results in zero or an insignificant degree of change in demand. Price elasticity is of two kinds: ‘own’ price elasticity and cross price elasticity. ‘Own’ price elasticity relates to the degree of change in the quantity of a good or service due to changes in the price of that good or service. Cross price elasticity relates to the degree of change in the quantity of a given good or service due to changes in the price of substitute goods.

---

26 Michael Trebilcock, Ralph Winter, Paul Collins and Edward Lacobucci, Law and Economics of Canadian Competition Policy (University of Toronto Press, 1st edition, 2002) 73.
27 Ibid.
28 Ibid.
29 Ibid.
30 Ibid.
service due to price of other alternatives. However, implicit in Opderbeck’s argument is that the higher the marginal utility of a patented invention, the higher the likelihood that it can command incremental prices.

This can be gleaned from his recommendation that in circumstances where the entire profits made by the infringer are attributable to the infringed patent, the courts are to ‘presume that any amount the infringer made on sales of the patented product, or process, or method above its marginal costs are attributable to price elasticity and can be disgorged by the plaintiff’. A hypothetical example can be given to show how Opderbeck’s conception of disgorgement would apply. If an infringer makes $12 profit from the sale of an infringing item, it is to be assumed that that sum of profit is attributable to the effects of the interaction between the patent’s value and price elasticity. Therefore, the $12 is to be disgorged to the patentee.

Further to this, Opderbeck submits that in circumstances where the infringed patent only contributed to part of the infringer’s profits, then the substitutability or cross price elasticity of demand should govern the application of the disgorgement remedy. Although he does not provide a precise signpost on how this will work, one can deduce that he intends a comparison between the infringed patent and NIAs. Moreover, he adds that in circumstances where a patent accounts for profits to convoyed goods—i.e. goods sold along with patented goods— restitution must be made to the patentee for the profits attributed to the patent as having impacted on such sale of convoyed goods.

Some significant weaknesses can be identified in Opderbeck’s case for disgorgement. They are as follows:

1) it does not account for non-monetary economic advantages derived by infringers e.g. reduction in the costs of production, efficiency gains, satisfying regulatory requirements, and other non-monetary improvements in the economic conditions of the infringer that the infringement enabled;

31 Ibid,
33 Ibid 184.
34 Ibid 186-187.
2) in circumstances where the infringed patent accounts for the entire profits earned, the infringer's opportunity costs, in terms of NIAs, are not factored in and thus the incremental profits gained from infringement are not the central to disgorgement; and

3) in circumstances where the infringement accounts for only part of the infringer's profits, regard is to be had to NIAs, but the parameters for such comparison are not provided.

Having addressed Opderbeck's justification for disgorgement and highlighted its weaknesses, focus now turns to the justification for disgorgement provided in this thesis. As the central justification is restitution of incremental gains, it is argued that an incremental/differential standard of disgorgement, as proposed in Chapter 6 is most ideal to patent law. This standard of disgorgement requires a comparison to be made against NIAs available before the infringed patent was issued, in assessing the quantum of economic gains derived from an infringing activity. The monetary value of the incremental difference in economic advantage is then disgorged to the patentee.

The features of this approach are essentially threefold. The first is that outcomes of disgorgement could be cumulative, as a result of accounting for both monetary gains and non-monetary gains. The second is that gains that the infringer could have derived from adopting NIAs would always have to be considered, whether or not the entire profits earned results from the infringed patent. The third is that critically, only NIAs available before the grant of the infringed patent are to be taken into consideration. This is because the incremental value of a patent is always measured against the prior art at the time of patent grant, but also because:

a) to extend NIA considerations to alternatives available before infringement will make it convenient for infringers to infringe and argue they could have pursued alternative measures; and

b) this will also create room for uncertainties in determining the scope of NIAs the infringer can claim could have been adopted.

However, it is possible that the gains secured by infringement may not particularly reflect the incremental value of a patent, in the sense that technical superiority of the patent is not able to facilitate a market price advantage. This is possible where market factors
other than price impact on the demand for goods. In such situations, the likely outcomes of disgorgement may not be satisfactory to the patentee as the incremental profit to be disgorged would almost certainly be low. Notwithstanding this likely outcome, the incremental approach should govern the application of the disgorgement remedy. This is because the essence of the remedy is founded on the question: what profits did the infringer derive from infringement, which but for infringement would not have been possible? That the infringer could not secure favourable prices reflects the fact that the infringer did not gain from the infringement. Opderbeck suggests that in situations of this kind the patentee should get nothing because the infringer is not able to recoup marginal costs. 35 However it is strongly submitted in this thesis that the patentee should be awarded restitutionary reasonable royalties as analysed in Chapter 5. Such remedy makes the infringer pay a price for the use of the patent determined on the incremental value of the infringed patent over NIAs. Thus, the fact that the infringer made no profits at all does not obviate restitution to the patentee for the price of use.

7.4 The Modalities for the Application of the Recommended Gain-based Remedies

i. Where the Infringed Patent Accounts for the Entire Economic Advantage(s) or Profits Derived

In circumstances where the infringed patent accounts for the entire economic advantage or profits derived by the infringer, it is first important to determine if the infringer could have secured these same profits by applying NIAs. If it appears, based on evidence before the court, that there were no NIAs that would have enabled the infringer to secure equal economic advantage or profits, then it is only fair to attribute the infringer's gain to the patent and award the patentee the monetary value of whatever advantage is derived. Where, however, the infringer could have derived an equal economic gain by employing NIAs, it would not be fair to attribute the profits gained by the infringer to the infringed

patent. Barring infringement, the infringer could still have obtained the same gains. Therefore, the infringer would have no reason to disgorge gains to the patentee. In either of these situations the court might award punitive sums where it appears that the infringer acted willfully and in utter disregard of the patent right, and the court also considers an award of such sums would is not against the utilitarian spirit of patent law.

ii. Where the Infringed Patent Is Partly, But Not Entirely, Responsible For The Entire Profits or Economic Advantage(s) Derived By the Infringer

Where the infringed patent is not entirely responsible for the gains derived, then it would be important for the court to enquire, based on evidence provided, into the incremental/differential gain derived from the infringer by comparing against NIAs available before the patent was granted. As already explained, the incremental difference between the gain derived from infringement over what could have been gained in applying an NIA should be disgorged. Assume, for example, an infringer secured a net profit of $4 per infringing item, but without infringing and instead applying an NIA, $2 would have been earned. The differential or incremental advantage or profit derived from infringement (and to be disgorged) would consequently be $2 per item. Further, if it is proven that the infringer acted willfully, punitive sums can be awarded.

Upon determining the quantum of economic advantage from infringement (in i, above, and ii, here), an incremental costing method should be used to determine the infringer's deductible expenses. This involves the infringer deducting only expenses necessarily incurred towards the infringing activities, to the exclusion of costs that would have been incurred anyway. Expressed differently, the infringer should only be able to deduct variable costs and that proportion of fixed costs that arose as a result of pursuing infringing activities.

iii. Where Disgorgement is Difficult to Pursue or Its Outcome Would Be Unsatisfactory To The Patentee

It is, however, possible that the patentee might find it difficult to prove differential profits or that the differential profits that are provable might be unsatisfactory to the patentee. Where this is the case, it is recommended that reasonable royalties be computed on the basis of a restitutionary model. This model would comprise the following factors:
1. an account of the NIAs available to the infringer at the time of infringement, and a comparative assessment of their differential values;

2. ignorance of considerations such as switching costs to NIAs that would have been incurred by the infringer had an injunction been issued, as well as other considerations that would inflate the value of the patent and that do not bear on its incremental utility; and

3. the fact that the infringer made no profits would not obviate the need to pay reasonable royalties.

Thus, to the extent that the patent is substitutable with NIAs, no reasonable royalties should be paid. Justice Posner made a point to this effect in *Apple v Motorola*. Where, however, the infringed patent provides an incremental advantage over NIAs, the quantum of incremental advantage assessed by the court would determine the reasonable royalty to be paid.

Admittedly an incremental advantage inquiry is not without its flaws, difficulties and degrees of speculation. This notwithstanding, it bears comparative advantage over compensatory remedies (i.e. compensatory damages and the compensatory model of reasonable royalties) as these pitfalls are better controlled. What is sought is a measure of the marginal value derived by the infringer, and in the generality of cases its outcomes would usually be tolerable even if they were occasioned by margins of error. However, in the case of compensatory remedies, which are concerned with redressing losses suffered by the patentee, errors can have very serious opportunistic implications for either patentee or infringer as demonstrated in Chapters 4 and 5 of this thesis.

### 7.5 Conclusion

In conclusion, it is submitted that the application of gain-based remedies in the fashion proposed in this thesis would help obviate patent opportunism or at least significantly

---

36 No. 1:11-cv-08540, N D IL., June 22, 2012)
37 See Chapters 4 and 5 of this thesis.
minimize it so that the social welfare intendments of the patent system are not suppressed or defeated. By adopting this proposed scheme of monetary legal remedies the incentives and constraints of actors in the patent market will change significantly, such that the illiquidity of the patent market would be subservient to the attainment of the social welfare objectives of the patent system. Patentees, knowing that they can only obtain reparations in court that approximate the incremental value of their patented inventions, would have lesser incentives to unduly exploit users of their inventions. Infringers, also knowing that the economic rents they have usurped from patentees would be confiscated, with a likelihood of additional awards should willful infringement be proven, would have lesser incentive to infringe. For this reason, it is submitted that the patent market, both at the ex-ante and the ex-post level, would unequivocally function more efficiently.

However, it should be borne in mind that monetary remedies are not the only set of remedies applied in the enforcement of patents. There are prohibitory remedies, chiefly injunctions, which can be applied against infringers. The effect of these remedies is to prevent infringements or the continued unauthorized exploitation of patented inventions. As discussed in Chapter 3 of this thesis, injunctions have the most palpable likelihood of producing opportunistic effects. If injunctions are not awarded with due regard to economic efficiency considerations, the potential benefits of the scheme of monetary remedies proposed in this thesis could be significantly reduced. Therefore, this proposed scheme of monetary remedies should be applied in tandem with a judicial culture of making economic efficiency the governing consideration in the award of injunctions. With this regime of patent remedies, the utilitarian purposes of the patent system, predominantly represented in social welfare advancement, will be more easily attainable.

7.5.1 Justifying the Conclusion Reached

In Chapter 1 the problem for examination in this thesis was introduced as patent opportunism. This problem was explained as being attributable to the ideological gap between the patent system and the species of legal remedies applied towards enforcing patents. This ideological gap lies in the fact that patent law is meant simply to encourage inventors and their sponsors to invest in inventive activities, by furnishing them with
legal protection that enables them to secure their marginal costs. Also, the patent system captures a reasonable degree of the social surplus they have bestowed upon society through their inventions. However, the prevailing regime of legal remedies provides recompense that extends far beyond these aims. It provides strict exclusivity. Chapter 1 went further to identify the patent market as the context in which patent opportunism occurs. Further, in that chapter, transaction costs were identified as the main factor often blamed for the unworkability or illiquidity of the patent market which in turn makes opportunism possible.

In Chapter 2 the unworkability or illiquidity of the patent market is extensively addressed. Transaction costs, which are often blamed for the state of illiquidity were examined and judged as unavoidable in the patent market because of the inherent information costs that affect patents. These information costs comprise the indeterminacies of patent scope, validity and existence. The discussion revealed, however, that transaction costs are not limited to information costs as they also extend to bargaining and negotiation costs. Patent thickets and human behaviour, in turn, compound these factors. Three major ‘Coasian’ solutions—‘creating a firm’ (privately pooling patents), government interventions and raising patentability standards—were identified as having been trialled to solving the transaction costs problem. It was concluded, however, that transaction costs are insoluble in the patent market. It was further concluded that, rather than hope wistfully to ease transaction costs, one viable avenue for adjustment in order to avoid patent opportunism resulting from transaction costs, is the optimal application of patent law remedies.

In Chapter 3 patent remedies were recognised as informing the incentives and constraints of economic actors in the patent market. The chapter included analysis of the patent policy debate over whether monetary remedies are better applied to patents than prohibitory remedies. However, the chapter concluded that this debate fails to address the ideological dissonance or mismatch between patent remedies, both prohibitory and monetary, and the social welfare objectives of the patent system. It was submitted that focus should be placed on monetary remedies with a view to showing their opportunistic implications on the patent market and system.
In Chapter 4 compensatory damages, the first of the monetary remedies, were dealt with. Here it was submitted that compensatory damages overreach the purposes of the patent system being a legal remedy that has the effect of diverting focus away incremental value of patents by focusing inordinately on compensating for the lost profits of patentees. Further, the legal rules for determining the computation of damages, both in terms of causation and the quantification of lost profits, unavoidably result in opportunism—‘lottery effects’. These ‘lottery effects’, although commonly arising unduly to the favour of patentees, can also have favourable implications for infringers. In this chapter it was submitted that compensatory damages should be abolished from patent law.

In Chapter 5, reasonable royalties, the second monetary remedy for examination were dealt with. The opportunistic effects of the customarily applied compensatory model of reasonable royalties computation were brought to the fore. The three major components of the compensatory model— the holdup power of injunctions, established licensing value of a patent, and timing of licensing negotiation—were shown to be factors that enable opportunistic effects. Against the continued application of the compensatory model, the restitutious model was suggested as a better and more suitable model. This suggestion was based on two simple grounds: 1) the restitutious model is concerned only with the incremental value of the infringed patent over available NIAs; and 2) it conforms to the peculiarities of the patent market, especially the problems of transaction costs.

In Chapter 6 the third and last monetary remedy, disgorgement, was analysed. This gain-based remedy is not one that particularly lends itself to opportunistic outcomes as its aim is simply to confiscate whatever economic advantages an infringer has obtained from an infringing activity. However, it could have slight tendencies to produce opportunistic effects (usually in favour of the patentee), in circumstances where an apportionment approach is applied to determining the quantum of disgorgement. It could also have opportunistic effects (usually in favour of the infringer), where either the full or proportional costing method is applied in deducting the expenses of the infringer. To ensure that tendencies of opportunism are avoided, it was suggested in this chapter that an incremental or differential approach should govern the determination of the quantum of disgorgement, while the incremental costing method should be applied in computation of deductible expenses of the infringer.
In this concluding chapter, **Chapter 7**, it is proposed that gain-based remedies should be applied exclusively to the enforcement of patents. Further, the scheme of gain-based remedies should be as follows: disgorgement should be the primary remedy, but in circumstances where the quantum of disgorgement is difficult to prove or the outcome of the remedy is unsatisfactory to the patentee, then the restitutionary model of reasonable royalty can be elected by the patentee.

### 7.5.2 Concluding Remarks

The chief concern of this thesis has been solving the vexed problem of patent opportunism. To do so would have the positive effect of obviating the social costs that often result from the unworkability or illiquidity of the patent market and would also avoid impediments to realising the social welfare objectives of the patent system. This problem, as reflected in this thesis, has the adverse potential of defeating the social welfare objectives of the patent system. However, as this thesis demonstrates, the most viable solution to the problem of opportunism is not by investing increasingly on exchange-easing mechanisms. Reality shows that facilities such as compulsory licensing, patent pools and heightened patentability requirements have not been able to obviate transaction cost problems, especially with the emergence of the patent thicket phenomenon. The solution to this state of affairs lies instead in changing the incentives and constraints of players or actors in the patent market in a way that measures up to the utilitarian foundation of the patent system. The most effective way to achieve this change is through the adoption of gain-based remedies.

However, if gain-based remedies are not exclusively applied to patents as strongly advocated in this thesis, then the campaign and lamentation against patent opportunism would be pointless. This would mean the continued application of a socially costly and inefficiently designed patent system. The natural implication of this would be that the social welfare objectives of the patent system would continue to be at risk of being defeated. Inventors would procure more than they are entitled to, in terms of being able to secure more than is reasonably necessary to incentivise inventive outcomes and cover their marginal costs. Also, the resulting implication of this for society is that goods and services incorporating patented inventions could become inordinately expensively for members of the public and so too could the cost of follow-on innovation requiring access
to patented technologies. In sum, the continued application of patent law remedies as presently designed, would create a significant likelihood of the social marginal benefits of the patent system being dwarfed by the social marginal costs of the patent system.
Bibliography: Books, Journal Articles and Monographs

A

Abdussalam, Moshood, 'Identifying ‘the invention’ in Inventorship Disputes’ (2014) 11 SCRIPTed 44.


Aljalian, Natasha, 'The Role of Patents in Biopharmaceutical Patents' (2005) 11 Boston University Journal of Science and Technology Law 1-79


__________, ‘Do Patents Perform like Property?’ (2008) Boston Univ. School of Law Working Paper No. 08-08, 21


Chao, Bernard, ‘Causation and Harm in a Multicomponent World’ (2016) 164 University of Pennsylvania Review Online 61.


__________, ‘Holding up and Holding Out’, (2014) 21 Michigan Telecommunications and Technology Law Review 1


D


---

**E**


F


Hovenkamp, Herbert, ‘Notice and Patent Remedies’ (2011) 88 Texas Law Review (See Also) 221.


J


K


Lindsay, Michael and Skitol, Robert, ‘New Dimension to the Patent Holdup Saga’ (2013) Antitrust 34.


M


McGregor, Harvey, McGregor on Damages (Sweet and Maxwell, 18th edition, 2009) 104-133.


O


P


Park, Jae Hun, Patents And Industry Standards (Edward Elgar Publishing, 2010).


S


Say, Jean Baptise, *A Treatise on Political Economy; Or the Production, Distribution and Consumption of Wealth*, (Translated by Prinsep from the 4th edition originally written in French; Philadelphia: Claxton, Remsen & Haffelfinger, 1880)


Schmookler, Jacob, ‘Economic Sources of Inventive Activity’ (1962) XXII The Journal of Economic History 1-20


Stephen, Frank, *The Economics of Law* (Billing and Sons Ltd, 1988)


T


X


Y

Z


List of Cases

Australian Cases

Colbeam Palmer Ltd v Stock Affiliates Pty Ltd (1968) 122 CLR 25.
Hampton v BHP Billiton Minerals Pty (No 2) [2012] WASC 285
Tenderwatch Pty Ltd v Reed Business Information Pty Ltd [2008] FCA 931.
Unilin Beeher BV v Huili Building Materials Pty (No 2) [2007] FCA 1615.
Zupanovich v B and N Beale Nominees Pty [1995] FCA 1424

Canadian Cases

Allied Signal Inc v Du Pont Inc 78 CPR (3d) 129.
Bristol-Myers Squibb Co v Canada (Attorney General) [2005] 1 RCS 533.
British Thomson-Houston Co Ltd v Charlesworth Peebles 1923 SC 599.
Diversified Products Corp. v Tye-Sil Corp (1990), 32 CPR (3d) 385 (FCTD).
Ductmate Industries Inc. v Exanno Products Ltd 1987 CLB 4452
Jay-Lor v Penta Farm System (2007) FC 358
Lubrizol Corp v Imperial Oil Ltd 1996 CLB 7481.
Monsanto Canada v Charles Rivett 2009 FC 317.
Monsanto Canada Inc v Rivett and Monsanto Canada Inc v Janssens [2010] FCA 207
Percy Schmeiser and Schmeiser Enterprises Ltd v Monsanto Canada Inc. and Monsanto Company [2004] 1 SCR 902.
Unilever PLC v. Procter & Gamble (1993), 47 CPR (3d) 479.

New Zealand Cases

ABB v NZ Insulators (2007) 74 IPR 172.

United Kingdom Cases

Aerotel Ltd v Telco Holdings Ltd and others (2007) Business Law Report
Boyd v Tootal Broadhurst Lee Co Ltd [1894] RPC 175.
British Motor Syndicate Ltd v John Taylor [1900] RPC 723.
Cartier v Carlile (1862) 31 Beavan 292.
Coflexip v Stolt [2002] EWHC 1686 (Ch).
Electric and Musical Industries Ltd v Lissen Ltd (1938) 56 RPC 23
Electrolux v Electrix Ltd and Another [1953] RPC 158
European Central Bank v DSS EWCA Civ 192 (19 March 2008)
HTC v Nokia [2013] EWHC 3778 (Pat) (UK)
Gafford v Graham (1999) 77 P&CR 73
General Tire & Rubber Company v Firestone Tyre & Rubber Company Limited and Others [1973] FSR 79

Inverugie Investments Ltd. v Hackett [1995] 1 WLR 713

Jaggard v Sawyer [1995] 1 WLR 269

Kirin-Amgen Inc v Transkaryotic Therapies Inc (No.1) [2001] RPC 1.


Markem Corp v Zipher Ltd [2005] EWCA Civ 267

Merck & Co Inc. v Generics UK (Ltd) 2004 RPC 31

Ministry of Defence v Thomas (1993) 25 HLR 552

My Kinda Town Ltd. (Trading As Chicago Pizza Pie Factory) v Soll and Another [1982] FSR 147.

Novoship (UK) Ltd and others v Mikhailyuk and others [2015] QB 499

Odeon Theatres v Jones ((1973) Ch 288

Pell Frischmann Ltd v Bow Valley [2011] 1 WLR 2370.

Penn v Jack (1867-68) L.R. 5 Eq. 81.

Peter Pan Manufacturing Corp v Corsets Silhouette Ltd [1964] 1 WLR 96.


Potton v Yorkclose Ltd [1990] FSR 11


Shi v Jiangsu Native Produce Import and Export Corp [2010] EWCA Civ 1582

Siddell v Vickers [1892] RPC 152


The Owners of the Steamship “Mediana” v The Owners, Master and Crew of the Lightship “Comet” [1900] A.C. 133.

Tire and Rubber Co v Firestone Tyre and Rubber Co. Ltd [1976] RPC 197.


Watson Laidlow v Potts, Cassel and Williams [1914] RPC 104.
Wrotham Park Estate Co Ltd v Parkside Homes Ltd [1974] 1 WLR 798.

Ultraframe (UK) Ltd v Eurocell Building Plastics Ltd [2006] EWHC 1344 (Pat).


US Cases


Apple v Motorola No. 1:11-cv-08540, 2012 WL 1959560.


Cincinnati Car Co v New York Rapid Transit Corporation, 66 F.2d 592 (1933).

Fromson v. Western Litho Plate & Supply Co, 853 F.2d 1568 (1988)

Gyromat Corp. v. Champion Spark Plug 735 F. 2d 549 - Court of Appeals, Federal Circuit 1984


Levin Bros. v. Davis Mfg. Co., 72 F.2d 163 (8th Cir. 1934).

Lam Inc v Johns-Manville Corp 718 F. 2d 1056 - Court of Appeals, Federal Circuit 1983


Microsoft Corp. v. Motorola Inc., 696 F.3d 872 (9th Cir. 2012)

Monsanto Company v Stauffer Chemicals [1988] FSR 57


Oiness v Walgreen Co 88 F. 3d 1025 - Court of Appeals, Federal Circuit 1996
Pall Corp v Micron Separations 792 F. Supp. 1298 - Dist. Court, D. Massachusetts 1992


Rite-Hite Corp v Kelley Inc., 56 F. 3d 1538 - Court of Appeals, Federal Circuit 1995

Sargon Enterprises, Inc. v. University of Southern California, 55 Cal.4th 747 (2012)

SmithKline Diagnostics, Inc v Helena Laboratory Corp 926 F. 2d 1161 - Court of Appeals, Federal Circuit 1991

Sun Products Group Inc v B&E Sales Co 700 F. Supp. 366 - Dist. Court, ED Michigan 1988


Yarway Corp. v. Eur-Control USA, Inc., 775 F.2d 268, 276 (Fed. Cir. 1985)
Legislation

Australia
Australian Patent Act (Cth) 1990

Canada

United Kingdom
United Kingdom Patent Act 1977

United States of America