

Aberdeen Student Law Review

**With thanks to our sponsor
Stronachs LLP**

**March 2017
Volume 7**

www.abdn.ac.uk/law/student-activities/aberdeen-student-law-review-95

**THE EDITORIAL BOARD
VOLUME 7**

Editors-in-Chief
Augustinus Mohn
Euan West

Editors
Mukarrum Ahmed
Kieran Buxton
Domenico Carolei
Iliyana Futkova
Dimitar Ivanov Kadiyski
Edward Russell
Evangelos Stefanis

**FOREWORD BY THE HON. LORD WOOLMAN
SENATOR OF THE COLLEGE OF JUSTICE**

Over the course of my career I have found one general truth: all branches of the law are fascinating. Each topic presents its own degree of intellectual interest. Our task as lawyers involves analysing the law and the facts. To find the right solution in an individual case, however, the analysis requires to be tempered. We may have to take account of human frailties, of financial considerations, of moral and policy factors. The articles in this volume cover a wide range of topics, from agricultural tenancies to homicide. Each one is an absorbing read. The authors write with cogency and clarity. It is difficult to achieve that effect. JK Galbraith said that his prose only appeared spontaneous after the eighth draft. So I congratulate the authors on both their industry and their scholarship.

**Stephen Woolman
February 2017**

ACKNOWLEDGEMENTS VOLUME 7

We would like to thank Sarah Duncan for her organisational expertise and readiness to help, Tamas Gyorfi for his technical assistance and Anne-Michelle Slater for her advice and continuing support. We would also like to thank Lord Woolman, the patron of the ASLR, for his encouragement and kind words throughout the publication process. Additional thanks are due to those of our predecessors who furnished us with advice regarding the editing and distribution process, namely Dominic Scullion, Leanne Bain, Colin Mackie, Ilona Cairns, Phil Glover and Constantinos Yiallourides. Without their input, this would have been a much more daunting task.

This publication could not have taken shape without the meticulous work and dedication of the editors, who have devoted countless hours of their study time to reviewing incoming submissions and editing the articles chosen for publication. Plaudits are also naturally due to the authors, not only for the excellent quality and rigour of their academic work, but for their general patience and co-operation during the publication process. In a similar vein, we are very grateful to the peer-reviewers, on whose expertise the quality of the ASLR irrefutably depends.

Last but not least, we would like to thank Stronachs LLP for their sponsorship and James Downie in particular. His enthusiasm for the ASLR has been a true source of inspiration.

**Augustinus Mohn and Euan West
Editors-in-Chief
February 2017**

CONTENTS

VOLUME 7

Articles

- The Fifth Element: Should Exclusive Possession Be Considered an Essential Requirement for the Constitution of a Contract of Lease in Scots Law? Mitchell Skilling
Page 1
- Agricultural Tenancies: Making Room for Generation Y – Will Only the Land Endure? Sam Read-Norrie
Page 27
- Trading Places: Benefits of Invoice Finance for Small and Medium Sized Enterprises as Opposed to Bank Lending Kahmun Goh
Page 56
- The Market Stability Reserve Strikes Back: A New Hope for the EU’s Greenhouse Gas Emissions Trading System Daria Draganova
Page 80
- The Lost Symbol: A Semiotic Analysis of the Psychoactive Substances Act 2016 Nicholas Burgess
Page 109
- The Classification of Murder and Slaughter in the Justiciary Court from 1625-1650: Malice, Intent and Premeditation – Food ‘Forethought’? Stephanie A Dropuljic
Page 121
- Guidelines for Contributors** *Page 149*

The Fifth Element: Should Exclusive Possession Be Considered an Essential Requirement for the Constitution of a Contract of Lease in Scots Law?

MITCHELL SKILLING*

Abstract

Historically, opinions have varied on whether exclusive possession or occupation of heritable subjects was an essential factor for a contract to be considered a valid lease. However, recent developments increasingly appear to suggest that this is correct. This is especially so in the context of the residential sector, though some doubt also exists in the agricultural sector. The two main arguments for this stance are that it promotes legal certainty and that the increased protection granted by security of tenure ought to be more difficult to access. However, this does not mesh well with the broad nature of the Scottish lease, especially considering the material differences that can exist between a residential lease under a statutory regime and a lease at common law. Accordingly, the article concludes by proposing an alternative approach to exclusivity more consistent with the spirit of the Scots law lease.

Keywords: Lease, Licence, Exclusive Possession, Residential Tenancy, Security of Tenure

1. Introduction

Scots law has long recognised the contract of lease, whereby a person known as a tenant gains the right to occupy another person's land for a set period of time in exchange for a consideration known as rent.¹ The Leases Act 1449, which establishes the circumstances in which a lease at common law may give the tenant a real right against a landlord's singular successor, shows that a common law lease has existed in Scotland since the days of feudalism.

Over time, leases have increased in variety and complexity. New social pressures have created a social rented sector that has developed alongside older methods of private renting, commercial leases and agricultural leases. Contracts for certain rights in relation to land usage, such as those relating to fishing and shooting, have also been recognised as leases in Scots law,² leading Rankine to comment that the concept of a lease runs far broader in Scotland than it does south of the border.³ In more recent times, the various Rent Acts and Housing Acts have created a

* LLB (Hons) Law with Spanish Language Graduate, University of Aberdeen (2016). The author would like to thank Kieran Buxton for his invaluable assistance during the editing process of this article.

¹ A McAllister, *Scottish Law of Leases* (4th edn, Bloomsbury Professional 2013) 3.

² See, for example: *Macpherson v Macpherson* (1839) 1 D 794; G Paton and J Cameron, *Law of Landlord and Tenant in Scotland* (W Green 1967) 73-84.

³ J Rankine, *A Treatise on the Law of Leases in Scotland* (3rd edn, Bell & Bradfute 1916) 2. See, in this regard, *Clapperton v Edinburgh Magistrates* (1840) 2 D 1385.

The Fifth Element: Should Exclusive Possession Be Considered an Essential Requirement for the Constitution of a Contract of Lease in Scots Law?

number of statutory tenancy regimes into which a contract may fall,⁴ offering protection for those who can prove that their contract meets the requirements of a lease for the purposes of that statute.

Traditionally, a contract of lease is considered to comprise four essential requirements: two or more different parties to the contract, heritable property to form the subject of the contract, a set period of time for the contract to be in force over (or a definite 'ish' or end date), and a price, usually in money, to constitute rent.⁵ There is also a fifth element that is often considered when determining whether a contract is a lease: whether or not the tenant has exclusive possession of the property.⁶ In the law of leases, this term is taken to refer specifically to exclusive possession vis-à-vis *the landlord*, who will have no right to retain any part of the let property unless such a right is provided for in the contract,⁷ which is why multiple tenants under the *same* lease agreement can still be regarded as exclusive possessors of the leased property.⁸ However, multiple apparent tenants under separate agreements regarding the same subjects would *not* be in exclusive possession, and indeed such a state of affairs would be impermissible under the original lease contract due to the landlord's implied obligation not to derogate from the initial grant.⁹ The 'exclusive possession' factor arises most frequently when the courts are called upon to distinguish between a lease and contracts that fall short of being a lease, most often referred to as licence agreements.¹⁰ This distinction is particularly important in the context of commercial leases, which are largely governed by contract law.¹¹

Exclusive possession is an essential requirement of a contract of lease in English law.¹² By contrast, Scots law points to exclusive possession as being a normal, but not yet cardinal, feature of a contract of lease.¹³ That is despite earlier cases, especially *Conway v Glasgow City Council*,¹⁴ giving the impression that Scots law may be moving towards viewing exclusive possession as an essential requirement of a contract of lease.

This article considers the context within which exclusive possession currently exists under Scots law and considers arguments for and against exclusive possession as an essential requirement for the creation of a contract of lease. At the same time, consideration is paid to a distinct, but closely related, issue: whether the presence of exclusive possession is not only essential for the constitution of a contract of lease but, together with the four traditional requirements (parties, property, ish and rent)

⁴ See, for example: Rent (Scotland) Act 1984, s 1; Housing (Scotland) Act 2001, s 11 (H(S)A 2001).

⁵ Paton and Cameron (n 2) 5-8; McAllister (n 1) 34-38.

⁶ See, for example: *Bradford Properties Ltd v British Telecom Plc* 1992 SC 12; *Conway v Glasgow City Council* 1999 Hous LR 20.

⁷ *Baxter v Paterson* (1843) 5 D 1074; McAllister (n 1) 62.

⁸ McAllister (n 1) 34.

⁹ *ibid* 62

¹⁰ *Bradford* (n 6); *Conway* (n 6).

¹¹ McAllister (n 1) 235.

¹² *Street v Mountford* [1985] 1 AC 809, 816 (Lord Templeman).

¹³ *South Lanarkshire Council v Taylor* [2005] CSIH 6, 2005 1 SC 182 [7] (Opinion of the Court); *Cameron v Alexander* 2012 SLCR 50 [58]-[59] (Opinion of the Court).

¹⁴ *Conway* (n 6).

decisive to that end. Does the presence of these five elements *guarantee* the existence of a contract of lease, or is even *this* arrangement, this seemingly strong, clear-cut example of a contract of lease, vulnerable to the contrary inference that no contract of lease was intended? Simply stated, where a contract permits the occupier of the property exclusive possession thereof, can this ever be interpreted as something other than a lease contract if, exclusive possession notwithstanding, it can be concluded that the parties did not *intend* their contract to be a lease? These two issues (whether exclusive possession is necessary and, secondly, whether it is decisive) are distinct, but they are also two sides of the same coin. Both relate to the tension that can arise between the *labels* that two parties use to describe their agreement (e.g. ‘landlord and tenant’ or ‘licensor and licensee’) and the rights and duties to which that agreement *actually* gives expression (e.g. ‘exclusive possession’). It is largely for that reason that these two issues (whether exclusive possession is essential and whether it is decisive) are considered together in this article. For the sake of simplicity, however, both issues shall generally be couched in the language of the first issue: whether exclusive possession should be deemed ‘essential’ for the constitution of a contract of lease.

This article establishes what the case law is able to tell us about these matters and what the prevailing academic opinions associated with the subject are. Having examined the merits and demerits of having exclusive possession as an essential requirement, it also proposes some possible alternatives to this position that could be more consonant with the principles of Scots law.

2. *The Established Law*

A. Background

A lease at common law is, ‘a contract of location by which the use of land or any other immovable subject is let for a period of time to the [tenant] in consideration of a determinate rent or duty to be paid or performed to the [landlord] either in money, the fruits of the ground or services,’¹⁵ with the term being used interchangeably between the contract itself and the rights associated therewith. At its most basic level, a lease is merely a contract between two parties: the original landlord and tenant. Consequently it does not, in itself, confer any protection or security of tenure on the tenant should the leasing party dispose of their property. This derives from the original non-sasine nature of leases, which, in feudal law, prevented them from being enforceable against singular successors.¹⁶ The Leases Act 1449 introduced a method of turning the tenant’s personal right into a real right enforceable against such a singular successor, paving the way for later statutory regimes to offer similar protection. It reads:

Item it is ordanit for the sauftie and fauour of the pure pepil that labouris the grunde that thai and al vthiris that has takyn or sal tak landis in tym to cum fra lordis and

¹⁵ Paton and Cameron (n 2) 5.

¹⁶ Craig, *Jus Feudale* II,7,28.

The Fifth Element: Should Exclusive Possession Be Considered an Essential Requirement for the Constitution of a Contract of Lease in Scots Law?

*has termes and yeris thereof that suppose the lordis sel or analy thai landis that the takaris sall remayn with thare takis on to the ische of thare termes quhais handis at euir thai landis cum to for sic lik male as thai tuk thaim of befoir[.]*¹⁷

Despite its brevity, the Act encapsulates all of the essential requirements of a *real right* of lease. In addition to containing the four requirements for a *contract* of lease, it is the fifth element of possession, introduced by the Act, that converts the tenant's personal right into a real right enforceable against singular successors. It should be noted, however, that it is an *implied term* of a contract of lease (hence, a term that only arises upon the classification of that contract as a lease by reference to the essential requirements of such a contract) that the landlord is obliged to grant the tenant exclusive possession of the subjects on the agreed date of entry.¹⁸ That said, certain reservations are allowed.¹⁹ A landlord might, for example, reserve the right to park a car within the area leased.²⁰

As illustrated above, on this view at common law, the right to exclusive possession arises from the conclusion of a contract of lease, as traditionally understood according to the four aforementioned requirements. This is distinct from viewing an express right to exclusive possession as constituting a fifth essential requirement for the contract to be classified as a lease. This implied term also deters landlords from avoiding the consequences of the 1449 Act, lest they argue that the tenant has not yet obtained exclusive possession. In other words, the tenant must have exclusive possession to obtain a real right under the 1449 Act,²¹ and he or she enjoys an implied right to obtain such possession under an antecedent contract. Failure to provide exclusive possession at the date of entry will provide the incoming tenant with a contractual claim for breach of the implied term.

B. Role of Exclusive Possession in the Leases Act 1449

The wording of the 1449 Act clearly indicates that it will come into effect only when the land has been, 'taken or shall [be taken] in time'. This offered retroactive protection to current tenants as well as extending that protection to those who would, at the time of the enactment, form contracts of lease in the future, something that the courts have always upheld unambiguously. The reasoning behind this was best explained by Lord President Cooper in *Millar v McRobbie*:

It has been well settled for centuries that possession under a lease is the equivalent of sasine in relation to feudal property. Without possession the tenant is merely the personal creditor of the [landlord]. By entering into

¹⁷ Translation from McAllister (n 1) 40: 'It is ordained for the safety and favour of the poor people that labour the ground that they and all others that have taken or shall take lands in time that come from lords and have terms and years thereof, that suppose the lords sell or alienate these lands, that the tenants shall remain with their leases until the end of their terms, no matter into whose hands that ever the lands come to, for the same rent as they took them for before.'

¹⁸ Rankine (n 3) 200; Paton and Cameron (n 2) 127.

¹⁹ *ibid.*

²⁰ Cameron (n 13) [59] (Opinion of the Court).

²¹ As interpreted in *Millar v McRobbie* 1949 SC 1, 8 (Lord President Cooper).

possession the [tenant] publishes to the world in general, and to singular successors in particular, the fact of his lease[.]²²

In this case, an incoming tenant entered into possession of parts of a farm before the starting date of his lease in order to prepare the ground and he argued that this constituted the conversion of his lease into a real right. This argument was rejected by the Court of Session, which held that a real right of lease could not come into existence without *exclusive* possession.²³ As such, it can be said that exclusive possession is vital to converting a mere personal right under a contract of lease into a real right under the 1449 Act.

C. Exclusive Possession and Security of Tenure

The various Housing Acts that apply to the public and private rented accommodation sectors in Scotland define ‘tenancies’ in a different way from a contract of lease, as traditionally understood.²⁴ These Acts are responsible for the creation of Scottish secure tenancies in the public sector and assured tenancies in the private sector, although not all private rented accommodation will attract the protection of an assured tenancy.²⁵ In many circumstances, a contract for a residential statutory tenancy will still resemble a lease capable of becoming a real right using the 1449 Act, for it will possess the essential requirements for doing so. As such, exclusive possession still fulfils the role of conferring a real right upon a tenant who falls within one of the various statutory regimes.

There are, however, circumstances where a residential statutory tenancy may not resemble the requirements of the 1449 Act but still fall within the ambit of one of the Housing Acts. The Housing (Scotland) Act 2001 construes a tenancy as, ‘an agreement under which a house is made available for human habitation,’ with related terms such as ‘lease’ to be interpreted accordingly.²⁶ This terminology was broadly shared by its 1987 predecessor,²⁷ and by its private sector counterpart.²⁸ These types of tenancy grant security of tenure as long as the aforementioned statutory definitions are met, with no need to meet the four essential requirements of a *contract* of lease.²⁹ The most common types of tenancy that do not meet the common law requirements but still fulfil the definitions of the Housing Acts are agreements with no fixed duration or rent.³⁰

²² *ibid* 6 (Lord President Cooper).

²³ *ibid* 8 (Lord President Cooper).

²⁴ See text to n 5.

²⁵ The Private Housing (Tenancies) (Scotland) Act 2016 introduces a new form of statutory tenancy, and dispenses with the current assured tenancy regime.

²⁶ H(S)A 2001, s 41.

²⁷ Housing (Scotland) Act 1987, s 82 (H(S)A 1987).

²⁸ Housing (Scotland) Act 1988, s 12 (H(S)A 1988).

²⁹ McAllister (n 1) 462.

³⁰ For an example of the latter see *Andrew v North Lanarkshire Council* (Lands Tribunal for Scotland, 1 June 2011) <www.lands-tribunal-scotland.org.uk/decisions/LTS.TR.2010.10.html> accessed 12 February 2016.

The Fifth Element: Should Exclusive Possession Be Considered an Essential Requirement for the Constitution of a Contract of Lease in Scots Law?

In these tenancies, the role of exclusive possession is that of an implied prerequisite to meeting the definition of 'house', which, in both the social and private rented sector is defined as a building or part of a building intended to be occupied 'as a separate dwelling'.³¹ Additionally, there is a requirement that the tenant be an individual.³² Taken together, these three factors provide a strong inference that the tenant will be required to show exclusive possession in order to claim security of tenure. This is more explicit in assured tenancy cases where the tenant shares part of their accommodation with someone other than their landlord, as they will only have an assured tenancy if they have exclusive occupation of some part of the accommodation.³³

The powerful influence that exclusive possession can have on statutory tenancies was highlighted by the decision of the House of Lords in *Uratemp Ventures Ltd v Collins*.³⁴ Although this was an English case, the wording of the relevant statutory provision is nearly identical to that of the equivalent Scottish statute;³⁵ hence this case can be considered a highly persuasive authority on the construction of the latter. Here, a long-term hotel resident was able to claim an assured tenancy of his room because he had exclusive possession thereof and used it as his principal home. The absence of any cooking facilities was deemed irrelevant, because the Housing Act was intended, 'to protect people in the occupation of their homes, not to encourage them to cook their own meals.'³⁶

For the purposes of Agricultural Holdings Act tenancies, a lease is defined as, 'a letting of land for a term of years, or for lives, or for lives and years, or from year to year.'³⁷ As a practical point, someone claiming to be an agricultural tenant would want to rely, first and foremost, on the statutory security of tenure provisions.³⁸ By all means, it would be possible, in theory, for them to rely on the 1449 Act; nay, the 1449 Act was expressly targeted at those who work the land (*labouris the grunde*). However, given the more lenient criteria typifying the statutory regimes (which do not prescribe the same four requirements applicable to a contract of lease at common law),³⁹ there is unlikely to be any case in which the 1449 Act would be useful to the intending agricultural tenant. Those preliminary points aside, the key point to note is, once again, how important a role exclusive possession plays in the context of agricultural tenancies. In *Clydesdale Bank plc v Davidson*,⁴⁰ Lord Hope of Craighead confirmed that exclusive possession was necessary to claim security of tenure under a 1991 Act tenancy.⁴¹ This position is unlikely to have changed for the purposes of

³¹ H(S)A 1988, s 12(1); H(S)A 2001, s 11(1)(a). See *Uratemp Ventures Ltd v Collins* [2001] UKHL 43, [2002] 1 AC 301 [57]-[58] (Lord Millett) and, by analogy, in the context of an English statute, *Westminster City Council v Clarke* [1992] 2 AC 288, 300-302 (Lord Templeman).

³² H(S)A 1988, s 12(1)(a); H(S)A 2001, s 11(1)(c).

³³ H(S)A 1988, s 14(1).

³⁴ *Uratemp* (n 31).

³⁵ Housing Act 1988, s 1. Cf H(S)A 1988, s 12.

³⁶ *Uratemp* (n 31) [57] (Lord Millett).

³⁷ Agricultural Holdings (Scotland) Act 1991 s 85(1). The Agricultural Holdings (Scotland) Act 2003, s 1(1) (AH(S)A 2003) implies that this definition still applies to new agricultural tenancies.

³⁸ *ibid*.

³⁹ *ibid*.

⁴⁰ 1998 SC (HL) 51.

⁴¹ *ibid* 54 (Lord Hope of Craighead).

the Agricultural Holdings (Scotland) Act 2003, given that the latter preserves the definitions from the 1991 Act.⁴² As such, it can be stated with some certainty that the role played by exclusive possession in agricultural holdings cases is similar to its role in statutory residential tenancies.

D. Other Types of Lease

In most other types of lease (that is, those contracts not falling within one of the aforementioned statutory regimes), exclusive possession will be relevant for the purposes of the 1449 Act. However, there are some types of lease to which the 1449 Act does not apply: notably, long leases of twenty or more years.⁴³ In these cases, registration is equivalent to possession, and so will suffice to confer a real right upon the tenant.⁴⁴ Having already been afforded a real right, the tenant will not need to demonstrate exclusive possession of the subject. Quite apart from possession no longer being necessary, it is no longer *possible* to create a real right in a long lease via the 1449 Act: this can *only* be achieved by means of registration.⁴⁵

In addition, there are some contracts of lease to which the 1449 Act *does* apply but which, exceptionally, do not require *actual* exclusive possession for the creation of a real right. The Land Tenure Reform (Scotland) Act 1974 legalised the interposed lease in Scotland, which is a lease granted by a landlord to a new tenant whilst the original tenant's lease continues to subsist.⁴⁶ The new tenant is interposed between the parties to the original lease and becomes the head tenant and effective landlord to the original tenant, now a subtenant of the interposed tenant.⁴⁷ One of the effects of such a lease is that the head tenant is *legally* deemed to have entered into possession of the leased subject,⁴⁸ thus claiming a real right under the 1449 Act whilst side-stepping the need for factual exclusive possession, which will typically remain with the subtenant.⁴⁹

Scots law also recognises as leases a number of contracts relating to sporting rights, for instance fishing and shooting rights.⁵⁰ The applicability of the 1449 Act to these contracts is a subjective matter, dependent upon the particular wording of the contract under consideration. The safest way to achieve a real right in such cases appears to be to form a contract for the lease of *land* for the purposes of the activity (e.g. shooting or fishing) rather than a contract simply relating to the activity itself, as the courts have generally held the latter type of contract incapable of conferring a real right.⁵¹ In *Farquharson*,⁵² which concerned the leasing of part of the estate of

⁴² AH(S)A 2003, s 1(1). Note, also, the absence of 'lease' from s 93.

⁴³ Registration of Leases (Scotland) Act 1857, s 20B (RL(S)A 1857), inserted by Land Registration etc. (Scotland) Act 2012, s 52.

⁴⁴ RL(S)A 1857, s 16(1).

⁴⁵ *ibid* s 20C.

⁴⁶ Land Tenure Reform (Scotland) Act 1974, s17(1) (LTR(S)A 1974).

⁴⁷ *ibid* s 17(2).

⁴⁸ *ibid* s 17(1)-(2).

⁴⁹ *ibid* s 17(1); McAllister (n 1) 176.

⁵⁰ Paton and Cameron (n 2) 73-84.

⁵¹ *Pollock, Gilmour & Co v Harvey* (1828) 6 S 913.

⁵² *Farquharson* (1870) 9 M 66.

Invercauld for shooting, the Lord President stated that it was material to observe that the tenant, '[had], or [was] intended to have, under this lease the exclusive occupation of the ground.'⁵³

E. Summary

This section has demonstrated that an implied term of a contract of lease, as traditionally understood, is that exclusive possession must be granted to the tenant at the date of entry. It has also established that exclusive possession is generally required to elevate a common law contract of lease to a real right under the 1449 Act. The notion of exclusive possession also plays a role, albeit sometimes in a roundabout, incidental way, if a contract of lease is to attract security of tenure under one of the various statutory regimes. From this, it is clear that exclusive possession still plays a key role in allowing a tenant to obtain a real right of lease (or analogous security of tenure). However, none of this is necessarily to imply that exclusive possession is an essential requirement for the *classification* of a *contract* as one of lease.

3. Opinions Regarding the Role of Exclusive Possession

Aside from being used to determine whether a lease confers a real right, it has often been argued that exclusive possession is a factor in distinguishing between a contract of lease and a lesser form of contract known as a licence agreement. This practice began in earnest with a series of cases in the 1920s and then arose again in the 1990s, where it was compared, in greater detail, to the related English law. Since the turn of the millennium, the issue has arisen infrequently, so there is no modern case that can be said to have resolved it definitively. Instead, two approaches to the problem have developed, each of which focuses on a different aspect of leasing. The first is an intention-based approach that focuses on comparing the factual circumstances with what the contract implies they *should* be; the second approach focuses purely on the nature of the rights held by the tenant under the express terms of the contract relating to the property. In the twenty-first century, it appears that the second approach has become the dominant one, with the first playing a limited role in the prevention of attempts to subvert contractual terms by creating a weaker licence agreement. The relevant case law is considered below.

Before proceeding, it is important to address two definitional issues. The first is that there is no official statutory definition of a licence agreement. For the purposes of this analysis, reference will be made to the definition provided by Paton and Cameron, which states that a licence is, 'a contract, falling short of a lease, whereby not the heritage itself but a right to use a particular part of it or to put a particular part of it to some use is granted.'⁵⁴ From this definition, the possible utility of exclusive possession as a distinguishing factor is made clear, since the spirit of the Scottish licence appears to exclude control of the whole of the subject. A model

⁵³ *ibid* 69 (Lord President Glencorse).

⁵⁴ Paton and Cameron (n 2) 12.

example of a licence that highlights the situational use of such a right is a football season ticket, defined as, ‘a licence giving access to and egress from the allocated seat and a right to occupy that seat when matches are played.’⁵⁵

The second term requiring clarification is ‘exclusive possession’. The term ‘exclusive possession’ may appear below in contexts referring to what is provided in the contract or whether the tenant enjoys exclusive possession *as a matter of fact*.⁵⁶ Thus, the simple case of ‘exclusive possession’ is where the owner of property expressly grants to the occupier of his property a right to exclusive possession. On the other hand, the owner of property may stipulate in the contract between him and the occupier that the occupier does not have a right to exclusive possession, and that the owner is free to come and go as he likes. Even so, if it is clear that the landlord does not, and does not plan to, act on that right, the occupier may be deemed to have exclusive possession as a matter of reality. For current purposes, these two understandings of exclusive possession are examined side-by-side.

A. Early Twentieth Century Cases

The first significant case on whether a contract is a lease or a licence is *United Kingdom Advertising Co v Glasgow Bag-Wash Laundry*.⁵⁷ This case related to a contract for the display of advertising in part of a post-office that the advertising company had the right to use. At issue was the precise nature of that right. The laundry defended the action for payment by stating that the contract that they had was a lease. This was in their interests as, if categorised as a lease, the contract would be unenforceable due to its failure to comply with formalities of proof specifically applicable to leases. The Inner House held that the contract did not amount to a lease, with the Lord President stating that, ‘*prima facie* that contract bears no resemblance to a lease of heritable property. It is not a contract by which possession is given, as to a tenant, of any of the post-offices or of any part of them.’⁵⁸ This was supported by Lord Hunter, who stated:

There is at the best a licence to exhibit advertisements in certain post-offices subject to certain conditions. That gives the licensees no exclusive right of possession of any part of the post-offices for any purpose whatever. It gives them merely a limited personal right.⁵⁹

⁵⁵ *Joint Administrators of Rangers Football Club plc, Noters* [2012] CSOH 55, 2012 SLT 599 [30] (Lord Hodge).

⁵⁶ Although there is a general rule against construing a contract by reference to subsequent conduct, this has been noted as possible in the present context where there is an ambiguity: *Scottish Residential Estates Development Co Ltd v Henderson* 1991 SLT 490, 491-2 (Lord Dunpark). See, generally, W McBryde, *The Law of Contract in Scotland* (3rd edn, W Green 2007) paras 8.30-8.31. For an analysis of the English approach in this context, see ACL Davies, ‘Sensible Thinking About Sham Transactions’ (2009) 38(3) ILJ 318, 320-22.

⁵⁷ 1926 SC 303.

⁵⁸ *ibid* 306 (Lord President Clyde).

⁵⁹ *ibid* 307 (Lord Hunter).

From this, it appears that, at the very least, the Inner House considered exclusive possession to be a feature absent from licences. In other words, a licence was a lesser personal right because it would not have been possible to convert this into a real right under the 1449 Act. This position would be further explored in a series of valuation cases starting with *Broomhill Motor Co v Assessor for Glasgow*.⁶⁰ *Broomhill* concerned agreements relating to the occupation of lock-up garages. The occupants were provided with keys, but the owners were permitted under the agreement to enter the premises for inspection and cleaning purposes. This arrangement was described by the Court of Session as, 'a temporary allocation of space for a limited purpose and for a limited time.'⁶¹ It did not involve a parting with the occupation of the heritage on the part of the proprietors. In contrast, *Chaplin v Assessor for Perth*⁶² established that the contracts in question, whereby garages were rented out by an individual with no connections to the motor trade by oral agreement (terminable upon a week's notice), were leases. The key factor distinguishing *Broomhill* from *Chaplin* was that, in the latter case, no services were performed by the garage owner, nor were there provisions in the contract sanctioning such use of the property.⁶³ In other words, the occupiers of the garages in that case had been afforded exclusive possession thereof.

These two cases, along with *United Kingdom Advertising Co*, used objective criteria to determine whether or not a lease existed. This makes their use of exclusive possession understandable given that it can clearly indicate who is the user of the heritage in question. If the purported landlord still has a significant amount of access to the property, it follows that the other party will be more a limited occupier than a tenant. From these cases alone, it would appear that the courts place value on exclusive possession in distinguishing leases from licence agreements. However, certain cases decided in the late twentieth century appear to follow an alternative route.

B. Late Twentieth Century Cases

One case to adopt a different approach to the lease-licence distinction was that of *Bradford Properties Ltd v British Telecom Plc*.⁶⁴ Here an attempt was made by a tenant to subvert a prohibition against subletting in a lease by entering into an agreement with the features of a lease but language suggesting that it was a licence. The tenant argued that there was no sub-lease because the alleged sub-tenant did not enjoy exclusive possession of the subjects under the disputed agreement; in this regard, the tenant relied on the English authority of *Street v Mountford*.⁶⁵ The Lord Justice-Clerk stated that *Street* was of no assistance in the current matter because it only related to the law of England.⁶⁶ This indicates a rejection of exclusive possession as a deciding

⁶⁰ 1927 SC 447.

⁶¹ *ibid* 454 (Lord Hunter).

⁶² 1947 SC 373.

⁶³ *ibid* 379 (Lord Russell).

⁶⁴ 1992 SC 12.

⁶⁵ *Street* (n 12).

⁶⁶ *Bradford* (n 6) 19 (Lord Justice-Clerk Ross).

factor in distinguishing between a lease and licence.⁶⁷ Instead, the Lord Justice-Clerk stated that the agreement itself must be closely scrutinised and that the court was entitled to consider whether the purpose of the agreement was to defeat the provision against subletting in the original contract of lease when reaching a conclusion as to the nature of the new contract.⁶⁸ It was therefore held that the agreement in question amounted to a sub-let impermissible under the original contract of lease.⁶⁹

The possible emergence of an intention-based test, in place of a more objective test involving exclusive possession, is also supported by the earlier case of *Scottish Residential Estates Development Co Ltd v Henderson*.⁷⁰ Here, a property development company allowed a woman to occupy a cottage for an indefinite period of time. When they subsequently asked her to vacate it, she refused on the grounds that her agreement with them constituted a lease: all the essential elements existed bar a definite ish, which would be inferred by the court if the other elements were present, a position supported by the case of *Gray v Edinburgh University*.⁷¹ In the court's opinion, Lord Dunpark stated:

As appears from *Gray*, before an ish may be inferred, there must be a lease and that means that the terms of this offer and acceptance must be such as to constitute a lease and nothing other than a lease. In our opinion that depends upon whether on a sound construction of these documents the pursuers intended to create a tenancy in favour of the defender and the defender accepted occupation of this cottage qua tenant[.]⁷²

The opinion went on to state that the only reasonable construction of the offer was to view it as a contract for the use of the property until the developers needed it, and that this could not amount to a lease.⁷³ However, were this case to be examined without taking the intentions of the parties into account, all of the essential requirements of a lease could be seen, counsel for the pursuers having accepted the payment of repair bills and an obligation to maintain the surrounding woodlands as rent,⁷⁴ and it being possible to infer an ish using the aforementioned principle in *Gray v Edinburgh University*.

Both *Bradford* and *Scottish Residential Estates* support the proposition that the intention of the contracting parties ought to be the court's primary concern in determining the existence of a lease or licence, rather than the presence or absence of a right to exclusive possession under the contract. In *Scottish Residential Estates*, the defender exclusively occupied the cottage,⁷⁵ yet the missives between the two parties did not amount to a lease. To have held this would have been to frustrate the

⁶⁷ *ibid.*

⁶⁸ *ibid* 20 (Lord Justice-Clerk Ross).

⁶⁹ *ibid* 23 (Lord Justice-Clerk Ross).

⁷⁰ 1991 SLT 490.

⁷¹ 1962 SC 157.

⁷² *Scottish Residential Estates* (n 56) 491 (Lord Dunpark).

⁷³ *ibid* 492 (Lord Dunpark).

⁷⁴ *ibid* 491 (Lord Dunpark).

⁷⁵ *ibid.*

intentions of the developers, who had simply made a goodwill gesture while they did not need the land themselves. In *Brador*, accepting the defender's propositions would have defeated the purpose of the contractual provisions agreed to by both parties under the original contract of lease. These cases both create a strong argument for disregarding exclusive possession as a litmus test, on the grounds that it would limit parties' freedom to contract. For example, if A allowed B to occupy an unused second home that A owned for a fee until B found suitable alternative accommodation, that agreement could be converted into a real right of lease by inferring an ish of one year, thereby completely disregarding the parties' original intentions and forcing the obligations of a landlord upon A. Additionally, in cases similar to *Brador*, where a tenant appears to be giving a lesser right to a third party, the original contract of lease between the tenant and landlord could be undermined if the tenant confers too great a right on the third party (i.e. a sub-lease), whether he or she does so inadvertently or by design.

C. The 1999 hostel cases

In 1999, two cases regarding hostel accommodation raised the issue of exclusive possession. The first of these, *Conway v Glasgow City Council*,⁷⁶ concerned a woman who was evicted from her hostel accommodation without notice as a result of an alleged incident that took place earlier that same day. She claimed that she had a common law tenancy and therefore required reasonable notice before she could be evicted. *Esto*, it was argued that, even if she were not a tenant, the Rent (Scotland) Act 1984 would prevent the termination of her occupancy without notice.⁷⁷ The Sheriff was unconvinced by the argument that the pursuer was a tenant. Counsel for the defender argued that exclusive possession was an essential condition of a contract of lease, distinguishing the arguments in *Brador* that suggested otherwise. This was based on the ground that no issue was taken in that case as to the supposed requirement of exclusive possession and that, instead, *Brador* was concerned with the question of what needed to be exclusively possessed.⁷⁸ Sheriff Gordon was persuaded by these arguments:

Similarly, the law has come increasingly to talk of exclusive possession as a necessary condition of a lease, as can be seen in the *Brador* case itself and in other cases such as *Chaplin* and the *Commercial Components* case, where, whatever the sheriff may have been prepared to infer, the sheriff principal said that exclusive possession of the subjects was one of the badges of a lease. Counsel for the pursuer, of course, says that this applies only to leases covered by the 1449 Act and not to common law leases. I accept that there are common law leases, but that does not mean that they do not require exclusive possession.⁷⁹

⁷⁶ *Conway* (n 6).

⁷⁷ Rent (Scotland) Act 1984, s 23(2A).

⁷⁸ *Conway* (n 6) 26 (Sheriff Gordon).

⁷⁹ *ibid*, citing *Commercial Components (Int) Ltd v Young* 1993 SLT (Sh Ct) 15, 17 (Sheriff Principal Hay).

Commercial Components (Int) Ltd v Young,⁸⁰ one of the cases referred to by Sheriff Gordon, was a Sheriff Principal appeal concerning an agreement for the occupation of two fields, which was held by the original Sheriff to be a grazing lease converted into an agricultural tenancy because there was never any removal of stock.⁸¹ The Sheriff Principal held that the Sheriff erred in law when determining the nature of the parties' agreement. The presence of ponies owned by the pursuers on the land, which denied exclusive possession to the defenders, was an important factor in determining the non-existence of a statutory agricultural tenancy.⁸²

If the Sheriff's judgment on the first issue raised in *Conway* is to withstand scrutiny, the position will represent a radical step forward in how a lease is perceived. However, there are problems with this judgment.

Firstly, no comments are made to substantiate the point that common law leases require exclusive possession. The *Clydesdale Bank* judgment makes it clear why such a requirement matters in relation to a statutory agricultural tenancy: the protection of the Agricultural Holdings Act can only be enjoyed if one has a 'real right of tenancy', obtained by way of exclusive possession.⁸³ However, all the statutory and case law evidence points to possession only being relevant for the acquisition of such a real right (or analogous security of tenure). Additionally, the comment from *Commercial Components* regarding exclusive possession being a badge of a lease appears to have been stretched in its interpretation. It is submitted that what the Sheriff Principal meant was that exclusive possession is an obvious outward sign that a contract is likely to be a lease, although it may still fail to be one. The classic example here is *Mann v Houston*, where exclusive possession was present but the contract was not a lease due to a lack of rent.⁸⁴ The position taken in that case therefore is not a novel one for Scots law, but rather a restatement of what has come before.⁸⁵ As a final note on this case, Sheriff Gordon stated that *Brador* did not explicitly dispute the general proposition that exclusive possession was required in leases.⁸⁶ Although this may be true, only sufficient possession as to amount to a sub-let, and thereby to rule out the possibility of a mere licence agreement, was required to satisfy the court in that case,⁸⁷ as against a requirement that the tenant enjoy *exclusive* possession. This, combined with the explicit rejection of *Street v Mountford*, would suggest that the Inner House had a larger margin of appreciation in mind and that it did not tacitly insist upon exclusive possession as a dogmatic requirement. The language of 'sufficient' possession suggests that there is a degree of possession – less than exclusive – at which a licence agreement would be ruled out by a court.

The second hostel case was *Denovan v Blue Triangle (Glasgow) Housing Association Ltd*.⁸⁸ The facts here were similar to those in *Conway*: a young homeless person was summarily evicted from his hostel accommodation. He subsequently

⁸⁰ *Commercial Components* (n 79).

⁸¹ *ibid* 15 (Sheriff Principal Hay).

⁸² *ibid* 18 (Sheriff Principal Hay).

⁸³ See text to nn 40-41.

⁸⁴ 1957 SLT 89.

⁸⁵ Paton and Cameron (n 2) 14.

⁸⁶ *Conway* (n 6) 26 (Sheriff Gordon).

⁸⁷ *Brador* (n 6) 22 (Lord Justice-Clerk Ross).

⁸⁸ 1999 Hous LR 97.

argued that he had a common law tenancy (i.e. a common law lease) as a gateway to obtaining an assured tenancy under the 1988 Act, since 'tenant' is undefined in that Act. However, a material difference in facts between *Denovan* and *Conway* is that, in the former case, the pursuer had exclusive occupation of a room alongside access to shared facilities, while the pursuer in *Conway* was in a two-bed room where the hostel owners were entitled to place another occupant in the second bed without her consent.

Despite *Denovan's* exclusive occupation of the room, it was held that his residency agreement did not contain terms that would prove a lease at common law.⁸⁹ This is significant, as, if he was a common law tenant, he would be able to use the exclusive possession of his room to claim security of tenure under the 1988 Act despite sharing the property with other people.⁹⁰ However, to do so, it first had to be established that the four traditional requirements of a common law lease were present.

Whereas *Conway* suggested that exclusive possession was a requirement for a contract of lease, *Denovan* favoured the opposite, more traditional, view of matters. The latter case suggests that, in determining whether a contract of lease exists, recourse is had to the first four elements only, with exclusive possession becoming relevant only *after* it has been established that a contract of lease exists (for instance as a prerequisite for statutory security of tenure). In contrast, the Sheriff in *Conway* used exclusive possession as a fifth element of a contract of lease (as opposed to a mechanism for claiming statutory protection), taking it into consideration at the same time as the other four requirements. As such, *Denovan* makes it easier to distinguish *Conway* as deviating from accepted practice. The view expressed in *Conway*, namely that exclusive possession was required to determine the existence of a contract of lease, was one for which more evidence would arise in later years. *Denovan's* continuation of the traditional view of exclusive possession would not prevent the ideas seen in *Conway* from resurfacing in later cases.

D. Twenty-First Century Cases

Since the turn of the millennium, two cases have considered exclusive possession in detail. In *South Lanarkshire Council v Taylor*,⁹¹ the defender occupied a racecourse under a contract titled as a grazing lease, which required her to vacate the area to allow events to take place if given twenty four hours' notice. The pursuers sought declarator that no valid lease existed. This was due to the defender's inability under the contract to exclude other persons from the land for the entirety of her lease. Implicit in this argument was the point that if the tenant were able to exclude other persons under the contract, she could be deemed a tenant. In other words, the argument was founded on the concept of exclusive possession being a requirement of a lease. Lord President Cullen made the following statement on the matter:

⁸⁹ *ibid* 109 (Sheriff Drummond).

⁹⁰ *ibid* 111 (Sheriff Drummond). HSA 1988, s 14(1).

⁹¹ *South Lanarkshire Council* (n 13).

In the light of these arguments we consider that it would be going too far and too fast to decide that merely by reason of the terms of cl 4 [i.e. the clause requiring her to vacate the area at short notice] the defender could not establish that the character of her occupation under the alleged agreement was that of a tenant under a lease. The authorities to which we have been referred indicate that a limited reservation in favour of the landowner or a limitation in the nature of the use to which the occupier can [put] the land would not necessarily be inconsistent with the existence between them of the relationship of landlord and tenant. Plainly, however, it would be a matter of degree, according to the circumstances of the individual case.⁹²

This opinion is consistent with the decisions in the valuation cases of the 1920s in prioritising the *content* of the agreement itself, namely whether it allows the supposed tenant to possess the property exclusively, over the parties' express contractual intentions, for two persons could, quite conceivably, describe themselves as 'tenant' and 'landlord' even if the contract's terms do not afford the avowed 'tenant' exclusive possession. What emerges from *South Lanarkshire* is an objective approach for reconciling the decision in *Broomhill* (where extensive access by the owner was seen to preclude the possibility of a lease) and the decision in *Chaplin* (where negligible interference by the owner was viewed to permit the possibility of a lease). Together, *Broomhill*, *Chaplin* and *South Lanarkshire* create the impression that it is not simply a rigid requirement of exclusive possession, enjoyed by the tenant, that is the deciding objective factor in the classification of a contract as a lease, but instead a certain *degree* of possession, the sufficiency of which is to be judged on a case-by-case basis. It is for these reasons that *South Lanarkshire* marks a return to the more objective reasoning visible in the earlier valuation, and more recent hostel, cases. *Brador* is referred to only to support the proposition that it is enough for the tenant to obtain certain uses of land to have sufficient possession not to be deemed the mere holder of a licence; that case is not invoked for its contractual, intention-based approach to the distinction between the different rights.⁹³

The decision in *South Lanarkshire* was followed in the Land Court case of *Cameron v Alexander*.⁹⁴ Here, the purported tenant sought declarator that he was entitled to possess the whole of a leased steading, subject to reservations of certain parts thereof by the landlord. The landlord objected, arguing that the agreement only constituted a licence agreement due to his right to use certain parts of the land. The landlord's argument was strongly founded on Sheriff Gordon's decision in *Conway*, but the court here found little persuasive authority for a fifth cardinal feature of a contract of lease, instead calling exclusive possession an 'important pointer' in determining the nature of an indeterminate agreement.⁹⁵ With this in mind, it was held that exclusive possession was not an absolute requirement for the creation of a contract of lease in a non-statutory, agricultural context.⁹⁶ The court in *Cameron* did, however, conclude that occupation of a residential property by tenants

⁹² *ibid* [7] (Opinion of the Court).

⁹³ *ibid* [6] (Opinion of the Court).

⁹⁴ *Cameron* (n 13).

⁹⁵ *ibid* [53] (Opinion of the Court).

⁹⁶ *ibid* [58] (Opinion of the Court).

under different leases, as against co-tenants under the same lease agreement,⁹⁷ could be such an impediment to the enjoyment of the property that an absolute rule of exclusive possession would be necessary in that context.⁹⁸ However, the Land Court had no difficulty in providing examples, both historical and modern, of situations where shared use would not impact on the legitimacy of a contract of lease in a residential context.⁹⁹

E. Academic Opinion

The uncertainty surrounding the role of exclusive possession is, in many cases, reflected in literature. This uncertainty has been observed from a comparative perspective. In a study of time-limited land interests of various legal systems, the Scottish stance on exclusive possession was described as 'not entirely free from doubt,'¹⁰⁰ although it was noted later in the text that it was most likely that a common law lease would be found where exclusive possession existed.¹⁰¹ Meanwhile, McAllister wrote that, although Rankine allowed for leases that only granted specific uses of land, such as leases for sporting or mineral rights, it did not follow that exclusive possession, '[could] be dispensed with in the case of leases (the vast majority) that [fell] outwith these special categories.'¹⁰² From this, an inference can be drawn that a lack of exclusive possession is the exception rather than the rule in a modern context. Thus, although some forms of lease can exist without exclusive possession, those comprise a narrow group of contracts with arguably limited relevance to modern law.

One of the principal arguments against exclusive possession as an essential requirement is that, as discussed in the second section, its role is largely confined to obtaining a real right or security of tenure. The lease, at its heart, is a personal contract, a statement oft-repeated in legal literature.¹⁰³ The concept of the personal lease was brought up recently in an article, analysing the decision in *Clydesdale Bank v Davidson*,¹⁰⁴ where it was noted that a personal lease, 'is no less a lease than a lease that has been made a real right,' and that such a common law lease is still capable of coming under a statutory regime, thus conferring protection akin to that of a real right.¹⁰⁵ The possibility of a personal lease that does not come under a statutory regime is left open, which would allow for a type of lease where exclusive possession has no bearing at all.

⁹⁷ See text to nn 7-9.

⁹⁸ *Cameron* (n 13) [59] (Opinion of the Court).

⁹⁹ *ibid.*

¹⁰⁰ C Van der Merwe and A Verbeke, 'Scotland' in C Van der Merwe and A Verbeke (eds), *Time-Limited Interests in Land* (CUP 2012) 118.

¹⁰¹ *ibid.*

¹⁰² McAllister (n 1) 54.

¹⁰³ Rankine (n 3) 6; Paton and Cameron (n 2) 5.

¹⁰⁴ D Bain and C Bury, 'A, B and C to A, Revisited' [2013] *Jur Rev* 77.

¹⁰⁵ *ibid.* 82.

Lord Gill disputes the arguments expressed in the latter article, arguing that the common law personal lease no longer exists in Scots law.¹⁰⁶ If this is the case, and only the ‘real right’ (1449 Act) lease remains, the court’s findings in *Millar v McRobbie* will require a tenant to take full and exclusive possession of the leased subjects before a ‘lease’ (as Lord Gill understands the term) can be said to exist. In support of his argument, Lord Gill examines the historical background of the ‘personal right’ theory and concludes that, by the time that *Millar* was decided, the Scots law position was that the tenant had a real right that was made effective by taking possession.¹⁰⁷ From his arguments, it would appear that Lord Gill interprets the opening statement of Lord President Cooper in *Millar* literally, with the prospective tenant not being a tenant at all until possession is taken, instead being a mere personal creditor who is ‘owed’ a tenancy. In response it might be countered that an agreement under which a person is owed a tenancy is still effectively a contract of lease, even if not explicitly so; nor does Lord Gill’s critique completely exclude the possibility of a lease without possession, as this could still be achieved with a registered long lease.¹⁰⁸

It is in the residential sector, more than any other, that exclusive possession is talked of as a requirement of a contract of lease. This may account for the reluctance of the Land Court in *Cameron v Alexander* to apply their judgment outside of an agricultural context. As an illustration of this at a basic level, the Scottish Qualifications Authority has seen fit to consider exclusive possession an essential requirement for a Scots law tenancy in their Higher National Unit on security of tenure, although they do note that this is ‘particularly contentious.’¹⁰⁹ It should be noted that the unit focuses solely on statutory residential tenancies in the context of security of tenure, where exclusive possession is clearly an important factor,¹¹⁰ so this could be viewed as a simplification for practical purposes. This also appears to conflict with the decision in *Denovan*, where exclusive possession was not considered a necessary element for the establishment of an assured tenancy.

In a 2015 article on ancillary rights in leases, Mike Blair offers a wider perspective on the matter.¹¹¹ Although Blair posits exclusive possession as a necessary condition of a lease, he allows for multiple leases to exist over the same space, each granting exclusive possession of a particular *use* of the land to their tenant. For example, an area of farmland with access to a river could be leased as an agricultural holding to one tenant, while another could have a fishing lease over the river. It is only in cases where two people share the *same use* of land that a lease would be found not to exist. This interpretation of exclusive possession appears wholly consistent with Rankine, who construes the term sufficiently broadly for it to include exclusive uses of land. However, exclusive use is materially different from exclusive possession, as it allows for cases where tenants do not necessarily have the

¹⁰⁶ B Gill, ‘Two Questions in the Law of Leases’ in F McCarthy, J Chalmers and S Bogle (eds), *Essays in Conveyancing and Property Law in Honour of Professor Robert Rennie* (Open Book Publishers 2015) 274.

¹⁰⁷ *ibid* 272.

¹⁰⁸ See text to nn 43-45.

¹⁰⁹ SQA, ‘National Unit Specification: Housing Law for Advisors: Introduction to Security of Tenure’ (20 June 2007) 7 <www.sqa.org.uk/files/hn/DF1E633.pdf> accessed 17 February 2016.

¹¹⁰ See text to nn 24-36.

¹¹¹ M Blair, ‘Ancillary Rights as Real Rights’ (2015) 60(10) *JLSS* 32.

right to remain on the property at all times, even if they are the only ones capable of exercising a certain right. As such, although Blair's arguments give the impression that exclusive possession is a pre-requisite of a lease, he does not always mean exclusive possession in the same way that it has been discussed in the case law, where it has always been referred to in terms of *occupation* as opposed to exclusive possession of a right, such as fishing, exercisable upon the land.

4. Exclusive Possession as a Requirement

With the current role of exclusive possession determined as a common, if not defining, feature of a contract of lease, and the main opinions on the subject established, it is important to ask why such a requirement might exist in the Scots law of leases. It is submitted that there are two main arguments in favour of such a requirement. The first is the desire for legal certainty that is so valued in property law.¹¹² This is a desire that has motivated many developments in the sector, catalysed by the need to know exactly what rights are possessed by anyone with some form of interest in property, especially heritable property. The second argument in favour of an exclusive possession requirement is that, in terms of time-limited rights relating to property, the stakes associated with distinguishing leases from lesser rights such as licences are higher than they have ever been before. When Rankine wrote of the broad definition of the Scottish lease,¹¹³ he could never have imagined the concept of security of tenure, which offers greater protection than even a real right by preventing landlords from terminating a statutory tenancy without a specific legal ground such as failure to pay rent.¹¹⁴ In this context, it is easy to see how an argument in favour of exclusivity in leases might be formed. However, there are still a number of opposing arguments and potentially better alternatives that could be explored.

A. Legal certainty

Legal certainty is perhaps the strongest argument that can be made in favour of a requirement for exclusive possession. It is certainly a value by which the Scottish Law Commission sets much store:

For people to be able to arrange their affairs in an orderly manner, the rules of property law must be clear and predictable; and that is as important in business as in private life. In property law, certainty is prized above all other virtues.¹¹⁵

If the law is left in a state of ambiguity, it generates confusion and leads to an increase in litigation. This is certainly evident as regards exclusive possession: each

¹¹² Scottish Law Commission, *Discussion Paper on Sharp v Thomson* (Scot Law Com DP No 114, 2001).

¹¹³ Rankine, (n 3) 2.

¹¹⁴ See, for example: H(S)A 2001, s 12 and sch 2.

¹¹⁵ Scottish Law Commission, *Discussion Paper on Sharp v Thomson* (n 112).

broad set of cases has presented a different approach to the issue, none of them providing a definite resolution. The case most favourably cited is *Brador*. However, *Brador* is unsuitable as a definitive answer in that it is not directly concerned with exclusive possession but instead the intentions of the parties and the degree of possession necessary for a sub-lease to exist. This allowed for the decision in *Conway*.

By settling this issue in the positive and establishing exclusive possession as a requirement for the characterisation of a contract as a lease, the uncertainties regarding this area of law would be brought to an end, creating an unambiguous distinction between leases and licences. Contracting parties would be in no doubt that, in order for an agreement to be classified as a contract of lease, the owner of the property must grant the occupier exclusive possession thereof; nor, under the approach propounded here, would there be any doubt that an agreement granting such exclusive possession, and which fulfilled the other four, uncontroversial elements of a contract of lease (parties, property, rent and ish), would be *sufficient* to create a contract of lease. On the other hand, this argument could be undermined by the existence of agreements that grant exclusive possession but which are not intended as leases. An example of this can be found in the unreported Sheriff Court judgment in *Caird Properties v Elbrask Contractors*.¹¹⁶ There, Elbrask Contractors entered into an agreement to cultivate an area of farmland and establish an initial crop. They later refused to vacate the land, arguing that the contract was a valid lease. The Sheriff at first instance ruled that the contract was in fact a *locatio operis faciendi*: a contract for services. This ruling was upheld on appeal to the Sheriff Principal. Paisley and Cusine note that there is no rule precluding such contracts from being carried out within the land of the instructing party, thus creating a contract that blends elements of an occupancy agreement with those of a services contract.¹¹⁷ Here, there was no question as to whether the agreement was a lease, since the defenders failed to establish a provision for the payment of rent, only attempting to argue that such a provision existed once the case was appealed. However, if such a contract had included some form of rent provision, it is entirely possible that it would have been held to constitute a lease even if that had not been the parties' intention, as per the *ratio* in *Brador*. If exclusive possession were to be used as such a decisive, pivotal element of a Scots law contract of lease, agreements of the type encountered in *Caird Properties*, or indeed any agreement vaguely resembling a lease, would have to be extremely well-drafted to avoid granting the occupant of another's property more rights than either party originally intended.

Having exclusivity as a requirement could also be argued to offer tenants a reassurance that parts of the leased property that they do not use will not be let to someone else. This scenario is suggested by the editor's note attached to the account of *Conway* in the Housing Law Reports as a reason why the decision in that case was generally correct.¹¹⁸ Otherwise, tenants would have no legal recourse if a landlord let the same subjects to multiple people.¹¹⁹ It is submitted that this argument is weak, as

¹¹⁶ (Peterhead Sheriff Court, 25 March 1993). The judgment text can be found in: RRM Paisley and DJ Cusine, *Unreported Property Cases from the Sheriff Courts* (W Green 2000) 1.

¹¹⁷ Paisley and Cusine (n 116) 2.

¹¹⁸ *Conway* (n 6) 31 (Sheriff Gordon).

¹¹⁹ *ibid*.

The Fifth Element: Should Exclusive Possession Be Considered an Essential Requirement for the Constitution of a Contract of Lease in Scots Law?

exclusive possession is, if not a constitutive requirement of a contract of lease, certainly an implied term of such a contract.¹²⁰ As such, were a landlord to attempt to break up an already-let property further, the original tenant would be able to pursue them for damages or even rescission of the contract,¹²¹ which reduces the need for an explicit statement on the matter.

B. Security of Tenure

The statutory innovation of security of tenure was first introduced into Scots law by various Rent Acts beginning in 1915. Now, it is present in the agricultural, public and private rented sectors and offers more protection to tenants than even a real right of lease does, allowing them to remain in occupation after the end date of their tenancy.¹²² With this in mind, it is no wonder that property owners have attempted to argue for the necessity of exclusive possession. This is because the establishment of a statutory tenancy could cause them to lose possession of their property for some time, or be unable to sell their property except at a heavily discounted price. A requirement of exclusive possession would restore a certain amount of power to landlords; there would be less risk of an agreement falling into a statutory regime by reason of ambiguous drafting or an error on the part of the landlord.

That said, it is arguable that security of tenure was intended to be easily accessible. In respect of the social housing sector, Adrian Stalker wrote that the specific statutory definition included in the Housing (Scotland) Act 2001 was significant because it precluded attempts by social landlords to subvert the Act by creating a lesser contract, regardless of whether the resulting tenancies were common law leases or not.¹²³ Given that this definition is also used in the private sector, it appears that the Scottish Parliament deliberately opened up the definition of tenancy for the purposes of the 2001 Act in order to extend the opportunity to as many as possible and to prevent landlords from exploiting their tenants. The looser requirements of a statutory residential tenancy also raise the point that the statutory tenancy is an entirely different species of contract from the original form of lease and a statutory residential tenancy benefits from the use of exclusive possession in a way that other leases do not. Thus, imposing a requirement of exclusive possession on all leases because the Housing Act tenancies would benefit from it would appear to create uniformity for its own sake, rather than for a practical benefit.

C. 'Rent' Asunder: A Potential Tear in the Logical Fabric of a Lease

One argument against exclusive possession as a requirement for the classification of a contract as a lease is that 'rent' (assuming Rankine's definition thereof is adopted) ought to be capable of definition independent of other requirements for the creation of a lease. This is particularly the case with exclusive possession since Rankine defined rent as being given *in return* for the possession and use of the leased

¹²⁰ Rankine (n 3) 200; Paton and Cameron (n 2) 127.

¹²¹ Rankine (n 3) 217.

¹²² McAllister (n 1) 467.

¹²³ A Stalker, *Evictions in Scotland* (Avizandum 2007) 92.

subjects.¹²⁴ His definition is imperfect as it fails to take into account the Registration of Leases Act 1857, which allows possession to be substituted with registration, a point raised in an argument against the judgment in *Clydesdale Bank v Davidson*:

Any general definition of ‘rent’ would have to be able to accommodate leases comprising real rights by means of the 1449 Act, and leases comprising real rights by means of the 1857 Act. It cannot therefore be absolutely predicated upon the taking of possession.¹²⁵

Following this line of reasoning, it would appear that, at the bare minimum, Rankine’s definition should be amended so that rent can be understood to be payable in return for possession and use in the context of short leases and in return for ‘registration and use’ in the context of long leases. However, this argument could be countered with the proposition that, since the 1857 Act provides that registration is equivalent to possession for the purpose of acquiring a real right,¹²⁶ the word ‘possession’ in Rankine’s text impliedly encompasses ‘registration’. Were this matter to be brought before a court, it appears unlikely that it would read Rankine’s definition narrowly to be an exception to the statutory rule when, in every other case, the agreement would satisfy the criteria for a long lease. In both cases, rent is exchanged in return for protection against a singular successor (whether achieved via possession or registration) and the right to use the property.

However, if Rankine’s definition of rent is accepted and it is understood as being given in return for possession or registration, a different argument can be ventured against allowing either of these to be a requirement of a lease.

As regards registration, for a lease to be eligible under the 1857 Act, it must first be a contract of lease, satisfying the four traditional requirements: parties, property, rent and ish.¹²⁷ If the definition of rent were tied to registration as a fifth element, the establishment of a long lease would be logically impossible, because registration of a lease would be required in order to register a lease to provide protection against singular successors.

If the registration of a lease is dependent on the requirements of a contract of lease being present, and rent cannot be established without registration, the establishment of a long lease becomes a self-validating circular definition that provides no way of actually establishing such. This is as unhelpful for practical purposes as it would be to define a property right as a right capable of being a property right.¹²⁸

It should be noted that this problem does not arise with exclusive possession itself as per the 1449 Act, as the taking of physical possession is an easy enough matter to perform and prove. In contrast, registration is a formal process with specific statutory requirements, which relies on there being a personal contract of lease to be given real right status. With exclusive possession as a requirement of a

¹²⁴ Rankine (n 3) 309.

¹²⁵ D Bain, ‘The Use and Management of Common Property in Scots Law’ (PhD Thesis, University of Aberdeen 2014) 281.

¹²⁶ RL(S)A 1857, s 16.

¹²⁷ *ibid* s 1.

¹²⁸ See *National Provincial Bank v Ainsworth* [1965] AC 1175, 1226.

lease, there would no longer be a distinction between a lease conferring a personal right and one conferring a real right, as their requirements would be identical.

The retention of the distinct personal lease allows for the definition of rent to be construed more widely as being given in exchange for what will be given under a real rights lease, as opposed to what is already there, which is impossible for a long lease. Removing that distinction would not be fatal to a lease under which the real right is conferred by exclusive possession. However, it would be fatal to the long lease, where the real right is conferred by registration. As such, it is submitted to be in the interests of consistency and simplicity that the four requirements of a contract of lease are kept as they are.

Further, tying rent to another requirement would also be inconsistent with the other three requirements for a common law lease, all of which are capable of being identified independently due to being physical or legal objects or identities or a contractual term that can either be explicitly provided or implied.¹²⁹ As such, the introduction of a possessory requirement could overcomplicate the process of establishing a *prima facie* lease.

5. Future Legal Developments and Possible Alternatives to Exclusive Possession

On 17 March 2016, the Scottish Parliament passed the Private Housing (Tenancies) (Scotland) Bill, which received Royal Assent on 22 April of the same year.¹³⁰ A commencement date is to be confirmed. This Act introduces the private residential tenancy ('PRT'), intended to replace the current assured tenancy regime.¹³¹ The Act has its own criteria for determining the existence of a tenancy thereunder, further distinguishing the residential tenancy in the private sector from the common law model. This continuing pattern of residential tenancies breaking away from the common law lease highlights the need to find an alternative to exclusive possession as a requirement; a new requirement that is sufficiently broad to encompass all the different types of lease without alienating any of them. In this section, some possible alternatives are considered.

A. The Private Housing (Tenancies) (Scotland) Act 2016

Under the 2016 Act, a PRT is created where a property is let to a tenant as a separate dwelling, which the tenant occupies as their only or principal home.¹³² Furthermore, the Act explicitly extends the meaning of tenancy to include agreements where there would be a tenancy but for the absence of an ish or where an agreement to pay rent is subsequently removed or ceases to have effect.¹³³

¹²⁹ *Gray* (n 71).

¹³⁰ Private Housing (Tenancies) (Scotland) Act 2016.

¹³¹ *ibid* s 1.

¹³² *ibid*.

¹³³ *ibid* s 4.

The former measure – the liberal definition of a PRT – is of more significance, as it completely removes one of the essential requirements of the common law lease and the 1449 Act from the equation, namely the *ish*. This measure explicitly removes the PRT from the common law model to the extent that a technical provision is required in the Act to ensure that the PRT is included within the meaning of the word ‘tenancy’ in previous legislation.¹³⁴ With this new development, the various private residential leases under numerous Acts become more difficult to reconcile with other types of lease because of their departure from the four essential requirements of the common law lease. This indicates that attempts to treat all leases using the same exclusivity criteria are counterintuitive, and that a new distinguishing criterion must be sought to create a clear boundary between the diverse range of leases and licence agreements.

The latter measure – extending the meaning of tenancy – is designed to prevent landlords from denying the protections of a statutory regime to tenants via later modification, further reinforcing Stalker’s proposal that the Scottish Parliament follow a liberal statutory framework for the creation of private residential tenancies.¹³⁵

B. Possible Alternatives to Exclusive Possession

Even if exclusive possession is not the litmus test between a contract of lease and a licence agreement, any sort of test used to distinguish between these statuses still ought to examine the necessary degree of control. The reasoning behind using a test, which would be applied on a case-by-case basis, rather than a fifth requirement, which would be applied in the same way in all cases, is that, with a test, there is greater scope for creating reservations; this approach better preserves the parties’ freedom of contract. By accounting for the post-contract reality, it would also provide greater protection against ‘sham’ agreements, in the broad sense of the term.¹³⁶ It is therefore submitted that a starting point for a test is to examine *how much* control the landlord retains. Strictly speaking, this is what the ratio in *South Lanarkshire Council v Taylor* addressed, as the case centred on a provision requiring the tenant to vacate the property at any time within twenty-four hours’ notice by the landlord and whether or not that prevented the contract from being a lease.¹³⁷ As such, it could be suggested that a ‘deprivation of use’ test be employed, in line with this case and the valuation cases. Under such a test, the primary focus would be the nature of the landlord’s continuing rights in the leased property and how much they interfere with the rights of the tenant. If it could be shown that the landlord was sufficiently deprived, the contract would be considered a lease rather than a licence, provided that all the other requirements for a lease were present or could be inferred via *Gray*. This would be carried out on a predominantly objective basis, looking at the terms of the contract, in order to create the least amount of ambiguity and to

¹³⁴ *ibid* s 5.

¹³⁵ Stalker (n 123) 92.

¹³⁶ Note, in this regard, the English law of leases, where the courts have adopted an expanded approach to the ‘sham’ doctrine, known as the doctrine of pretence: Davies (n 56) 320-22.

¹³⁷ *South Lanarkshire Council* (n 13) [7] (Opinion of the Court).

generate legal certainty. However, to obviate agreements that do not reflect the reality, which is a possibility in light of the nature of this proposal, provision could be made to 'to look beyond the formal written documentation to determine the true nature of the agreement between the parties'.¹³⁸ Under this test, the terms of the contract would be determinative as to the existence or non-existence of a contract of lease, unless shown to misrepresent the reality of the situation. For example, an agreement for the occupation and use of a parking garage where the terms of the agreement allow the landlord to enter the premises at will as well as keep their own car there would appear to be a licence agreement, as the landlord can use the property to a significant degree. However, should it be found that as a matter of fact the landlord does not exercise these rights at all, the contract would be a lease, the contractual provisions having been drafted to avoid the consequences thereof.¹³⁹

A potential problem with this approach is that, in most cases, it will be the tenant who is attempting to establish a lease rather than the landlord, as they will have the most to gain from such a contract. As such, it seems strange that the focus should be on the landlord's rights, especially since they already have a real right of ownership in the property.

A more practical approach would be to require the tenant to prove that they have a right to exclude the landlord from exercising the right let to them. This would appear to be the essence of modern exclusive possession, as Blair defines it,¹⁴⁰ and works alongside his argument that two people sharing the same right do not have exclusivity. If a tenant can positively establish the four essential requirements of a lease, and that they are themselves capable of preventing other people, including their landlord, from exercising *the same right*, there is very little room to conclude that their contract is something other than a lease. This approach also sidesteps any potential conflict between the Housing Act tenancies, which are the sectors that would most benefit from a possession requirement, and other forms of lease. Under this system, a residential tenant would be required to prove that they can stop other people from living in their leased property, which is analogous to exclusive occupation while at the same time not imposing that requirement on leases that would see no practical benefit therefrom. A tenant under a fishing lease, for example, would have to prove that they had the right to prevent others from fishing on a given piece of land under the same test.

As is usually the case, hypothetical scenarios could be foreseen where this test would face difficulties. An example would be a contract farming case analogous to *Caird Properties* but with provisions for a rent-like payment. However, in practice, payments in those sorts of contracts tend to be lump sums, and the legal documentation well-drafted enough to exclude a relationship of landlord and tenant.¹⁴¹ The former, alone, excludes the contract from being a lease on the basis of *Mann v Houston*, so in practice there is likely to be little difficulty in adapting these arrangements to a right to exclude test. Similarly, there are no difficulties applying

¹³⁸ Davies (n 56) 322.

¹³⁹ See, for a residential example and approach in the context of English Law, *Antoniades v Villiers* [1990] 1 AC 417, 462 (Lord Templeman).

¹⁴⁰ Blair (n 111) 32.

¹⁴¹ Paisley and Cusine (n 116) 3.

this test to long leases without possession, as the existence of the lease will still entitle the tenant to prevent others from occupying the land.

6. Conclusion

As the rights available to Scottish tenants have increased, so too has the significance of exclusivity within the law of leases. At the foundational level of a personal contract, it is, if Lord Gill's view is to be accepted, irrelevant. The Leases Act 1449 brought possession into the equation to secure a real right, and it has long been accepted that this means exclusive possession.¹⁴² Finally, a third, even greater level of protection came in the twentieth century, under which a strong current of exclusivity also appeared to run. Given the diversity of the sector, a conflict was inevitable. Scots law recognises a number of different contracts as leases, from offering places to live to sporting rights to long-term rights, each of which has their own set of needs that must be met to allow the law as a whole to function effectively. Additionally, the twentieth century has prompted the judiciary to make decisions on how to distinguish licences from leases, but these decisions have not always been consistent, creating an unnecessary amount of legal ambiguity.

Calls for the introduction of exclusive possession in modern cases have mostly come from a residential perspective, where security of tenure, a right stronger than a real right, has become available. Comparatively, in an agricultural context, the courts have not viewed exclusive possession as a vital component of a *contract* of lease. The best explanation that can be found for this is that, in an agricultural context, the definition of a lease is not far removed from how it was originally conceived hundreds of years ago. Meanwhile, the Housing Acts have redefined, for their purposes, what a tenancy is in order to cast the net wide and empower as many tenants as they can. In the process those Acts have sacrificed the absolute need for rent, which is, more than any other essential requirement, sacrosanct in Scots law.

With this in mind, it is not difficult to see why, when examining whether or not secure or assured tenancies exist, the courts might look favourably upon exclusive possession. However, the fact that exclusive possession greatly assists in identifying one type of lease – those under the Housing Act tenancies – is not a sufficient reason to impose it as a requirement upon all leases. What is instead required, if Scotland is to create a proper distinction between leases and licences, is a test that captures the spirit exclusive possession is meant to indicate without forgetting the broad nature of the Scottish lease.

By using a wider interpretation and taking exclusive *use* rather than exclusive possession as the basis for a 'right to exclude' test, Scots law will be able to do this. Such a test would examine each contract based on the power that it affords to the occupier as regards the particular right in question. If it is strong enough to allow the occupier to exclude others from exercising that right, within reason, it will be a lease. This would be determined based on the provisions in the contract. However, should

¹⁴² *Millar* (n 21) 8.

The Fifth Element: Should Exclusive Possession Be Considered an Essential Requirement for the Constitution of a Contract of Lease in Scots Law?

the reality of the situation differ from what is provided in the contract, as an attempt at manipulating the true agreement of the parties, then the actions of the two parties will take precedence. In this way, modern challenges to lease law are overcome without sacrificing the uniquely Scottish identity and character of this type of contract.

Agricultural Tenancies: Making Room for Generation Y – Will Only the Land Endure?

SAM READ-NORRIE

Abstract

*The level of agricultural tenancies in Scotland has been in steady decline for over a century, with the result that new tenant farmers are struggling to gain access to the agricultural sector. The present study considers whether recent legislation has provided landlords with the confidence to bring forward more land to the letting market, and whether recent legislative reform has learnt from previous mistakes. In particular, this paper assesses the impact that measures and proposals considered under the Agricultural Holdings (Scotland) Act 2003 – such as the pre-emptive right to buy, the absolute right to buy and security of tenure – have had on the confidence of landlords to lease land, and how this is reflected in the circumstances surrounding the case of *Salvesen v Riddell*. Furthermore, this paper considers whether proposals and measures brought forward under the Land Reform (Scotland) Act 2016 have learnt from the *Salvesen* case, and whether the balance of rights between landlords and tenants needed for a healthy tenanted sector has been met.*

Keywords: Agricultural Holdings, Scotland, Land Reform

1. Introduction

In 2007, the Tenant Farming Forum (TFF) was commissioned by the Cabinet Secretary, Richard Lochhead, to investigate barriers which may be preventing a new generation of farmers from entering the agricultural industry.¹ It had become apparent to the Scottish Government that the average age of those currently working in the agricultural industry was not being lowered sufficiently by the entry of new blood into the industry, with 51 per cent of the current workforce aged over fifty-five, a quarter of whom were older than sixty-five.² In a recent study of tenant farmers commissioned by the Scottish Government in anticipation of the review of agricultural holdings legislation, it was found that only 7 per cent of tenant farmer respondents were under the age of forty, which is in stark contrast to the 74 per cent who were over the age of fifty.³

The Agricultural Holdings Legislation Review Group (AHLRG) raised concerns, therefore, that current legislation may be preventing farmers from retiring earlier and, thereby, preventing a younger generation of farmers from, ‘taking the

¹ Tenant Farming Forum, ‘Tenant Farming Forum Consultation Helping New Entrants into Agriculture’ (2007)

<www.tenantfarmingforum.org.uk/eblock/services/resources.ashx/000/241/600/Helping_New_Entrants_into_Agriculture.pdf> accessed 21 March 2016.

² *ibid.*

³ C Martin and others, *Survey of Agricultural Tenant Farmers* (Scottish Government Social Research 2014) para 2.4.

reins', before the productivity of a holding is compromised by the incapacity of an aging tenant to farm efficiently.⁴ Indeed, the survey commissioned by the Scottish Government revealed that almost a fifth of respondents stated that they were experiencing 'long term illness, disability or health problem[s] (...) which limits the daily activities or work that they can do.'⁵ The Scottish Government also raised concerns that there is a risk that the skills, which many of these aging farmers have honed over the years, may be at risk of dying out, especially the skills of those involved in livestock.⁶ Despite this, the Scottish Government retained some hope that the research conducted by the TFF would reveal a strong contingency of young farmers available to succeed to farms, where previous research focussed solely on the age of the principal farmer, concealing the existence of a successor.⁷ Instead, however, the TFF's report confirmed that there is major concern in the industry regarding its aging population, revealing that 83 per cent of its respondents agreed that the sector is faced with a 'real new entrant problem'.⁸

The TFF published a report in 2008 after extensive consultation, which sought to establish the most important barriers preventing new entrants into the farming sector. The report revealed that one of the most important barriers facing the industry was a lack of available land to let.⁹ Given the high capital cost of farming equipment and land, leasing provides an essential access route for new entrants. It was recognised by the Cabinet Secretary, Richard Lochhead, that, 'Scotland now has one of the lowest proportions of rented land anywhere in Europe', and that action was essential to stop further reductions thereof; as Lochhead explained, '[n]o industry can expect to flourish when there are barriers to new entrants and obstacles faced by successful businesses that want to improve, diversify and flourish'.¹⁰

The reduction of agricultural holdings available to let is not a new issue. Since the start of the twentieth century, the dualistic pattern of agricultural land tenure has been converted from one dominated by tenanted holdings, to one where there is now a greater number of owner-occupants than tenant farmers. The Land Reform Review Group (LRRG) revealed in their Final Report that the tenanted sector in 1913 accounted for 90 per cent of farms in Scotland; however, since that time, the sector has been in continual decline, with tenancies reducing from 71,740 in 1913 to just

⁴ Scottish Government, *Review of Agricultural Holdings Legislation* (Interim Report, Scottish Government 2014) para 224.

⁵ C Martin and others, *Survey of Agricultural Tenant Farmers* (Scottish Government Social Research 2014) para 2.4.

⁶ Tenant Farming Forum, 'Tenant Farming Forum Consultation Helping New Entrants into Agriculture' (2007) <www.tenantfarmingforum.org.uk/eblock/services/resources.ashx/000/241/600/Helping_New_Entrants_into_Agriculture.pdf> accessed 21 March 2016.

⁷ *ibid.*

⁸ P Cook and others, 'Barriers to New Entrants to Scottish Farming: An Industry Consultation for the Tenant Farming Forum' (2008) para 1.3 <www.tenantfarmingforum.org.uk/eblock/services/resources.ashx/000/244/597/58_final_report_forum_contractors.pdf> accessed 22 February 2017.

⁹ *ibid.*

¹⁰ Scottish Government, *Review of Agricultural Holdings Legislation* (Interim Report, Scottish Government 2014) para 224.

14,274 by 1980.¹¹ The latest statistics published also reveal that the decline of agricultural tenancies has shown no sign of abating in the modern era: the current level sits at almost half that of the 1980s, reducing from 41 per cent to 23 per cent between 1983 and 2014.¹²

The LRRG explained that there were various historical factors contributing to the reduction of the tenanted sector, such as the amalgamation of smaller tenancies,¹³ the rise of owner-occupation,¹⁴ the mechanisation of labour in the 1950s,¹⁵ and increasing subsidy payments.¹⁶ Importantly, however, the reduction of the tenanted sector also comes as a result of legislation. After the Second World War, food security was a major concern of public policy, influencing the then Labour Government to enact radical measures, greatly interfering with the terms of agricultural leases.¹⁷ Up to that point, landlords would have expected to retrieve vacant possession of their land at the end of the lease; however, the Agricultural Holdings (Scotland) Act 1948 provided tenant farmers with security of tenure, interfering with the operation of notices to quit.¹⁸ The consequence of this was effectively to lock up land by unlimited succession to tenancies, where generation after generation of farmers could retain their lease within their families.¹⁹ After the 1948 Act's consolidation in the Agricultural Holdings (Scotland) Act 1949, these tenancies were preserved in the modern era of agricultural tenancy law by the Agricultural Holdings (Scotland) Act 1991.²⁰ This was a major deterrent to land leasing as the capital value of land subject to a secure 1991 Act tenancy was reduced significantly. Unsurprisingly, landlords effectively ceased creating secure tenancies by the 1980s.²¹

In light of the continuing reduction of tenancies, the Land Reform Policy Group was established by the newly formed Scottish Executive in 1997 to consider the most effective legislative means of remedying this situation.²² The group

¹¹ Land Reform Review Group, *The Land of Scotland and the Common Good* (Scottish Government 2014) 195.

¹² Scottish Government, *Tenanted Agricultural Land in Scotland 2014* (Statistical Publication Agriculture Series 2015) 4.

¹³ Land Reform Review Group, *The Land of Scotland and the Common Good* (Scottish Government 2014) 195.

¹⁴ P Cook and others, 'Barriers to New Entrants to Scottish Farming: An Industry Consultation for the Tenant Farming Forum' (2008) para 1.7

<www.tenantfarmingforum.org.uk/eblock/services/resources.ashx/000/244/597/58_final_report_from_contractors.pdf> accessed 22 February 2017.

¹⁵ Land Reform Review Group, *The Land of Scotland and the Common Good* (Scottish Government 2014) 196.

¹⁶ *ibid.*

¹⁷ H Lean, 'A Short History' (Farm North East 2013) <www.stronachs.com/wp-content/uploads/2013/12/A-short-history.pdf> accessed 21 March 2016.

¹⁸ *ibid.*

¹⁹ Land Reform Review Group, *The Land of Scotland and the Common Good* (Scottish Government 2014) s 28.1, para 9.

²⁰ S Notley, *Scottish Agricultural Law Handbook* (Avizandum 2009) 99-100.

²¹ Land Reform Review Group, *The Land of Scotland and the Common Good* (Scottish Government 2014) s 28.1, para 9.

²² Land Reform Policy Group, *Identifying the Problems* (Scottish Office 1998) para 1.2.

proposed several recommendations with that intention, eventually incorporated into a White Paper, forming the basis for the Agricultural Holdings (Scotland) Act 2003.

The present study examines the effectiveness of this act in achieving one of its primary aims of encouraging landowners to bring forward more land to the letting market, which, as identified above, is an essential route of access for new entrants into the farming industry. Additionally, this research considers critically whether the Land Reform (Scotland) Act 2016 has learnt from the legislative process surrounding the 2003 Act, which saw some difficulties in terms of balancing the rights of tenants and landlords, as was reflected in the case of *Salvesen v Riddell*.²³ In particular, it is considered whether the newest addition to the agricultural legislative regime will go any further towards achieving an increase of land on the letting market.

2. *Agricultural Holdings (Scotland) Act 2003*

One of the major concerns highlighted by a recent review of agricultural holdings legislation was the lack of confidence which the sector has felt as a result of political uncertainty.²⁴ The following section assesses the impact which recent legislation has had on the confidence of landlords to contribute further land to the letting market in this sector. It has been established by research that the current shortage of tenancies is the greatest barrier preventing new entrants from gaining a foothold in the industry,²⁵ which is linked to legislative uncertainty and its effect on the confidence of landlords to let land.²⁶

A. Pre-Emptive Right to Buy

With the establishment of the new Scottish Parliament, the opportunity arose for land tenure to be reformed in order, 'to determine what needs to change to fit Scotland's land legislation for the twenty-first century'.²⁷ Agricultural land accounted for 80 per cent of rural space in Scotland; therefore, it was inevitable that the tenanted farming sector would also be reviewed within this process. In October 1997 The Land Reform Policy Review Group (LRPG), chaired by Lord Sewel, was created and tasked with the remit of, 'identify[ing] and assess[ing] proposals for land reform, taking account of their cost, [and their] legislative and administrative implications'.²⁸ One such proposal identified by the group was the argument that

²³ *Salvesen v Riddell* [2013] UKSC 22, 2013 SC 236.

²⁴ Scottish Government, *Review of Agricultural Holdings Legislation* (Interim Report, Scottish Government 2014) para 87.

²⁵ P Cook and others, 'Barriers to New Entrants to Scottish Farming: An Industry Consultation for the Tenant Farming Forum' (2008) para 1.3
<www.tenantfarmingforum.org.uk/eblock/services/resources.ashx/000/244/597/58_final_report_from_contractors.pdf> accessed 22 February 2017.

²⁶ C Martin and others, *Renting-Out Agricultural Land in Scotland Survey* (Scottish Government Social Research 2014) 52.

²⁷ Land Reform Policy Group, *Identifying the Problems* (Scottish Office 1998) para 1.2.

²⁸ *ibid.*

tenants of secure 1991 Act tenancies should be afforded the right to buy their holdings.²⁹

An ‘absolute’ right to buy had already been established by the Crofting Reform (Scotland) Act 1976, allowing crofters to acquire their holdings on application to the Land Court where the landlord had refused to sell their property voluntarily.³⁰ This absolute right to acquire an ownership interest can be contrasted with that of the pre-emptive right to buy, which can only be triggered by the owner's volition to sell their property.

The LRPG established, however, that an absolute right to buy (ARTB) posed a major risk to the sector: this would deter landlords from letting land, as well as leaving the Scottish Executive exposed to compensation claims potentially exceeding £100 million; the proposal was regarded as a ‘non-starter’.³¹ After recommendations from the group, Ross Finnie, the then Minister for Environment and Rural Development, rejected the ARTB in his Draft Agricultural Holdings Bill, which was brought before the Scottish Parliament on 16 September 2002.³² Furthermore, the LRPG chose to omit the pre-emptive right to buy in the White Paper that it produced for the Scottish Government.³³ Despite this, Ross Finnie made the following declaration two years later in the Draft Agricultural Holdings (Scotland) Bill:

For some time I have been considering the case for a pre-emptive right to buy for tenant farmers. I have concluded that a pre-emptive right to buy can make a valuable contribution to our land reform agenda [and] I have therefore decided that the Bill should provide for [this].³⁴

When one considers the Scottish Government's rationale for this change of direction, it would seem justified to have afforded secure 1991 Act tenants a pre-emptive right to buy. The nature of those tenants' tenure, with continuous rights of succession, would have already diminished the value of any land encumbered thereby. In these circumstances, it would have seemed prudent that a landlord should negotiate a sale price with the tenant, who may have been willing to pay for a value somewhere between that of vacant possession, and that of a sale on the open market with a sitting-tenant; in contrast, sale to a third party would have only amounted to 50 per cent of the price at vacant possession.³⁵ One could even question why such a measure had to be incorporated into statute, given that it was already the practice of many landlords to offer the property to the tenant.³⁶ In any case, the measure was commendable for emboldening tenants to invest in their holdings. However, the manner in which this pre-emptive right to buy came into being undermined the

²⁹ Land Reform Policy Group, *Identifying the Solutions* (Scottish Office 1998) para 8.1.

³⁰ The Crofters (Scotland) Act 1993, s 12.

³¹ Land Reform Policy Group, *Identifying the Problems* (Scottish Office 1998) 65.

³² Scottish Executive, *Draft Agricultural Holdings (Scotland) Bill – Consultation Paper* (2002) para 4.4 <www.gov.scot/Resource/Doc/46905/0029913.pdf> accessed 23 February 2017.

³³ Scottish Executive, *Agricultural Holdings (Scotland) Bill – Policy Memorandum* (The Stationery Office 2002) para 20.

³⁴ *ibid* 2.

³⁵ Scottish Executive, *Agricultural Holdings (Scotland) Bill – Policy Memorandum* (The Stationery Office 2002) para 18.

³⁶ S Notley, *Scottish Agricultural Law Handbook* (Avizandum 2009) 99-100.

primary aim of the Agricultural Holdings (Scotland) Act 2003 – namely to ‘stimulate the rented sector’.³⁷

It was highlighted recently by the AHLRG that legislative intervention and political uncertainty have played a major role in deterring landlords from bringing more land forward to the tenant farming sector.³⁸ Much of landlords’ mistrust of the Scottish Parliament emanated from the legislation passed in 2003.³⁹ Indeed, the manoeuvre of the Scottish Government to include the pre-emptive right to buy in its Draft Bill was the beginning of a sequence of events, which would eventually undermine the aforementioned policy aim.⁴⁰

The LRPG were tasked with assessing proposals for legislation and to consider their ‘legislative and administrative implications’, a remit which led the group to consider the pre-emptive right to buy.⁴¹ After the group’s consultation, it was agreed between the Scottish Ministers and the LRPG that this right to buy would not be pursued any further, and this measure was, therefore, not included in the Agricultural Holdings White Paper.⁴² However, within two years, the Scottish Executive had made a complete ‘U-turn’ on its policy,⁴³ rejecting the advice of both the LRPG and key stakeholders.⁴⁴ The reason behind this sudden change was the growing pressure from certain campaigns, such as that of Liberal Democrat MSP, George Lyon, calling on the Scottish Executive to reconsider its stance on the right to buy.⁴⁵ This pressure eventually led the Executive to reconsider its investigation into the pre-emptive right to buy,⁴⁶ and, by January 2002, Ross Finnie announced to the Scottish Parliament that this right was to be included in its Agricultural Holdings Bill.⁴⁷

Despite receiving support from the key stakeholder, NFU Scotland, the Executive’s contradictory move was immediately criticised by the Scottish Landowners’ Federation (SLF).⁴⁸ The then convener of SLF, Robert Balfour,

³⁷ Scottish Executive, *Draft Agricultural Holdings (Scotland) Bill – Consultation Paper* (2002) 1 <www.gov.scot/Resource/Doc/46905/0029913.pdf> accessed 23 February 2017.

³⁸ Scottish Government, *Review of Agricultural Holdings Legislation* (Interim Report, Scottish Government 2014) para 87.

³⁹ *ibid* para 90.

⁴⁰ Scottish Executive, *Draft Agricultural Holdings (Scotland) Bill – Consultation Paper* (2002) 2 <www.gov.scot/Resource/Doc/46905/0029913.pdf> accessed 23 February 2017.

⁴¹ Land Reform Policy Group, *Identifying the Solutions* (Scottish Office 1998) 65.

⁴² Scottish Executive, *Agricultural Holdings (Scotland) Bill – Policy Memorandum* (The Stationary Office 2002) para 20.

⁴³ BBC News, ‘Minister’s U-turn over Land Purchase’ (*BBC News*, 16 April 2002) <<http://news.bbc.co.uk/1/hi/scotland/1932255.stm>> accessed 21 March 2016.

⁴⁴ ‘Right-to-Buy for Tenants by Next April, says Finnie’ *The Scotsman* (Edinburgh, 31 January 2002) <www.scotsman.com/business/companies/right-to-buy-for-tenants-by-next-april-says-finnie-1-595549> accessed 21 March 2016.

⁴⁵ *ibid*.

⁴⁶ Scottish Executive, *Agricultural Holdings (Scotland) Bill – Policy Memorandum* (The Stationary Office 2002) para 20.

⁴⁷ Scottish Parliament Justice Committee 2, 30 January 2002, Scottish Parliament Official Report, col 985.

⁴⁸ ‘Landowners Slam New Right to Buy Proposals’ *The Scotsman* (Edinburgh 2002) <www.scotsman.com/business/companies/landowners-slam-new-right-to-buy-proposals-1-503401> accessed 21 March 2016.

lambasted the measure for its intervention in private tenancy agreements, and for its retrospective effect on arrangements not originally intended to afford a right to buy.⁴⁹ The group argued that this last minute change to the Draft Bill would only discourage landowners from letting as a result of the ‘uncertainty’ that it had caused.⁵⁰ The Conservative MSP, Alex Fergusson, also denounced the measure, arguing that it would instantly depreciate the value of estate lands, where enterprises would be broken up by the offending rights.⁵¹ It would appear, therefore, that even with the benefits of this measure for tenants, which could only be triggered by the landlord’s volition to sell and terminate his interest in a property, the mere fact that the Scottish Parliament could interfere with agricultural leases to favour tenants caused anxiety to landlords. The balance of rights which was vital to achieve stability in the rented sector already appeared under threat at this early stage.

B. The Case for an Absolute Right to Buy (ARTB)

Although 82 per cent of members of the key stakeholder group NFU Scotland had shown their support for the pre-emptive right to buy, 66 per cent voted against an ARTB.⁵² However, another tenant farmer group questioned the Scottish Executive’s rejection of the latter. Malcolm McCall, the then chairman of the Scottish Tenant Farmer Action Group (STFAG), wrote to Ross Finnie immediately following his announcement of the pre-emptive right to buy’s inclusion in the Draft Bill, criticising the Minister’s decision for being undemocratic, having prevented any discussion of the ARTB being brought before the Scottish Parliament.⁵³ Andrew Thin compiled a submission for STFAG in June 2002, stating that there was a public interest argument for the transformation of the pre-emptive right into an absolute one.⁵⁴ It was argued that the ARTB could address economic, social and environmental policy concerns by removing constraints to sustainable rural development,⁵⁵ and that it was in the public interest to encourage investment in rural communities, as well as allowing tenants to express themselves within those areas where they were deeply-rooted.⁵⁶ Furthermore, the ARTB would give secure tenants the confidence to take up environmentally friendly measures, which the landlord could otherwise veto.⁵⁷

⁴⁹ T Peterkin, ‘Tenant Farmers May Be Given the Right to Buy’ *The Telegraph* (London, 1 February 2002) <www.telegraph.co.uk/news/uknews/1383405/Tenant-farmers-may-be-given-the-right-to-buy.html> accessed 21 March 2016.

⁵⁰ *ibid.*

⁵¹ *ibid.*

⁵² S Wright, ‘Give Tenants Right to Buy Land on Market’ (Farmers’ Weekly 2002) <www.fwi.co.uk/news/give-tenants-the-right-to-buy-land-on-market.htm> accessed 21 March 2016.

⁵³ ‘Tenant Farmers Press Finnie on Absolute Right-to-Buy’ *The Scotsman* (Edinburgh, 29 May 2002) <www.scotsman.com/business/companies/tenant-farmers-press-finnie-on-absolute-right-to-buy-1-568963> accessed 21 March 2016.

⁵⁴ Submission from Scottish Tenant Farmers Action Group to the Scottish Parliament (June 2002) <www.parliament.scot/S4_RuralAffairsClimateChangeandEnvironmentCommittee/General%20Documents/ECHR_-_STFA.pdf> accessed 23 February 2017.

⁵⁵ *ibid* para 1.2.1.

⁵⁶ *ibid* para 1.2.6.

⁵⁷ *ibid* para 1.2.3.

Despite these exhortations, the Executive made clear in its policy memorandum that its stance on the ARTB would remain the same, explaining that this measure would ‘decimat[e]’ the agricultural rented sector.⁵⁸ Furthermore, this would make the Executive liable for compensation exceeding £100 million, and impinge upon Article 1 of the First Protocol to the European Convention on Human Rights, which enshrines the right to peaceful enjoyment of one’s property.⁵⁹ Despite this, STFAG claimed that, although the Executive’s stance had remained the same, ‘the public interest case for an absolute right to buy had been largely accepted by policy makers in Edinburgh’.⁶⁰ STFAG perceived the exclusion of the ARTB as a reflection of the ‘paralysis’ suffered by the Executive and Parliament, occasioned by a ‘sense of institutional caution (...) not (...) necessarily in the public interest’.⁶¹ STFAG and land-reform campaigners continually sought to dispel the ‘myths’ surrounding the ARTB, such as the threat of a decrease in tenancies.⁶² This eventually swayed many of the senior Ministers overseeing the legislative scrutiny of the Bill, who stated in their Stage 1 Committee Report that, ‘the majority of the Committee is (...) sympathetic to an absolute right to buy for secure tenants under the 1991 Act’.⁶³

Once again, it appeared that there was to be a complete change of direction from the policy measures agreed in the White Paper. The Rural Development Committee had chosen to debate the merits of ARTB, despite the evidence provided by representative groups, arguing that landowners, ‘must be given the confidence to allow them to [let land]. Concern that politicians might, at some stage, extend preemptive rights to ARTB wholly militates against such confidence.’⁶⁴ It became clear to the Scottish Parliament that there was some veracity to this statement as anecdotal evidence began to surface that landlords were responding to a potential ARTB by terminating tenancies across the country.⁶⁵

C. Security of Tenure and Limited Partnerships

Importantly, however, there was a second element to this reaction. The anecdotal evidence provided concerned a specific type of tenancy arrangement devised by

⁵⁸ Scottish Executive, *Draft Agricultural Holdings (Scotland) Bill – Consultation Paper* (2002) para 4.4 <www.gov.scot/Resource/Doc/46905/0029913.pdf> accessed 23 February 2017.

⁵⁹ Scottish Executive, *Agricultural Holdings (Scotland) Bill – Policy Memorandum* (The Stationery Office 2002) para 45.

⁶⁰ Submission by the Scottish Tenant Farmers Action Group to Rural Development Committee (October 2002) para 1.2 <<http://archive.scottish.parliament.uk/business/committees/historic/x-rural/reports-02/rar02-13-03.htm>> accessed 23 February 2017.

⁶¹ *ibid* para 1.3.

⁶² ‘Submission from Andy Wightman to Rural Development Committee’ (5 November 2002) <<http://archive.scottish.parliament.uk/business/committees/historic/x-rural/reports-02/rar02-13-04.htm>> accessed 23 February 2017.

⁶³ Rural Development Committee ‘Stage 1 Report on the Agricultural Holdings (Scotland) Bill’ (Report No 13, 2002) para 77.

⁶⁴ Submission from Scottish Landowners’ Federation to the Rural Development Committee (October 2002) <<http://archive.scottish.parliament.uk/business/committees/historic/x-rural/reports-02/rar02-13-03.htm>> accessed 23 February 2017.

⁶⁵ Rural Development Committee 4 February 2003, Scottish Parliament Official Report, col 4161.

landlords to avoid the effects of legislation consolidated under the Agricultural Holdings (Scotland) Act 1991. Predominantly, this statute sought to strengthen the tenants' position, providing them with confidence to invest in their holdings, as well as to share some of the capital burden required by a farming business.⁶⁶ This was achieved by measures such as security of tenure, compensation at waygo, and obligations on the landlord to provide fixed equipment.⁶⁷

Landlords attempted various measures to avoid these cumbersome legislative controls by, for example, using sub-tenancies. One practice was for a landlord to let land to their spouse, wherefrom a sub-let would be granted to a de facto tenant, in an attempt to obviate the latter's challenge to a notice to quit.⁶⁸ The courts have tended to reject these measures on the basis that they undermined public policy.⁶⁹ By 1997, however, the Court of Session finally upheld an arrangement used by landowners to avoid security of tenure.⁷⁰ Landowners found that they could use the Limited Partnership Act 1907 by granting a tenancy to a limited partnership body, where the landlord and tenant were limited and general partner respectively. This circumvented the restrictive process of notices to quit under the 1991 Act by allowing the landlord to bring the tenancy to an end through the dissolution of the partnership body.

Although this arrangement received legal validation from the Court of Session in *MacFarlane v Falfield Investments Ltd*,⁷¹ previous legislative attempts from within the UK Parliament were made to outlaw the use of these arrangements in 1983.⁷² MP Martin O'Neill argued that there was an urgent need to amend the Agricultural Holdings (Amendment) (Scotland) Bill to prevent large estates from offering limited partnerships to young farmers, who were effectively forced into these arrangements as a result of the shortage of tenancies.⁷³ The proposed abolition of these arrangements was, however, unsuccessful. The UK Conservative Government of the period observed that the 1979 Northfield Report into agriculture had shown that not all of these arrangements prejudiced the rights of tenants, and that very few tenants had made any complaints regarding the operation of these tenancies.⁷⁴ The Government also feared that the proscription of the arrangement would lead to further reductions in tenancy levels by removing the only popular letting vehicle of that time; in the absence of limited partnerships, landlords would likely take farming operations in-hand.⁷⁵

The Scottish Executive made clear, however, that history would not repeat itself: the limited partnerships were to have no role in the future of Scotland's agriculture.⁷⁶ With the position of tenants being further strengthened by the

⁶⁶ S Notley, *Scottish Agricultural Law Handbook* (Avizandum, 2009) 2.

⁶⁷ *ibid.*

⁶⁸ *Gisborne v Burton* [1988] 3 All ER 760; Agricultural Holdings (Scotland) Act 1991, s 23(3) and (8).

⁶⁹ S Notley, *Scottish Agricultural Law Handbook* (Avizandum 2009) 2.

⁷⁰ *MacFarlane v Falfield Investments Ltd* (1998) SC 14.

⁷¹ *ibid.*

⁷² HC Deb 21 April 1983, vol 41, col 454.

⁷³ *ibid.*

⁷⁴ *ibid* col 459.

⁷⁵ *ibid.*

⁷⁶ Scottish Executive, *Agricultural Holdings: Proposals for Legislation* (HMSO 2000) para 2.9.

Agricultural Holdings Bill, providing them with a minimum term in the Limited Duration Tenancy (LDT), as well as rights to diversify and to claim compensation at waygo, the Scottish Executive wanted to ensure that landlords would not be able to evade these provisions.⁷⁷

The stakeholder groups argued, however, that the limited partnership tenancy arrangement should be left to continue, as this provided the flexibility for landowners and tenants to agree on the length best suited to their needs.⁷⁸ This was already the case in England and Wales, where the Agricultural Holdings Act 1995 allowed freedom to negotiate.⁷⁹ The Scottish Landowners' Federation argued, furthermore, that in practice the dissolution of partnership could only operate in accordance with the deeds recorded in the partnership agreement.⁸⁰ NFU Scotland believed that the continuation of the limited partnership agreement was vital as a commercial option and, therefore, any change to these arrangements would result in a reluctance to let land.⁸¹

Despite these claims, the Scottish Executive would not allow the arrangements to continue as an evasive measure to the new Bill.⁸² It was established in the Draft Bill, therefore, that protection was to be afforded to de facto tenants of limited partnerships. General partners were to be given the right to continue the tenancy in place of the limited partnership, where they had been served a notice to quit, and the intention of that notice was to deprive the tenant of his rights.⁸³ Importantly, this proposal was only applied to tenancies created after the enactment of the legislation. The Scottish Executive made it clear that it had no intention of extending these measures to have retrospective effect on existing limited partnerships agreements.⁸⁴

D. The Consequences of Legislative Instability

Despite this statement, and the numerous warnings from stakeholders regarding the interference with limited partnerships, the Scottish Parliament eventually did extend these measures to existing limited partnership agreements. Furthermore, Ross Finnie made suggestions that a pre-emptive right to buy might be considered to allow the general partners of existing partnerships to acquire ownership of their holdings, despite the fact that there was no clear support for such a right during the

⁷⁷ Scottish Executive, *Draft Agricultural Holdings (Scotland) Bill – Consultation Paper* (2002) para 5.21 <www.gov.scot/Resource/Doc/46905/0029913.pdf> accessed 23 February 2017.

⁷⁸ *ibid* para 3.17.

⁷⁹ Land Reform Policy Group, *Identifying the Solutions* (Scottish Office 1998) 58.

⁸⁰ Scottish Executive, *Draft Agricultural Holdings (Scotland) Bill – Consultation Paper* (2002) para 3.17 <www.gov.scot/Resource/Doc/46905/0029913.pdf> accessed 23 February 2017.

⁸¹ *ibid* para 3.16.

⁸² *ibid* 11.

⁸³ Scottish Executive, *Agricultural Holdings (Scotland) Bill – Policy Memorandum* (The Stationery Office 2002) para 17.

⁸⁴ Scottish Executive, *Agricultural Holdings: Proposals for Legislation* (HMSO 2000) para 2.9.

consultation process.⁸⁵ This, coupled with the fear of an imminent ARTB, led to the service of dissolution notices on general partners across the country.⁸⁶

The Scottish Parliament reacted to this situation with amendment 169: this proposed to allow de facto tenants, who were served a notice of dissolution on or after 4 February 2003 to a ‘relevant date’ in the future, to resume the tenancy in place of the partnership.⁸⁷ This completely contradicted the Scottish Executive’s previous statement, which asserted that existing limited partnership arrangements would remain unaffected by the 2003 Act’s provisions. Instead, amendment 169 would have retrospective effect on these tenants, with the severe consequence that they would become 1991 Act tenants with the benefit of security of tenure. One commentator highlights the draconian effect of this, explaining that the limited partnership arrangements were motivated solely by an intention to provide a fixed term tenancy, concluded in private by the parties.⁸⁸ Instead, these tenancies would now have full security of tenure, as well as an impending right to buy, thereby significantly reducing the capital value of the property.

The ramifications of this were equally severe: the intensity of these dissolution notices markedly increased to the point that in one evening on 3 February 2003, the night before this amendment was proposed to take effect, more than two hundred landowners informed their general partners of the end of their tenancy.⁸⁹ After receiving news of this, the Scottish Parliament responded with a further amendment: the date initially proposed for amendment 169 was brought forward to 16 September 2002, ensuring that those two hundred or so dissolved partnerships would also fall within the scope of this provision,⁹⁰ which became law under section 72 of the Agricultural Holdings (Scotland) Act 2003.

Importantly, however, this section created two separate regimes preventing landlords from taking evasive action. A statutory instrument established a division between the application of sections 72 and 73 by enacting a ‘relevant date’ as a cut-off point between the two provisions, falling on 30 June 2003.⁹¹ For those landlords who served notices after this date, section 73 would apply, which one commentator has described as the ‘successor’ regime.⁹² If the service fell before this date, section 72, appropriately named the ‘transition’ regime, would apply.⁹³ The importance of this distinction is discussed below.

⁸⁵ *Salvesen v Riddell* [2012] CSIH 26, 2013 SC 69 [22] (Gill LJ).

⁸⁶ Rural Development Committee 4 February 2003, Scottish Parliament Official Report, col 4161.

⁸⁷ SP Bill 62A Agricultural Holdings (Scotland) Bill [Marshalled List of Amendments Selected for Stage 2] Session 1 (2003).

⁸⁸ A Fox, ‘How the Leopard Changed Its Spots’ (2003) 48(5) JLSS 58.

⁸⁹ A Wightman, ‘Tenant Farming Crisis’ (Land Matters, 2012)

<www.andywightman.com/archives/794> accessed 21 March 2016.

⁹⁰ *Salvesen v Riddell* [2012] CSIH 26, (2013) SC 69 [27] (Gill LJ).

⁹¹ Agricultural Holdings (Relevant Date and Relevant Period) (Scotland) Order 2003, SSI 2003/294.

⁹² M Combe, ‘Human Rights, Limited Competence and Limited Partnerships: *Salvesen v Riddell*’ [2012] SLT 195.

⁹³ *ibid.*

3. *Salvesen v Riddell*

A. Lord Gill's Critique

In the case of *Salvesen v Riddell*,⁹⁴ Alastair Salvesen was one among many landowners who informed their general partners on 3 February 2003 that they had decided to dissolve their limited partnerships. The general partners in this case took advantage of the anti-avoidance mechanism afforded by sub-section 72(6), and continued the tenancy as secure 1991 Act tenants in their own right.⁹⁵ Salvesen appealed the continuation of the tenancy by his former general partners to the Land Court under sub-section 72(8) of the 2003 Act. For the appeal to succeed, however, the Land Court would have to find that the dissolution of the partnership was for a purpose, other than to deprive the de facto tenants of the rights bestowed upon them under the new legislation.⁹⁶ It was in this regard that the Land Court dismissed Mr Salvesen's application.⁹⁷ The case was appealed by the landowner to the Court of Session, but on this occasion Salvesen brought forward another argument: section 72 had violated his rights under the European Convention of Human Rights (ECHR).⁹⁸

Salvesen argued that section 72 was not a 'legitimate' provision, and that it arbitrarily punished him through the instant 'material reduction' of his property's capital value as a result of the indefinite tenancy to which his land was subject.⁹⁹ As a result of this 'material interference' with the landlord's peaceful enjoyment of his property, the appellant argued that Article 1 Protocol 1 had been engaged.¹⁰⁰ The appellant also argued that Article 14 of the ECHR had been violated: section 72 arbitrarily discriminated against landlords through differing and unequal treatment of them based purely upon whether they had served notice before or after 30 June 2003, placing a 'disproportionate and excessive burden' on those falling in the former category without any 'objective justification'.¹⁰¹ It was argued by the general partners, however, that the approach taken by the Scottish Parliament was justified as a remedy to the 'mischief' occurring externally to the Bill's passage, and that this was not as 'draconian' as applying section 72 to all limited partnership tenancies, regardless of the date on which notice was served.¹⁰² Furthermore, although section 72 was retrospective in nature, this effect was limited by targeting that 'flurry' of notices served before 4 February 2002.¹⁰³

The Lord Justice-Clerk began his judgment by criticising the drafting of section 72: if the provision were to be read literally, it would have no effect on those partnerships served with dissolution notices between 16 September 2002 and 22 April 2003. The

⁹⁴ [2013] UKSC 22, 2013 SC 236.

⁹⁵ Agricultural Holdings (Scotland) Act 2003, s 72(6).

⁹⁶ *ibid* s 72(9).

⁹⁷ *Salvesen v Riddell*, Scottish Land Court, SLC/3/09 <www.scottish-land-court.org.uk/decisions/SLC.03.09.rub.html> accessed 21 March 2016

⁹⁸ *Salvesen v Riddell* [2012] CSIH 26, 2013 SC 69 [53] (LJ-C Gill).

⁹⁹ *ibid* [55].

¹⁰⁰ *ibid*.

¹⁰¹ *ibid*.

¹⁰² *ibid* [56].

¹⁰³ *ibid*.

reason for this was that sub-section 72(9) required the court to consider whether the purpose of the dissolution was to deprive the tenant of any right afforded by the legislation.¹⁰⁴ However, the legislation was yet to be enacted and, therefore, none of these rights were in existence.¹⁰⁵ Furthermore, the main purpose of any termination of a contractual agreement is to bring the rights and obligations entailed therein to an end; therefore, the contents of sub-section 72(9) were devoid of meaning.¹⁰⁶ In any case, if sub-section 72(9) were to be applied as the Scottish Parliament had intended,¹⁰⁷ there would be a violation of Article 1 Protocol 1.¹⁰⁸ The Lord Justice-Clerk held that, on consideration of the statute as a whole, the 2003 Act pursued a social or economic policy in the public interest.¹⁰⁹ The court would need, however, to decide whether section 72 served that legitimate aim in an ‘appropriate and proportionate’ manner.¹¹⁰ This would entail consideration of the justifications for the enactment of section 72; the impact that the application of this provision would have upon the landlord’s interest in his property; and whether there existed any alternatives to this procedure, such as that provided by section 73 (discussed further below).¹¹¹ One of the justifications offered by the Scottish Government was that the legislation, as a whole, sought to meet a balance between the security of tenure afforded to tenants, and the freedom of landlords to use the land.¹¹² It was argued by the Scottish Government that section 72 was a furtherance of this aim as it prevented landlords from circumventing the aforementioned balancing of interests.¹¹³ The court found, however, that section 72 went beyond acting as a ‘furtherance’ of this aim: the Lord Justice-Clerk took the view that anti-avoidance mechanisms should seek only to remove the benefit achieved through circumvention; this measure went much further.¹¹⁴ It had created ‘excessive consequences’ for the landlord by encumbering him with a 1991 Act tenancy. It was the view of the court that less excessive means could have been used, such as preventing those dissolutions served within the defined period from being effectual,¹¹⁵ or providing the tenant with one of the new limited duration leases created by the statute.¹¹⁶ The Scottish Parliament could have also extended the measure found under section 73 to all cases, instead of creating two separate regimes.¹¹⁷

Indeed, the very existence of two regimes only highlighted the arbitrary nature of this provision: for those landlords who were fortunate enough to have served a dissolution notice after 30 June 2003, section 73 would afford them an

¹⁰⁴ Agricultural Holdings (Scotland) Act 2003, s 72(9)(a)(i).

¹⁰⁵ *Salvesen v Riddell* [2012] CSIH 26, 2013 SC 69 [61] (LJ-C Gill).

¹⁰⁶ *ibid* [64].

¹⁰⁷ Meeting of Scottish Parliament 12 March 2003, Scottish Parliament Official Report, col 16317.

¹⁰⁸ *Salvesen v Riddell* [2012] CSIH 26, 2013 SC 69 [73] (LJ-C Gill).

¹⁰⁹ *ibid* [75].

¹¹⁰ *ibid* [79].

¹¹¹ 16 September 2002 to 30 June 2003 as enacted by Agricultural Holdings (Relevant Date and Relevant Period) (Scotland) Order 2003, SSI 2003/294.

¹¹² *Salvesen v Riddell* [2012] CSIH 26, 2013 SC 69 [74] (LJ-C Gill).

¹¹³ *ibid* [74].

¹¹⁴ *ibid* [80].

¹¹⁵ *ibid* [82].

¹¹⁶ Agricultural Holdings (Scotland) Act 2003, ss 4 and 5.

¹¹⁷ *Salvesen v Riddell* [2012] CSIH 26, 2013 SC 69 [82] (LJ-C Gill).

incontestable notice to quit, whilst providing a notice period of two to three years.¹¹⁸ The benefit of this requirement of notice was that it acted as a ‘cooling off’ period to enable mediation to be undertaken, or a negotiated resolution to be reached.¹¹⁹ This was in stark contrast to the consequences produced by section 72, which placed an ‘excessive burden’ on landlords, encumbering their lands with secure tenancies, as well as imposing upon them a pre-emptive right to buy, to the significant detriment of the capital value of that land. No justification for this distinction had been offered by the Scottish Government,¹²⁰ which led the Lord Justice Clark to the conclusion that the two regimes amounted to ‘unreasonable discrimination’ between landlords.¹²¹ However, this did not amount to ‘discrimination’ in the sense of Article 14.¹²²

In an attempt to find any further justification for the enactment of section 72, the court turned its attention to comments made by the Deputy Minister for Environment and Rural Development, Allan Wilson, during the discussion of the amendments in the Scottish Parliament.¹²³ It was stated by the Minister during a parliamentary debate that this measure came in response to the ‘immoral’ actions of landlords,¹²⁴ ‘pending the final shape of the legislation’, and it was necessary, therefore, to protect, ‘general partners in a situation of great uncertainty (...) faced with the threat of imminent eviction’.¹²⁵ The Lord Justice-Clerk found that section 72 was, therefore, essentially a retaliatory measure against the ‘immoral’ actions which had been taken by the landlords and was, as a result, punitive in nature.¹²⁶

The court could not find any justification for the severity of this punishment, or for the generous level of protection afforded to general partners, which appeared to be based on a misconception of limited partnership agreements.¹²⁷ The Minister stated during the discussion of amendment 111 that this was a necessary measure to protect de facto tenants from the ‘great uncertainty’ caused by ‘imminent eviction’;¹²⁸ however, for vacant possession to be recovered in practice, the dissolution of limited partnerships had to be in accordance with the deeds recorded in the partnership agreement.¹²⁹ Furthermore, these legally valid contracts were entered into freely by the parties ‘at arm’s length’,¹³⁰ and it would have been known to both that the limited partner had the right to dissolve the relationship, so no greater ‘uncertainty’ could have been caused to the general partners.¹³¹

¹¹⁸ Agricultural Holdings (Scotland) Act 2003, s 73(4) and 73(5).

¹¹⁹ M Combe, ‘Remedial Works in Agricultural Holdings’ [2014] SLT 70.

¹²⁰ *Salvesen v Riddell* [2012] CSIH 26, 2013 SC 69 [97] (LJ-C Gill).

¹²¹ *ibid* [83].

¹²² *ibid*.

¹²³ *ibid* [87].

¹²⁴ Meeting of Scottish Parliament 12 March 2003, Scottish Parliament Official Report, col 16317.

¹²⁵ *ibid* col 16315.

¹²⁶ *Salvesen v Riddell* [2012] CSIH 26, 2013 SC 69 [90] (LJ-C Gill).

¹²⁷ *ibid* [91].

¹²⁸ Meeting of Scottish Parliament 12 March 2003, Scottish Parliament Official Report, col 16317.

¹²⁹ *Salvesen v Riddell* [2012] CSIH 26, 2013 SC 69 [91] (LJ-C Gill).

¹³⁰ *MacFarlane v Falfield Investments Ltd* 1998 SC 14.

¹³¹ *Salvesen v Riddell* [2012] CSIH 26, 2013 SC 69 [91] (LJ-C Gill).

The Lord Justice-Clerk also took the view that the action of these landowners could not have been ‘immoral’.¹³² Before the last amendment of section 72, this provision had not yet been extended to landlords issuing notices before 4 February 2003, who, up to that point, had received assurances in the Draft Bill,¹³³ as well as in the White Paper,¹³⁴ that there would be no interference with existing limited partnership agreements; therefore, they, ‘would [not] have (...) assumed that the Bill would be extended to existing leases’.¹³⁵ Perhaps the court was seeking to find a less cynical view of the landlords’ actions on 3 February 2003; however, Lord Gill’s statement here seems to disregard the increasing external pressure on the Rural Development Committee to extend the pre-emptive right to buy to an absolute one.

Indeed, this paper has shown that on numerous occasions, the Scottish Executive had already contradicted its own policy, becoming increasingly in favour of balancing rights to suit the needs of tenants. Furthermore, landowners were also likely to be aware of media coverage, broadcasting the unpredictable changes occurring within the legislative debate, and the increasing pressure for the introduction of ARTB.¹³⁶ When one also considers the sudden emergence of evidence of several hundred dissolution notices being issued in a single evening, occurring the night before amendment 169 was due to be discussed, this would lead one to believe that the landlords had, in fact, already, ‘assumed that the Bill would be extended to existing leases’.¹³⁷ Although, as Lord Gill states, they may have, ‘had no cause to fear that [their] action would be affected by any provision of the forthcoming legislation’,¹³⁸ they were likely to fear that *inaction* would have had serious consequences, in light of the Executive’s comments regarding an extension of the pre-emptive right to buy to limited partnerships.¹³⁹

Whether the action of landlords, attempting to avoid an imminent extension of severely punitive measures, was ‘immoral’ is a separate question; and whether that would have impacted upon the Lord Justice-Clerk’s view that section 72 was defective is improbable: the burden created by the measure was, in any case, ‘excessive’. The court did not believe that section 72 reasonably related to the overall aims of the legislation and concluded, therefore, that this provision *did* violate Article 1 Protocol 1 of the appellant’s Convention rights.¹⁴⁰

¹³² *ibid.*

¹³³ Scottish Executive, *Draft Agricultural Holdings (Scotland) Bill – Consultation Paper* (2002) 11 <www.gov.scot/Resource/Doc/46905/0029913.pdf> accessed 23 February 2017.

¹³⁴ Scottish Executive, *Agricultural Holdings: Proposals for Legislation* (HMSO 2000) para 2.9.

¹³⁵ *Salvesen v Riddell* [2012] CSIH 26, 2013 SC 69 [93] (LJ-C Gill).

¹³⁶ BBC News, ‘Minister’s U-turn over Land Purchase’ (*BBC News*, 16 April 2002) <news.bbc.co.uk/1/hi/scotland/1932255.stm> accessed 21 March 2016; T Peterkin, ‘Tenant Farmers May Be Given the Right to Buy’ *The Telegraph* (London, 1 February 2002) <www.telegraph.co.uk/news/uknews/1383405/Tenant-farmers-may-be-given-the-right-to-buy.html> accessed 21 March 2016; ‘Tenant Farmers Press Finnie on Absolute Right-to-Buy’ *The Scotsman* (Edinburgh, 29 May 2002) <www.scotsman.com/business/companies/tenant-farmers-press-finnie-on-absolute-right-to-buy-1-568963> accessed 21 March 2016.

¹³⁷ *Salvesen v Riddell* [2012] CSIH 26, 2013 SC 69 [93] (LJ-C Gill).

¹³⁸ *ibid* [94] (LJ-C Gill).

¹³⁹ Scottish Executive, *Draft Agricultural Holdings (Scotland) Bill – Consultation Paper* (2002) para 4.5 <www.gov.scot/Resource/Doc/46905/0029913.pdf> accessed 23 February 2017.

¹⁴⁰ *Salvesen v Riddell* [2012] CSIH 26, 2013 SC 69 [98] (LJ-C Gill).

Where questions of legislative competence arise with regard to the ECHR, the Human Rights Act 1998 then requires the court to consider whether any reading can be given to the offending provision to make it compatible with Convention rights.¹⁴¹ With the existence of both the 'transition' and 'successor regime' and their starkly contrasting results, section 72 could only be read as arbitrary and, therefore, it fell outwith the competence of the Scottish Parliament.¹⁴²

B. The Supreme Court and Remedial Order

The Court of Session granted leave to the Scottish Government to appeal this decision to the Supreme Court.¹⁴³ It was argued that the decision of the Inner House was premature as the case had not yet been remitted to the Land Court to allow for proof of Salvesen's averment that he had intended to end the tenancy for amalgamation purposes, rather than to deprive the tenant of any of his rights.¹⁴⁴ The court rejected this view, holding that it was no longer necessary to remit the case to the Land Court as the parties had reached an agreement.¹⁴⁵ In any case, it was necessary, in the public interest, to consider the violation that section 72 posed to the ECHR.

The Supreme Court's approach to the issue differed from Lord Gill's, as the dictum focussed on the discriminatory element of sections 72 and 73, rather than the purposes of the legislation as found in the statements of Scottish Ministers.¹⁴⁶ It was as clear to the Supreme Court as it had been to Lord Gill that the differing treatment of landlords lacked any proper justification and could only be seen as arbitrary, which confirmed that Article 1 Protocol 1 had been violated.¹⁴⁷ The court was careful to recognise, however, that both these provisions had, at the time of the decision, been in existence for ten years, meaning that affected parties had relied thereon.¹⁴⁸ Indeed, the Scottish Government identified several parties who had sold property as a result of these provisions;¹⁴⁹ therefore, it was necessary to avoid any legal uncertainty by limiting the defectiveness of the legislation to sub-section 72(10).¹⁵⁰ It was this particular provision which determined whether a limited partnership was to be treated under the arbitrary 'transition regime', or the more lenient 'successor regime', as found under sections 72 and 73 respectively.¹⁵¹

The court recognised, however, that any order seeking to rectify section 72(10) with retrospective effect would raise issues of legal certainty for tenants who had relied thereon.¹⁵² These tenants would have invested in their holding with the

¹⁴¹ Human Rights Act 1998, s 3.

¹⁴² Scotland Act 1998, s 29(1); *Salvesen v Riddell* 2013 SC 69 [102]-[103] (LJ-C Gill).

¹⁴³ A Fox, 'Holyrood out of Bounds' (2013) 58(6) JLSS 28-29.

¹⁴⁴ *ibid.*

¹⁴⁵ *Salvesen v Riddell* [2013] UKSC 22, 2013 SC 236 [25] (Lord Hope).

¹⁴⁶ *ibid* [39].

¹⁴⁷ *ibid* [44].

¹⁴⁸ *ibid* [51].

¹⁴⁹ Tenant Farming SB 14/52, 25 July 2014, 17.

¹⁵⁰ *Salvesen v Riddell* [2013] UKSC 22, 2013 SC 236 [50] (Lord Hope).

¹⁵¹ M Combe, 'Peaceful Enjoyment of Farmland at the Supreme Court' [2013] SLT 201-203.

¹⁵² *Salvesen v Riddell* [2013] UKSC 22, 2013 SC 236 [55] (Lord Hope).

expectation of retaining full security of tenure, as purported by section 72 of the 2003 Act.¹⁵³ It was argued in response, however, that an order that avoided this outcome by rectifying the provision in a prospective manner would prevent landlords from recovering vacant possession of their property, meaning that the law would still be in violation of the ECHR.¹⁵⁴

In light of this, it was necessary to allow the Scottish Parliament, as the democratic representative of the Scottish people, to consult stakeholders as to the optimum means of rectifying the offending provision; the Supreme Court suspended its judgment for twelve months to allow for this discussion.¹⁵⁵ On 22 November 2013, the Scottish Government brought forward the Draft Agricultural Holdings (Scotland) Act 2003 Remedial Order 2014, which sought to balance tenants' rights with those of the landlords affected by the retrospective nature of the order;¹⁵⁶ however, the compromise made as a result of this was controversial.

The Remedial Order sought to afford parties affected by the illegal provision the same route given to landlords under the 'successor' regime. This enabled landlords to reclaim vacant possession through notice procedure, as well as providing a 12-month 'cooling-off' period.¹⁵⁷ However, it would not extend section 73 to landlords who had made bilateral agreements as a result of the defect; therefore, where landlords had sold property or had granted new tenancies, this was deemed by the Scottish Government as moving 'beyond the defect'.¹⁵⁸ Certain advisory groups agreed with the proposal, believing that this upheld Lord Hope's dictum in *Salvesen v Riddell*, where the Supreme Court suggested that the Remedial Order should avoid interference with 'settled transactions'.¹⁵⁹ Therefore, even where the parties remained the same in subsequent contractual relationships, such as in an agreement to continue a 1991 Act tenancy, or to grant a Limited Duration Tenancy, this would leave the landlord with no recourse in the face of section 73, despite being unduly forced into these arrangements by the illegal 'transition' regime in the first instance.

One can understand the limitation of this remedial measure where the tenanted property had been sold on, as a result of which the parties involved in the initial agreement were no longer contractually related. However, it seems unjust to prevent such recourse in instances where the original parties remain the same in simply a new bilateral agreement.¹⁶⁰ Indeed, one commentator argues that this fails to recognise that section 72(10) was defective from its enactment.¹⁶¹

¹⁵³ *ibid* [55].

¹⁵⁴ *ibid* [54].

¹⁵⁵ *ibid*.

¹⁵⁶ Scottish Government, *The Agricultural Holdings (Scotland) Act 2003 Remedial Order 2014* (2013) para 6 <www.gov.scot/resource/0043/00438820.pdf> accessed 23 February 2017.

¹⁵⁷ *ibid*.

¹⁵⁸ *ibid*.

¹⁵⁹ Law Society of Scotland, 'Written Submission from Law Society of Scotland' (2014) <www.scottish.parliament.uk/S4_RuralAffairsClimateChangeandEnvironmentCommittee/General%20Documents/ECHR_-_Law_Society_of_Scotland.pdf> accessed 21 March 2016.

¹⁶⁰ A Fox, 'Salvesen: The Proposed Fix' (Journal Online, 2013) <www.journalonline.co.uk/Magazine/58-12/1013450.aspx> accessed 21 March 2016.

¹⁶¹ *ibid*.

In any case, the cumulative effect of *Salvesen v Riddell* has been to undermine one of the key objectives of the 2003 Act in its purported attempt to stimulate the agricultural tenanted sector. It is no surprise that recent research into the sector has revealed that one of the largest deterrents to landowners considering bringing forward land to the letting market is the lack of legislative stability.¹⁶² Lord Gill's critique of the defective section 72 revealed the drastic and ill-conceived measures which the Scottish Government had devised to ensure compliance with their statutory regime, showing complete disregard for the financial implications suffered by landlords as a result of this measure.

Furthermore, Lord Hope confirmed that the Scottish Parliament had illegally enacted measures which interfered with landlords' Convention rights, which, as one commentator observed, is 'hardly a ringing endorsement of the exercise of [its] devolved powers'.¹⁶³ However, one can go much further in arguing that this may have jeopardised any remaining confidence which landlords may have had in the legislature. This is particularly so when an institution believes that severely punishing one community is a furtherance of balancing the rights of two groups. One might even argue that this has revealed the Scottish Parliament's bias. It is no surprise, therefore, that 81 per cent of respondents to a recent survey of landowners perceive a lack of legislative confidence as the greatest challenge facing the agricultural tenanted sector.¹⁶⁴

4. Land Reform (Scotland) Act 2016

In light of the controversy surrounding the Agricultural Holdings (Scotland) Act 2003 and the failure of the Scottish Government to strike the correct balance of rights between tenants and landlords, as outlined above, it cannot come as a surprise that recent statistics have shown the inadequacy of this legislation to achieve one of its primary aims – secure greater levels of tenancies.¹⁶⁵ Indeed, within a decade of the 2003 Act's enactment, the proportion of agricultural land that is rented has dropped from 28.6 per cent to 22.2 per cent.¹⁶⁶ It is also no surprise that further legislative reform of agricultural law was necessary to correct this problem. The Cabinet Minister announced in 2013 that a group of experts had been appointed to conduct a review of agricultural holdings legislation, with the remit of finding the appropriate means to reinvigorate the tenanted sector, where the 2003 Act had so obviously

¹⁶² Scottish Government, *Review of Agricultural Holdings Legislation* (Interim Report, Scottish Government 2014) para 87.

¹⁶³ D Carr, 'Not Law (But Not Yet Effectively Not Law)' (2013) 17(3) Edin LR 370-76.

¹⁶⁴ C Martin and others, *Renting-Out Agricultural Land in Scotland Survey* (Scottish Government Social Research 2014) 52.

¹⁶⁵ Scottish Executive, *Draft Agricultural Holdings (Scotland) Bill – Consultation Paper* (2002) 2 <www.gov.scot/Resource/Doc/46905/0029913.pdf> accessed 23 February 2017.

¹⁶⁶ Scottish Government, *Review of Agricultural Holdings Legislation* (Interim Report, Scottish Government 2014) table 1, 22.

failed.¹⁶⁷ After the publication of the Review Group's recommendations in January 2015, many of these proposals were incorporated into the Land Reform Bill, which was passed by the Scottish Parliament on 16 March 2016 and which received Royal Assent on 22 April 2016.¹⁶⁸

A. The Absolute Right to Buy (ARTB) in the Wake of *Salvesen v Riddell*

Despite the clear misgivings expressed in the debate of ARTB in the previous reform of the tenanted sector, it was announced by the Cabinet Secretary that AHLRG would specifically consider the measure.¹⁶⁹ This came much to the disappointment of the key stakeholder group, Scottish Land and Estates (SLE), formerly the SLF, prompting the then chairman to state that, 'allowing the absolute right to buy to be brought back to the table when it has been rejected by the farming unions and landowners alike flies in the face of listening to the farming industry'.¹⁷⁰

In this regard, an industry-led body was created by the 2003 Act to consider certain issues regarding the tenant farming sector, with the remit of, 'promoting a healthy farm tenanted sector in Scotland'.¹⁷¹ The TFF, consisting of the major stakeholders in the industry, was commissioned by the Scottish Government in 2007 to investigate the greatest barriers obstructing access to the sector for new entrants.¹⁷² After the publication of the Report into new entrant barriers, the TFF then sought to make recommendations to the Cabinet Secretary.¹⁷³ Recommendation four made it clear that, '[a]ll TFF members recognise the importance of (...) evidence (...) regarding the perceived uncertainty about the "right to buy" legislation and the impact that this is having on the confidence of the owners [considering letting land]'.¹⁷⁴ It was recommended, therefore, that there should not be an extension of the right to buy beyond that embodied in the 2003 and 1991 Acts.¹⁷⁵ The then chairman of the TFF, Jeff Maxwell, publicly announced that it was, 'up to the [Cabinet Secretary] to make an unequivocal statement about the Scottish Government's intentions in relation to the right to buy'.¹⁷⁶

¹⁶⁷ Scottish Government, 'Agricultural Holdings Legislation Review' (2013) <<http://news.scotland.gov.uk/News/Agricultural-Holdings-Legislation-Review-5f2.aspx>> accessed 21 March 2016.

¹⁶⁸ Land Reform (Scotland) Act 2016 (asp 18), preamble.

¹⁶⁹ G Mackenzie, 'Tenants' Absolute Right to Buy Back on the Cards' (Farmers' Weekly 2013) 18.

¹⁷⁰ Scottish Land and Estates, 'Landowners' Response to Cabinet Secretary's Announcement on Absolute Right to Buy' (2013)

<www.scottishlandandestates.co.uk/index.php?option=com_content&view=article&id=2631:-landowners-response-to-cabinet-secretarys-announcement-on-absolute-right-to-buy&catid=71:national&Itemid=107> accessed 21 March 2016.

¹⁷¹ S Notley, *Scottish Agricultural Law Handbook* (Avizandum 2009) 7.

¹⁷² Tenant Farming Forum, 'Tenant Farming Forum Consultation Helping New Entrants into Agriculture' (2007) 1

<www.tenantfarmingforum.org.uk/eblock/services/resources.ashx/000/241/600/Helping_New_Entrants_into_Agriculture.pdf> accessed 23 February 2017.

¹⁷³ Tenant Farming Forum, 'Assisting New Entrant into Scottish Farming: Recommendations to the Cabinet Secretary' (2008) 2.

¹⁷⁴ *ibid* 6.

¹⁷⁵ *ibid*.

¹⁷⁶ I Ashbridge, 'Scots "Right to Buy" Rules a Barrier to New Entrants' (Farmers' Weekly 2008) 21.

The Scottish Government responded to this by recognising that confidence was key to, 'encouraging owners of farmland to offer up land for let to new tenants', and that an extension of the right was not in the, 'best interests of a thriving farm tenanted sector'.¹⁷⁷ Furthermore, the Scottish Government could not justify the interference that this would cause to property rights in terms of public benefit, and this, 'would do nothing to encourage new entrants'.¹⁷⁸ It was confirmed, therefore, that the Scottish Government would not bring forward any extension of the right to buy, clarifying that, '[the] Absolute Right to buy [was] not on [their] agenda'.¹⁷⁹ This statement came with an important caveat, however:

[O]wners should now have the confidence to get on and let land to new tenants. Indeed, we will monitor the uptake of new tenancies in the future and consider further action in the future if more land is not let[.]¹⁸⁰

It was likely, therefore, that the Cabinet Secretary's statement regarding the reconsideration of ARTB came as a response to the fact that there had been no increase in tenanted land in the five years following the statement made by the Scottish Government. Indeed, the announcement came within a week of the latest statistics revealing that tenanted levels had instead decreased.¹⁸¹ This would suggest that the Cabinet Secretary's statement came in retaliation to the failure of landlords to bring more land to the market; furthermore, one could even argue that if any future legislation were to enact ARTB, this could be interpreted as a punitive measure in consideration of the statement made by the Scottish Government above. Besides revealing the unpredictable nature of legislation in this area of law, the inclusion of ARTB as a punitive measure in any future legislation would certainly raise issues of ECHR compliance in light of *Salvesen v Riddell*.¹⁸²

B. The Future of the Absolute Right to Buy (ARTB)

It is interesting to consider, in view of AHLRG's reconsideration of ARTB, whether this measure has a place in the future of agricultural tenancy law in Scotland. ARTB has already been afforded to crofters,¹⁸³ as well as to crofting communities.¹⁸⁴ A similar measure was applied to ultra-long leases to allow for feuars to convert their property interest into dominium,¹⁸⁵ which the Scottish Government believed to be the final legislative instalment fully to remove feudal superiority from Scottish land

¹⁷⁷ Scottish Government, 'Barriers to New Entrants to Farming' (2008) <www.gov.scot/Topics/farmingrural/Rural/rural-land/agricultural-holdings/New-entrants> accessed 21 March 2016.

¹⁷⁸ *ibid.*

¹⁷⁹ *ibid.*

¹⁸⁰ *ibid.*

¹⁸¹ Scottish Government, *Tenanted Agricultural Land in Scotland 2012* (Scottish Government, 12 June 2013) <www.gov.scot/Resource/0042/00424370.pdf> accessed 23 February 2017.

¹⁸² *Salvesen v Riddell* [2012] CSIH 26, 2013 SC 69 [90] (LJ-C Gill).

¹⁸³ Crofters (Scotland) Act 1993, s 12.

¹⁸⁴ Agricultural Holdings (Scotland) Act 2003, pt 3.

¹⁸⁵ Long Leases (Scotland) Act 2012.

tenure.¹⁸⁶ Others would argue, however, that without providing secure 1991 Act tenants with ARTB, Scottish farming will remain within ‘the grip of feudal inheritance’.¹⁸⁷

The major stakeholders, excepting the Scottish Tenant Farming Association (STFA), have consistently asserted that the introduction of ARTB would undermine the confidence of landowners to let land. However, commentators, such as Andy Wightman have argued that this concern lacks any rationale.¹⁸⁸

The STFA, originally as STFAG, had only argued for the extension of ARTB to secure 1991 Act tenants.¹⁸⁹ Furthermore, the SNP, who supported the measure during the 2003 Act’s passage through Parliament, provided its support on the condition that this measure be limited to secure tenancies, recognising, ‘secure heritable tenants as a distinct group of people in Scotland’.¹⁹⁰ The same stance has been consistently maintained by the Cabinet Secretary.¹⁹¹ Wightman argues, therefore, that the effect of this proposal on the confidence of landlords has no justification, considering that the use of secure tenancies had ceased by the 1980s.¹⁹² Indeed, recent statistics have shown that, excepting seasonal leases, the only letting vehicles provided by statute that have increased are the Limited Durations Tenancies.¹⁹³

Wightman argues, furthermore, that the assertion that the introduction of ARTB would lead to a loss of available tenancies on the market also lacks any foundation.¹⁹⁴ The commentator believes that Scotland's level of tenancies could actually increase if the tenure system were shifted from a ‘landlord-tenant’ system to one where owner-occupier farmers rented out to one another, achieved through ARTB.¹⁹⁵ Historical reference to such an approach was cited by Professor James Hunter, explaining that this measure was applied by the UK Government to Ireland in 1903, when the ‘Wyndham’ Act, named after the then Chief Secretary for Ireland,

¹⁸⁶ Scottish Government, *Policy Memorandum – Long Leases (Scotland) Bill* (APS Group Scotland 2012) para 9 <[www.parliament.scot/S4_Bills/Long%20Leases%20\(Scotland\)%20Bill/Policy_Memo.pdf](http://www.parliament.scot/S4_Bills/Long%20Leases%20(Scotland)%20Bill/Policy_Memo.pdf)> accessed 23 February 2017.

¹⁸⁷ A Wightman, ‘Tenant Farming Crisis’ (*Land Matters* 2012) <www.andywightman.com/archives/794> accessed 21 March 2016.

¹⁸⁸ A Wightman, ‘Why Does Norway Have More Tenanted Land than Scotland?’ (*Land Matters* 2013) <www.andywightman.com/archives/3086> accessed 21 March 2016.

¹⁸⁹ ‘Submission from Scottish Tenant Farming Action Group to the Scottish Parliament’ (June 2002) 2 <www.parliament.scot/S4_RuralAffairsClimateChangeandEnvironmentCommittee/General%20Documents/ECHR_-_STFA.pdf> accessed 23 February 2017.

¹⁹⁰ Rural Development Committee 4 February 2003, Scottish Parliament Official Report, col 4188.

¹⁹¹ E Pate, ‘Richard Lochhead Takes Issue of Farm Tenants’ Absolute Right to Buy to New Level’ (*The Courier* 2014) <www.thecourier.co.uk/news/richard-lochhead-takes-issue-of-farm-tenants-absolute-right-to-buy-to-new-level-1.132911> accessed 21 March 2016.

¹⁹² A Wightman, ‘Why Does Norway Have More Tenanted Land than Scotland?’ (*Land Matters* 2013) <www.andywightman.com/archives/3086> accessed 21 March 2016.

¹⁹³ Scottish Government, *Review of Agricultural Holdings Legislation* (Interim Report, Scottish Government 2014) 27.

¹⁹⁴ A Wightman, ‘Why Does Norway Have More Tenanted Land than Scotland?’ (*Land Matters* 2013) <www.andywightman.com/archives/3086> accessed 21 March 2016.

¹⁹⁵ *ibid.*

allowed tenant farmers to buy their holdings with Government support.¹⁹⁶ Considering, however, that Ireland now has a lower level of tenancies than Scotland, this would seem to undermine Wightman's argument.¹⁹⁷

Wightman argues further that in other countries where systems of owner-occupation exist, tenancy levels remain higher, such as in France and Germany with 62 per cent and 74 per cent respectively,¹⁹⁸ in stark contrast to 23 per cent in Scotland.¹⁹⁹ The Agricultural Holding Review Group warned, however, that 'international comparisons have to be made with care',²⁰⁰ considering that there are various exogenous factors affecting land tenure in those countries, such as varying regulatory control, land prices and agricultural business structures.²⁰¹ Furthermore, these arguments fail to take into account that Scotland's current system of land tenure is already dominated by owner-occupiers: in 2013, the Scottish Government recorded that 71 per cent of all holdings were owned by their occupiers.²⁰² Despite this rise, there has been no mitigation of the reduction of rented land.²⁰³ It seems unlikely, therefore, that a system of tenure, relying purely upon owner-occupiers, would successfully deliver the land desperately needed for new entrants. Indeed, the Agricultural Holdings Review Group stated that it could find no evidence to support Wightman's claim; rather it seemed more likely that owner-occupiers would sell to other owner-occupiers.²⁰⁴

Wightman and Hunter have argued, however, that once the transition has been made to a system of farmer-to-farmer rental, freedom of contract could be made available as a means of encouraging greater levels of letting.²⁰⁵ That said, neither commentator has attempted to explain why freedom of contract could not be afforded to landlords under the current system. It is clear that, without such an explanation or justification, the proposal could not be enacted without engaging Article 1 Protocol 1 ECHR.

Some justification for the proposal might be addressed in the Land Reform Review Group's (LRRG) consideration of secure 1991 Act tenancies in its Final

¹⁹⁶ J Hunter, 'Land Reform – Let's Have the Lochhead Act!' (The Scottish Farmer 2013) <www.thescottishfarmer.co.uk/news/land-reform-lets-have-the-lochhead-act.21330968> accessed 21 March 2016.

¹⁹⁷ Scottish Government, *Review of Agricultural Holdings Legislation* (Interim Report, Scottish Government 2014) fig 4, 33.

¹⁹⁸ A Wightman, 'Why Does Norway Have More Tenanted Land than Scotland?' (*Land Matters* 2013) <www.andywightman.com/archives/3086> accessed 21 March 2016.

¹⁹⁹ Scottish Government, *Tenanted Agricultural Land in Scotland 2014* (Statistical Publication Agriculture Series 2015) 4.

²⁰⁰ Scottish Government, *Review of Agricultural Holdings Legislation* (Interim Report, Scottish Government, Scottish Government 2014) para 84.

²⁰¹ *ibid* paras 77-78.

²⁰² Land Reform Review Group, *The Land of Scotland and the Common Good* (Scottish Government 2014) fig 33, 197.

²⁰³ *ibid*.

²⁰⁴ Scottish Government, *Review of Agricultural Holdings Legislation* (Final Report, Scottish Government 2015) para 202.

²⁰⁵ A Wightman, 'Why Does Norway Have More Tenanted Land than Scotland?' (*Land Matters* 2013) <www.andywightman.com/archives/3086> accessed 21 March 2016.

Report.²⁰⁶ Although AHLRG were tasked with finding the optimum means of achieving a ‘sustainable’ and ‘dynamic’ tenanted sector, this ultimately involved balancing the rights of landlords and tenants for the good of agriculture.²⁰⁷ The focus of LRRG was much wider, considering measures which would make ‘stronger, more resilient and independent’ rural communities.²⁰⁸ LRRG highlighted the value of the inter-generational continuity which these secure 1991 Act tenancies provide, explaining that the resilience of rural communities operates in correlation with the density of kinships ties.²⁰⁹ This highlights the importance of these secure 1991 Act tenants, who are often born into land tenure by succession to these tenancies.²¹⁰ Indeed, in a recent survey of tenant farmers, it was found that 50 per cent of respondents and their families had held the same lease between, or over, 50 to 175 years.²¹¹

Furthermore, submissions to LRRG highlighted that as a result of the Scottish Parliament’s failure to enact ARTB for secure tenants, rural communities have continually deteriorated, particularly in the Highlands and Islands, where large estates, owned by overseas companies, were effectively diverting money from local economies through rents.²¹² LRRG recognised that, as the current debate regarding the right to buy continues, land agents are attempting to find any technicalities available to retrieve vacant possession of holdings at the cost of valuable social capital.²¹³ Indeed, one submission claims that if many of the holdings on Islay and Jura had been owner-occupied, local businesses previously supported by secure tenants would have remained open, preventing the increased depopulation of the islands.²¹⁴ LRRG considered, therefore, that the continuing deterioration of rural communities was not in the wider public interest, justifying the extension of a right to buy in appropriate circumstances.²¹⁵ The group rejected the idea of an ‘absolute’ right to buy, preferring the notion of an ‘actual’ right to buy, which would only be provided to tenants on a case-by-case basis, considering the public benefit that the enforcement of a sale would provide; this was with a view to countering the concerns of ECHR compliance.²¹⁶

Similarly, AHLRG considered a conditional right to buy, but the justification for this was premised upon balancing the rights between landlord and tenant.

²⁰⁶ Land Reform Review Group, *The Land of Scotland and the Common Good* (Scottish Government 2014) 202-207.

²⁰⁷ *ibid* para 28.

²⁰⁸ *ibid* para 34.

²⁰⁹ *ibid*.

²¹⁰ *ibid*.

²¹¹ Scottish Government, *Survey of Agricultural Tenant Farmers* (Scottish Government Social Research 2014) fig 2.2, 6.

²¹² Land Reform Review Group, ‘A Tenant Farmer’s Story’ <www.gov.scot/Resource/0043/00432634.pdf> accessed 21 March 2016.

²¹³ Land Reform Review Group, *The Land of Scotland and the Common Good* (Scottish Government 2014) 204.

²¹⁴ Land Reform Review Group, ‘A Tenant Farmer’s Story’ <www.gov.scot/Resource/0043/00432634.pdf> accessed 21 March 2016.

²¹⁵ Land Reform Review Group, *The Land of Scotland and the Common Good* (Scottish Government 2014) 206.

²¹⁶ *ibid*.

Tenants under the previous 2003 Act regime lacked a right to enforce contractual obligations owed by the landlord to the tenant. In contrast, the landlord could serve an incontestable notice to quit where the tenant had failed to remedy a contractual breach brought to their attention.²¹⁷ This form of a ‘conditional’ right to buy was incorporated into the Land Reform (Scotland) Act 2016 (LRSA).²¹⁸

No such provision was made, however, for the ‘actual’ right to buy considered by the LRRG for the prevention of wider rural depopulation. Indeed, AHLRG found that there was not enough evidence to support either the argument that owner-occupation leads to greater investment in holdings,²¹⁹ or that the introduction of a right to buy would lead to an increase in lettings.²²⁰ Without such evidence, there was no clear public interest argument for the extension of the right in this regard.²²¹ Instead, the group found arguments, ‘calling for the idea to be set firmly aside in order to bring confidence back into the market’ more persuasive.²²²

Despite AHLRG’s call for the issue of ARTB to be ‘put to bed’ once and for all, Scottish Ministers expressed concerns during the Land Reform Bill’s passage through the Scottish Parliament that the proposals embodied therein did not go far enough towards achieving ‘radical reform’ for land tenure in Scotland, with some members of the SNP having demanded the enactment of ARTB.²²³ On the other hand, the Conservative MSP Alex Fergusson claimed that the Bill did not achieve the primary policy objective of securing an increase in lettings, arguing instead that it would contribute further to the erosion of landlords’ trust in the legislature.²²⁴ Despite these concerns, the majority of Ministers on the Rural Affairs, climate Change and Environment (RACCE) Committee shared the view that some form of a conditional right to buy, similar to that proposed by LRRG, should be explored by the Scottish Government.²²⁵

The issue was raised again within a Stage 2 Committee, where Michael Russell MSP brought forward amendment 293 to provide ARTB for certain 1991 Act tenants; however, the minister withdrew the amendment, explaining that his intention had merely been to highlight the ongoing debate concerning the right to buy, and the need for the Scottish Parliament to address this issue once and for all.²²⁶ Russell explained that this issue had arisen from the failure of the UK Government to apply the same changes that had been implemented in other countries across

²¹⁷ Scottish Government, *Review of Agricultural Holdings Legislation* (Final Report, Scottish Government 2015) para 192.

²¹⁸ Land Reform (Scotland) Act 2016, s 100.

²¹⁹ *ibid* para 200.

²²⁰ *ibid* para 202.

²²¹ *ibid* para 206.

²²² *ibid*.

²²³ S Johnson, ‘SNP Members Rebel in Demand for Radical Land Reform’ *The Telegraph* (London, 16 October 2015) <www.telegraph.co.uk/news/politics/SNP/11937275/SNP-members-rebel-in-demand-for-radical-land-reform.html> accessed 21 March 2016.

²²⁴ RACCE, ‘Report Stage 1 on Land Reform (Scotland) Bill’ (2015) para 400 <www.parliament.scot/S4_RuralAffairsClimateChangeandEnvironmentCommittee/Reports/RACCES042015R10Rev.pdf> accessed 23 February 2017.

²²⁵ *ibid* para 561.

²²⁶ RACCE 3 February 2016, Scottish Parliament Official Report, cols 52-54.

Europe at the turn of the twentieth century.²²⁷ It was only now, with the inception of the Scottish Parliament in 1999, that Scotland was having to cope with these changes.²²⁸ The SNP Minister argued, therefore, that it was imperative to take the next step in land reform and provide secure tenants with the right to buy in circumstances where they could prove investment in their holdings and communities, in order to allow these tenants to move forward.²²⁹ However, the question would then arise as to how a legal construct could be applied to test ‘investment in a community’. Russell suggested as a qualification for this measure that a tenant, or their family, must have farmed the holding for a minimum period of fifty years.²³⁰

The issue with this proposal is that it is impossible to reconcile the consequence of this qualification with the aim of the Scottish Parliament to encourage landowners to create longer-term tenancies. Indeed, one could argue that such an approach would be ‘schizophrenic’, as the chairman of SLE did, explaining that, ‘on the one hand you have a Bill to create a vibrant tenanted sector and they are trying to encourage people to let for the long term, because we fundamentally believe (...) it gives parties security’; however, proposals, such as the conditional right to buy, suggest that, ‘those who have already let for the long term in the past’ will be punished for doing so.²³¹

The Scottish Government rejected the calls for ARTB, relying on AHLRG’s findings that such a measure would be outwith the legislative competence of the Scottish Parliament, and that the Scottish Government wished to avoid another *Salvesen v Riddell* scenario.²³² It believed instead that other measures now contained in LRSA would go some way towards strengthening tenants’ rights, such as the conditional right to buy and assignation, discussed further below.²³³ The Cabinet Secretary did accept, however, that this debate was likely to surface again in the near future.²³⁴

C. Late Amendments and the Land Reform (Scotland) Act 2016

With the Scottish Parliament’s continued interest in the introduction of ARTB, it would appear that the lessons from *Salvesen* have not been learnt. In addition, history seems to have repeated itself in another way with the late adoption of a controversial measure into Part 10 of the Land Reform Bill.²³⁵ Under the old

²²⁷ *ibid.*

²²⁸ *ibid.*

²²⁹ *ibid.*

²³⁰ SP Bill 76-ML3 Land Reform (Scotland) Bill [3rd Marshalled List of Amendments for Stage 2] Session 4 (2016).

²³¹ A Cramb, ‘Property Rights Threatened by “Aggressive” Change in Land Reform Bill’ *The Telegraph* (London, 18 December 2015) <www.telegraph.co.uk/news/earth/countryside/12057440/Property-rights-threatened-by-aggressive-change-in-land-reform-bill.html> accessed 21 March 2016.

²³² RACCE 3 February 2016, Scottish Parliament Official Report, cols 60-61.

²³³ *ibid.*

²³⁴ *ibid.*

²³⁵ SP Bill 76-ML4 Land Reform (Scotland) Bill [4th Marshalled List of Amendments for Stage 2] Session 4 (2016), amendment 325.

legislation, landlords might have expected to retrieve vacant possession of their land where their 1991 Act tenant was unable to find an eligible successor or assignee of his tenancy. Where a secure tenant sought to assign their tenancy, the 2003 Act required that the assignee fall under the same class of persons eligible to succeed to that tenancy.²³⁶ To be assigned a secure tenancy, therefore, the assignee would have had to fall within a class of 'near relatives', excluding cousins, nieces and nephews, and siblings.²³⁷ The prospect of retrieving vacant possession from a secure 1991 Act tenant seemed particularly likely in the modern era: it was highlighted by a report from the TFF that, as a result of ubiquitous access to tertiary education, potential successors no longer had the same attraction to returning to the family farm as they had had in the past.²³⁸

It was highlighted by AHLRG that there was a concern that the current lack of 'willing' or 'eligible' successors was inhibiting older tenant farmers from planning retirement, which in turn prevented the early release of land to be farmed at its full capacity by new entrants.²³⁹ The lack of successors, and of secure pension provision, led AHLRG to recommend that secure tenants be afforded the right to assign their tenancies for value on the open market.²⁴⁰ Importantly, however, this could only be achieved through the conversion of the 1991 Act tenancy to the new 'Modern Limited Duration Tenancy', removing security of tenure and replacing it with a fixed term agreement.²⁴¹

This measure was incorporated into the Land Reform Bill as it made its way through the Scottish Parliament;²⁴² however, this did not come without criticism. SLE provided evidence that if these measures were to be imposed on landlords, this could cost the public purse in excess of £600 million through compensation claims against the Scottish Government;²⁴³ furthermore, assignation of secure tenancies, albeit after conversion to LDT, would amount to losses of £279 million to landowners.²⁴⁴ The only consolation that landowners could derive from such a

²³⁶ Agricultural Holdings (Scotland) Act 2003, s 10A.

²³⁷ Scottish Executive, *Land Reform (Scotland) Bill – Policy Memorandum* (APS Group Scotland 2015) para 372 <[www.parliament.scot/S4_Bills/Land%20Reform%20\(Scotland\)%20Bill/b76s4-introd-pm.pdf](http://www.parliament.scot/S4_Bills/Land%20Reform%20(Scotland)%20Bill/b76s4-introd-pm.pdf)> accessed 23 February 2017.

²³⁸ P Cook and others, 'Barriers to New Entrants to Scottish Farming: An Industry Consultation for the Tenant Farming Forum' (2008) para 1.11.3 <www.tenantfarmingforum.org.uk/eblock/services/resources.ashx/000/244/597/58_final_report_from_contractors.pdf> accessed 23 February 2017.

²³⁹ Scottish Executive, *Land Reform (Scotland) Bill – Policy Memorandum* (2015) para 372 <[www.parliament.scot/S4_Bills/Land%20Reform%20\(Scotland\)%20Bill/b76s4-introd-pm.pdf](http://www.parliament.scot/S4_Bills/Land%20Reform%20(Scotland)%20Bill/b76s4-introd-pm.pdf)> accessed 23 February 2017.

²⁴⁰ Scottish Government, *Review of Agricultural Holdings Legislation* (Final Report, Scottish Government 2015) para 173.

²⁴¹ *ibid* recommendation 15, para 177.

²⁴² SP Bill 76 Land Reform (Scotland) Bill [as introduced] Session 4 (2015), s 79.

²⁴³ Scottish Land and Estates, 'Landowners Highlight Tenant Farming Risks and Opportunities to MSPs' (2015)

<www.scottishlandandestates.co.uk/index.php?option=com_content&view=article&id=4152:landowners-highlight-tenant-farming-risks-and-opportunities-to-msps&catid=71:national&Itemid=107> accessed 21 March 2016

²⁴⁴ S Gore, 'Effect of Values of Extending Succession and Assignment to Holdings Subject to the Agricultural Holdings (Scotland) Act 1991' (*Land Matters* 2014) 36

proposal would be the prospect of a fixed term tenancy; however, this would sometimes occur in circumstances where the tenant would have had no prospect of succession in the first place.²⁴⁵

However, the prospect of change to the rules governing assignation was to be made even worse for landlords: the Scottish Government wrote to the RACCE Committee to inform them of their proposal to allow secure tenants to assign their 1991 Act tenancies without any conversion to a fixed term tenancy.²⁴⁶ To counter-balance this, the new provisions of ‘assignation and relinquishment’ would offer landlords a right of pre-emption where the outgoing tenant notified the landowner of their intention to relinquish the tenancy,²⁴⁷ which the landlord would have to seize if they were to have any prospect of retrieving vacant possession. Despite this opportunity to purchase their land back from the tenant, it was highlighted by one writer that this proposal would undermine the confidence of landlords.²⁴⁸ The amendment was incorporated into the Bill at Stage 2, after the agreement of the RACCE Committee was obtained,²⁴⁹ and can now be found under Part 3A of the amended 1991 Act, despite claims that this measure ran the risks of ‘jeopardising’ the agricultural tenanted sector and breaching the ECHR. Indeed, this proposal had been considered by AHLRG, but the group rejected it on the basis that a public interest argument for this provision had not been established.²⁵⁰

This move could be commended for fulfilling one of the main policy aims of the legislation by providing opportunities to new entrants to gain access to the industry, whilst retaining the benefits of full security of tenure.²⁵¹ However, the initial proposal to afford open market assignation after conversion already achieved this aim without affording a right of pre-emption to landlords. Indeed, SLE stated that this proposal would fail to meet these aims, given that sale of the tenancy to the landlord would be the most likely outcome of this provision: they would have more capital than the new entrant and would likely absorb the costs of the purchase to

<http://www.andywightman.com/docs/SLE_smithsgore_valuation.pdf> accessed 27 February 2017.

²⁴⁵ A Cramb, ‘Property Rights threatened by “Aggressive” Change in Land Reform Bill’ *The Telegraph* (London, 18 December 2015) <www.telegraph.co.uk/news/earth/countryside/12057440/Property-rights-threatened-by-aggressive-change-in-land-reform-bill.html> accessed 21 March 2016.

²⁴⁶ Scottish Government, ‘Land Reform (Scotland) Bill: Proposed Stage 2 Amendment to Section 79’ (2015)

<www.scottish.parliament.uk/S4_RuralAffairsClimateChangeandEnvironmentCommittee/General%20Documents/20151204_Scot_Gov_on_amendments_to_section_79_LRB.pdf> accessed 21 March 2016

²⁴⁷ Agricultural Holdings (Scotland) Act 1991, pt 3A.

²⁴⁸ A Cramb, ‘Property Rights Threatened by “Aggressive” Change in Land Reform Bill’ *The Telegraph* (London, 18 December 2015) <www.telegraph.co.uk/news/earth/countryside/12057440/Property-rights-threatened-by-aggressive-change-in-land-reform-bill.html> accessed 21 March 2016.

²⁴⁹ RACCE 10 February 2016, Scottish Parliament Official Report, col 34.

²⁵⁰ Scottish Government, *Review of Agricultural Holdings Legislation* (Final Report, Scottish Government 2015) para 180.

²⁵¹ Scottish Tenant Farming Association, ‘Written Submission from Scottish Tenant Farmers Association: Replacement to Section 79 of the Land Reform Bill at Stage 2’

<[www.scottish.parliament.uk/S4_RuralAffairsClimateChangeandEnvironmentCommittee/General%20Documents/\(010\)_Scottish_Tenant_Farmers_Association.pdf](http://www.scottish.parliament.uk/S4_RuralAffairsClimateChangeandEnvironmentCommittee/General%20Documents/(010)_Scottish_Tenant_Farmers_Association.pdf)> accessed 21 March 2016.

retain vacant possession.²⁵² Furthermore, Michael Russell MSP, despite supporting the measure in the Stage 2 Committee, admitted that this would do nothing to achieve the Bill's policy of increasing lets; rather, it would undermine it.²⁵³ As seen with the 2003 Act and the subsequent case of *Salvesen v Riddell*,²⁵⁴ the Scottish Parliament has, once again, brought forward a measure which has been introduced at a late stage in the legislative process, and contradicts the advice and expertise offered to achieve a healthy tenanted sector.

5. Conclusion

Considering the unpredictable nature of legislative reform in this area of law, and the repeated contradictory changes that the Scottish Government has made to its own policies, it is understandable why landlords have such little faith in the competence of the Scottish Parliament to achieve a fair balance of rights for tenants and landlords. Indeed, in a system of land tenure which requires the confidence of both communities for the sector to flourish, it seems that the Scottish Parliament has struggled to understand the difference between ideological aspirations for tenants, and those measures which will truly find a balance for the good of the tenanted sector. Unfortunately, the ill-conceived measures proposed under the Agricultural Holdings (Scotland) Act 2003 have impeded the invigoration of the tenanted sector.

It was to be hoped, therefore, that the Scottish Ministers would learn from the consequences of *Salvesen v Riddell* and recognise the importance of equitable legislation. Instead, the Scottish Government has once again brought forward a hastily prepared,²⁵⁵ ill-conceived, last-minute amendment; one that was not based on the extensive consultation process conducted by the Agricultural Holdings Legislation Review Group. Just as the Scottish Executive did in the 2003 Act's passage, the Scottish Government has produced a proposal which wholly militates against the primary aim of ensuring greater levels of tenancies for new entrants. Landlords are much more likely to be prepared to incur the cost of pre-emption than to allow the tenancy to be transferred with full security of tenure to a new entrant, a factor that undermines the entire purpose behind the provision. More important, however, is the scenario of a landlord who is unwilling to pay that price: the proposals to amend assignation of 1991 Act tenancies may induce another *Salvesen v Riddell* scenario, considering the excessive burden that this provision places on landlords' property. This is a radical alteration of privately-made agreements, which were unlikely to have had any intention of providing secure tenants with as wide a class of assignees as the amendment proposes, albeit such is confined to a 'new

²⁵² Scottish Land and Estates, 'Written Submission from Scottish Land and Estates' <[www.scottish.parliament.uk/S4_RuralAffairsClimateChangeandEnvironmentCommittee/General%20Documents/\(007\)_Scottish_Land_and_Estates.pdf](http://www.scottish.parliament.uk/S4_RuralAffairsClimateChangeandEnvironmentCommittee/General%20Documents/(007)_Scottish_Land_and_Estates.pdf)> accessed 21 March 2016.

²⁵³ RACCE 10 February 2016, Scottish Parliament Official Report, col 27.

²⁵⁴ *Salvesen v Riddell* [2012] CSIH 26, 2013 SC 69.

²⁵⁵ B Henderson and A Arbuckle, 'Fears Land Reform Bill is Being Rushed Through' *The Scotsman* (Edinburgh, 17 December 2015) <www.scotsman.com/business/companies/farming/fears-land-reform-bill-is-being-rushed-through-1-3978038> accessed 21 March 2016.

entrant' or those 'progressing in farming'.²⁵⁶ Although such an alteration can be seen as beneficial in bringing new entrants into the agricultural sector, the perpetuation of secure tenancies favours tenants to such an extent that it jeopardises the future letting of land. This compromises the balance of rights needed to secure confidence of landlords in the letting market, and will ultimately reduce the supply of land available to new entrants, which is essential for affordable access to the agricultural industry.

It is hard to imagine how any legislature will be able to regain the trust of landlords following the events surrounding agricultural tenancies, detailed over the course of this article. However, the provisions developed by the 2003 and 2016 Acts, such as those covering rent reviews, compensation at waygo, amnesty agreements, break clauses for new entrants and fixed term tenancies, to name but a few, do strike the balance required for a healthy tenanted sector. Although the uptake of these tenancies may be slow, as seen with the limited duration tenancies created under the 2003 Act,²⁵⁷ this may improve over time as landlords regain confidence. However, this can only occur if the perpetuation of secure tenancies and the issue of the ARTB can be 'put to bed' once and for all.²⁵⁸ A healthy tenanted sector requires a balancing of rights, something which neither of these measures achieves. If the Scottish Government persists in proposing measures that weigh disproportionately in favour of tenants, the supply of affordable land to lease will likely dry up and the door to the agricultural industry for the next generation of tenant farmers will remain firmly shut.

²⁵⁶ SP Bill 76B Land Reform (Scotland) Bill [As Passed] Session 4 (2016), s 89A(2).

²⁵⁷ Scottish Government, *Review of Agricultural Holdings Legislation* (Interim Report, Scottish Government 2014) 27.

²⁵⁸ E Pate, 'Hamish Lean Adamant on ARTB Stance at NFUS Annual Meeting' (The Courier 2015) <www.thecourier.co.uk/business/farming/news/hamish-lean-adamant-on-artb-stance-at-nfus-annual-meeting-1.832088> accessed 21 March 2016.

Trading Places: Benefits of Invoice Finance for Small and Medium Sized Enterprises as Opposed to Bank Lending

KAHMUN GOH*

Abstract

The contribution of small and medium sized enterprises (SMEs) to the British economy is crucial due to their creation of jobs and their impetus to growth and innovation.¹ In order for SMEs to thrive, they must first have a steady cash flow facilitated through easy access to finance. However, following the financial crisis in 2008, banks have been increasingly hesitant to lend to SMEs and have imposed stricter conditions on their lending.² In addition, SMEs typically experience late payment from their debtors, which affects their cash flow, therefore restricting growth.³ Although bank lending to SMEs has improved in the last few years,⁴ the strict covenants that banks impose on SMEs restrict growth within the business. However, an alternative to bank lending exists in the form of invoice financing. This paper considers the benefits of such financing for SMEs and argues that it is an adequate alternative to the bank loans on which SMEs have traditionally relied.

Keywords: SME Funding Gap, Factoring, Invoice Discounting, Bank Lending

1. Introduction

*'A bank is a place that will lend you money if you can prove that you don't need it.'*⁵
- Bob Hope

In the early twentieth century, Jack Cohen started a small business selling groceries from a market stall in East London.⁶ His first day's sales were £4, resulting in a profit

* The author graduated from the University of Aberdeen with a First Class LLB (Honours) degree in 2016. She is now completing the Diploma in Professional Legal Practice (DPLP) at the University of Glasgow.

¹ Department for Business Innovation & Skills, 'SMEs: The Key Enablers of Business Success and the Economic Rationale for Government Intervention' (BIS Analysis Paper Number 2, December 2013) 9 <www.gov.uk/government/uploads/system/uploads/attachment_data/file/266304/bis-13-1320-smes-key-enablers-of-business-success.pdf> accessed 17 January 2016.

² I McCafferty, 'UK Business Finance since the Crisis – Moving to a New Normal?' (Speech at Bloomberg, London, 20 October 2015) <www.bankofengland.co.uk/publications/Documents/speeches/2015/speech851.pdf> accessed 29 January 2016.

³ T Edmonds, 'Late Payment of Debts' (House of Commons Library Briefing Paper No 6640, February 2016) 4 <<http://researchbriefings.files.parliament.uk/documents/SN06640/SN06640.pdf>> accessed 19 January 2016.

⁴ Bank of England, 'Bank of England and HM Treasury Announce Extension to the Funding for Lending Scheme' (Bank of England, 30 November 2015) 1 <www.bankofengland.co.uk/publications/Pages/news/2015/096.aspx> accessed 18 January 2016.

⁵ A Harrington, *Life in the Crystal Palace* (Alfred A Knopf 1959).

of £1.⁷ Now, approximately ninety years later, this business produces over £50 billion in annual revenue; employs around 500,000 people worldwide and generates £6,000 in profit per minute per day.⁸ That business is known as Tesco Plc. It serves as a striking illustration of the fact that many of the largest corporations in the world started life as small businesses. With a stable economy and access to working capital, they were able to grow into some of the most successful players in their field. Although the percentage of small firms that experience this kind of dramatic growth is minimal, collectively small businesses are vital to the UK economy.⁹

Small businesses are typically grouped together with medium-sized enterprises under the abbreviation 'SME'; both are very important to the economy and tend to be similar in nature.¹⁰ Under the Companies Act 2006, a small business is regarded as one with a turnover of not more than £6.5m and employing no more than fifty people; a medium sized business has an annual turnover of less than £25.9m, employing no more than 250 people.¹¹

SMEs make up 99.9 per cent of all UK businesses and the sector has made its mark in the last five years, supporting high, sustainable growth within the domestic economy and fostering innovation and job creation.¹² Figures show that, in 2015, employment by SMEs totalled 60 per cent of private-sector employment and that the combined annual turnover of SMEs amounted to £1.8 trillion.¹³ Essentially the backbone of the UK economy, it is vital that they have easily accessible working capital with which to maintain a stable cash flow or to expand.

SMEs are typically heavily dependent on bank lending to fund general growth or to fuel a start-up, especially on term-loans and overdrafts, as these are easy loans to obtain and ostensibly the cheapest.¹⁴ However, as a result of the financial crisis in 2008, bank lending has now hit an all-time low.¹⁵ SMEs in

⁶ T Clark and SP Chan, 'A History of Tesco: The Rise of Britain's Biggest Supermarket' *The Telegraph* (London, 4 October 2014) <www.telegraph.co.uk/finance/markets/2788089/A-history-of-Tesco-The-rise-of-Britains-biggest-supermarket.html> accessed 17 January 2016.

⁷ *ibid.*

⁸ Tesco PLC, 'Annual Report and Financial Statements 2015' (Tesco plc 2015) 3 <www.tescopl.com/files/pdf/reports/ar15/download_annual_report.pdf> accessed 21 January 2016.

⁹ Clark and Chan (n 6).

¹⁰ F Schmithusen and others, *Principles of Business Economics and Management Processes* (Routledge 2014) 103.

¹¹ Companies Act 2006, ss 382 and 465.

¹² The Royal Bank of Scotland, 'Independent Lending Review' (Royal Bank of Scotland 2013) 12 <www.independentlendingreview.co.uk/RBS_ILR_Full_Report.pdf> accessed 20 January 2016.

¹³ National Federation of Self Employed & Small Businesses Limited, 'Statistics – SMEs and the Economy' (FSB 2015) <www.fsb.org.uk/media-centre/small-business-statistics> accessed 19 January 2016.

¹⁴ D van der Schans, 'SME Access to External Finance' (BIS Economics Paper No 16, Department for Business Innovation and Skills 2012) 7 <www.gov.uk/government/uploads/system/uploads/attachment_data/file/32263/12-539-sme-access-external-finance.pdf> accessed 17 January 2016.

¹⁵ S Fraser, 'The Impact of the Financial Crisis on Bank Lending to SMEs' (Department for Business Innovation and Skills 2012) 5 <www.gov.uk/government/uploads/system/uploads/attachment_data/file/34739/12-949-impact-financial-crisis-on-bank-lending-to-smes.pdf> accessed 23 January 2016.

particular have been challenged with the difficulty of accessing working capital.¹⁶ In the period from 2008 onwards, 'rejection rates for both overdrafts and term loans were higher than ever'.¹⁷ Due to these circumstances, SMEs are increasingly prone to failure.

With this problem in mind, the present article seeks to propose that asset-based finance is an adequate alternative to bank lending where banks are either hesitant to lend or have tightened their terms of agreement. Whilst it is acknowledged that asset-based finance can take various forms, the focal point for the purposes of this work is invoice finance alone.

2. Problems Faced by SMEs

A. The 2008 Financial Crisis

The financial crisis was predominantly caused by failures in the financial sector as a result of under-regulation and insufficient supervision of the industry.¹⁸ Financial institutions failed to manage their businesses vigilantly and to appreciate the risks inherent in the activities that they were conducting.¹⁹ As a result of these failures, there has been an unprecedented level of worldwide government intervention to increase liquidity and to prevent even more banks from collapsing.²⁰

The crisis first became apparent in the UK during 2008, when the Bank of England provided financial support to the Northern Rock Building Society; the latter faced a liquidity crisis after borrowing large sums of money to fund mortgages for customers.²¹ As the crisis developed, the affected banks' lending capacity was significantly diminished, impelled as they were to readjust their balance sheets.²² Andrew Freeman testifies to this problem, explaining that banks began to deleverage through scanning loan agreements for covenants that the debtor may have breached; where a breach was discovered, the loan was withdrawn or the covenant's restrictions tightened.²³ Freeman describes this process as 'upsetting' for businesses, observing that 'debt was actually behaving like equity – the lender became the de facto owner of the business, able to dictate terms and potentially even put the business into administration'.²⁴

¹⁶ S Fraser, 'Evaluating Changes in Bank Lending to UK SMEs over 2001-12 – Ongoing Tight Credit?' (Department for Business Innovation and Skills 2013) 3

<www.gov.uk/government/uploads/system/uploads/attachment_data/file/193945/bis-13-857-evaluating-changes-in-bank-lending-to-uk-smes-2001-12.pdf> accessed 23 January 2016.

¹⁷ *ibid.*

¹⁸ Department for Business Innovation and Skills, 2011, *The Government Response to the Independent Commission on Banking* (Cm 8252, 2011) 11

¹⁹ *ibid.*

²⁰ *ibid.* 11.

²¹ P Kinglsey, 'Financial crisis: Timeline' *The Guardian* (London, 7 August 2012)

<www.theguardian.com/business/2012/aug/07/credit-crunch-boom-bust-timeline> accessed 15 February 2016.

²² A Freeman, *Challenging Myths about the Funding of Small Businesses* (Demos Finance 2013) 77.

²³ *ibid.*

²⁴ *ibid.* 77.

Evidence from autumn 2008 indicates that SMEs were, ‘finding it increasingly difficult to obtain finance; seeing the withdrawal of promised finance; experienc[ing] sharp increases in loan interest rates; and (...) having facility fees imposed’.²⁵ In addition to this, the UK Survey of SME Finances (UKSMEF) showed that businesses’ credit ratings were poorer on average following the financial crisis.²⁶ This accounts for the reluctance of banks to lend to smaller businesses, which were more likely to have poor credit ratings and to lack collateral.

B. The ‘Funding Gap’

Banks’ reluctance to lend to SMEs has contributed to the notion of a ‘funding gap’ existing in the SME sector, identified by the 2009 Rowlands Review.²⁷ The term refers to, ‘the situation where a firm has a profitable opportunity but there are no or insufficient funds either from internal or external sources to exploit the opportunity’.²⁸

Although banks’ disinclination to lend to SMEs accounts for most of the funding gap, there are other contributory factors thereto. A significant amount of SMEs do not even attempt to obtain debt finance on the assumption that their requests will be rejected.²⁹ For example, the Small Business Survey carried out in 2015 stated that just under a third (31 per cent) of Scottish SMEs had applied for debt finance over the previous three years, a figure which, when compared with the 45 per cent of SMEs that had sought such finance in 2012,³⁰ suggests that SMEs’ confidence in their ability to obtain bank debt finance has been diminishing since the 2008 crisis. Increasing rejection rates have only confirmed SMEs in their fear that they will be refused a loan.

Another key factor in widening the funding gap is the lack of awareness amongst some SMEs of the types of finance that are most suitable and the potential availability of finance from alternative sources. This is partly a result of the expectations that dominated the pre-crisis period, when the exercise of bank lending was unregulated and limited advice was available on the issue.³¹ A lack of awareness of alternatives led SMEs to believe that bank debt financing, consistently readily available and easy to obtain, was the *only* suitable financing option. Evidence collected by the SME Finance Monitor shows that fewer than a fifth of the SMEs surveyed were aware of banks’ strategies to improve access to alternative providers of finance such as asset-based finance and Business Angel finance.³² The lack of educational advice and support in this matter may have posed a particular threat to

²⁵ National Federation of Self Employed & Small Businesses Limited (n 13).

²⁶ *ibid.*

²⁷ C Rowlands, *The Provision of Growth Capital to UK Small and Medium Sized Enterprises* (TSO 2009) 39.

²⁸ R Jarvis, *Finance and the Small Business in Enterprise and Small Business: Principles, Practice and Policy* (Prentice Hall 2000) 340.

²⁹ The Scottish Government, ‘The Market for SME finance in Scotland’ (The Scottish Government 2015) 13 <www.gov.scot/Resource/0048/00484379.pdf> accessed 15 February 2016.

³⁰ Office of the Chief Economic Adviser, ‘SME Access to Finance Report’ (The Scottish Government 2015) 5 <www.gov.scot/Resource/0047/00474061.pdf> accessed 21 February 2016.

³¹ The Royal Bank of Scotland (n 12) 12.

³² *ibid* 55.

SMEs, for which one of the key challenges is the identification of an appropriate type of finance; one which has the potential to kick-start or increase growth within the business.

Although, in recent years, one could observe a slight improvement in bank lending to SMEs – cumulative gross lending increased from £38bn to £50bn between 2012 and 2014,³³ and the ‘Funding for Lending’ governmental scheme has been implemented³⁴ – this trend should not be embellished. John Allan, chairman of the National Federation of Self Employed and Small Businesses (FSB), states that: ‘Although these are promising signs (...) there are still large numbers of businesses, especially sole traders and micro businesses, that are not aware of their finance options and as a consequence don’t get the finance they need.’³⁵

C. Late Payment of Debts

In addition to the funding gap, SMEs are adversely affected by the late payment of debts due to them, in particular those owed by their larger customers.³⁶ Larger customers tend to enjoy greater bargaining power, primarily due to the influence that they have on the potential growth of an SME as a purchaser of their goods. As a result, such customers can often delay payment of invoices with impunity, as well as exert pressure on SMEs to obtain higher quality products at lower prices. This constitutes a serious problem for SMEs as it hinders the cash flow of their businesses and prevents them from paying back their own suppliers, leading to bad business relationships.

Even before the financial crisis, in 1993 to be precise, the Forum of Private Business estimated that 89 per cent of SMEs were paid late by fifty-one days after the average due date.³⁷ The Late Payment of Commercial Debts (Interest) Act 1998 was subsequently passed to allow SMEs to claim interest upon late payments from other businesses.³⁸ The statutory interest rate is 8 per cent plus the Bank of England Base Rate, currently making it 8.5 per cent.³⁹

Although there is some evidence of an improvement to the ‘late payment’ problem within business,⁴⁰ research also suggests that the position remains unsatisfactory for many small firms.⁴¹ The average late payment burden that SMEs face now stands at £31,901, which puts many at risk of bankruptcy.⁴² This is even

³³ Bank of England, ‘Trends in Lending’ (Bank of England, January 2015)
<www.bankofengland.co.uk/publications/Documents/other/monetary/trendsjanuary15.pdf>
accessed 23 February 2016.

³⁴ Bank of England (n 4).

³⁵ E Anderson and J Titcomb, ‘Lending to Small Businesses Rises Sharply as Confidence Improves’
The Telegraph (London, 8 March 2015)

<www.telegraph.co.uk/finance/newsbysector/banksandfinance/11458051/Lending-to-small-businesses-rises-sharply-as-confidence-improves.html> accessed 23 February 2016.

³⁶ Edmonds (n 3) 4.

³⁷ *ibid* 5.

³⁸ Late Payment of Commercial Debts (Interest) Act 1998.

³⁹ *ibid*.

⁴⁰ Edmonds (n 3) 15.

⁴¹ *ibid*.

⁴² *ibid*.

more worrying given the statement by 25 per cent of SMEs that a late payment burden of £20,000 or less is enough to imperil their business prospects.⁴³

One could, therefore, question how successful the 1998 Act has been in combatting the problem of late payments. One potential difficulty is that SMEs want to avoid jeopardising their trading relationships with the customers on whom they rely for survival. Amir Khanlari accentuates this point, observing that no matter how satisfactory a business' products are, that business will not be guaranteed continued customer loyalty if it has provided a poor service experience, a key example of which being the imposition of charges and interest.

In this troubling financial climate, characterised by a decline in bank lending and tightened lending criteria, along with the continuous problem of late payment of debts, the growth of SMEs will inevitably be constrained. As SMEs have, hitherto, tended to rely on bank lending, it is presumed that this would be their preferred means of reinstating their cash flow when faced with late payments. However, bearing in mind the increasing strictness of lending criteria and of the covenants imposed on higher risk businesses, in particular SMEs, the latter may be placed in an even more vulnerable position if they continue to borrow from banks.

3. A Potential Solution: The Law on Invoice Finance

A. What Is Invoice Finance?

Ensuring a healthy cash flow is the first and foremost priority of an SME. It is both a necessary ingredient of the growth of a business and an important measure of stability.⁴⁴ It is submitted that the best way to create such a cash flow is via invoices. With the biggest source of cash remaining in the hands of their customers, it appears that SMEs' invoices will often be the most valuable assets within their businesses.⁴⁵

As an alternative to bank lending, asset-based finance covers an array of devices that can be used to release locked up cash, such as plant and machinery, property and receivables.⁴⁶ For the purposes of this written work, the focus will be on releasing cash tied up in receivables, i.e. invoice financing. The benefits of invoice financing have the potential to overcome the problems that typically confront SMEs – by combatting restricted access to finance, providing credit expertise and obviating the risk of loss in profits where invoices are paid late. Moreover, the nature of SMEs' payment procedure qualifies for the invoice financing service; generally speaking, these businesses issue invoices to their customers along with credit terms, for example '30 days to pay this invoice'.⁴⁷ Indeed, SMEs' growth is built upon their

⁴³ *ibid.*

⁴⁴ B O'Halloran, 'Bridging the Gap' (2002) 93 GLSI 26, 26.

⁴⁵ *ibid.*

⁴⁶ Asset Based Finance Association, 'Asset Based Finance Industry Information' (Asset Based Finance Association 2017) <www.abfa.org.uk/About/industryInformation.asp> accessed 24 February 2017.

⁴⁷ UK Government 'Invoicing and Taking Payment from Customers' (UK Government 2017) <www.gov.uk/invoicing-and-taking-payment-from-customers/payment-obligations> accessed 24 February 2017.

trading relationships with other businesses where an invoice is issued immediately after the transfer of goods.

Invoice financing can be thought of as 'receivables financing', which refers to the transactions whereby a business raises money against the debts due to them.⁴⁸ According to Eva Lomnicka, legally this can be exerted through an 'outright sale of the receivables at a discount' or a 'loan secured on the receivables'.⁴⁹ Although both arrangements have the same economic outcome, in that the business is provided with immediate available finance, legally the two transactions are different.⁵⁰ However, 'outright sale of the receivables at a discount' is the method most associated with receivables financing, as it has been established that this transaction is seen to be carried out more as a sale than a security transaction.⁵¹ Hereinafter, a business selling its invoices to a factor can be understood as the 'client' and the term 'debtor' can be taken to refer to the customer of the client who owes the money.

The method of invoice financing is twofold: it can take the form of either 'factoring' or 'invoice discounting'. *Salinger on Factoring*, the leading text on this area of law, defines it as: 'The purchase of debts (other than [personal] debts (...) and debts payable on long terms or by instalments) for the purpose of providing finance, or relieving the seller from administrative tasks or from bad debts, or for any or all of such purposes'.⁵²

It is imperative at this point to elaborate on this definition by stating that, in the context of factoring, the debts are purchased at a discount reflecting the service that the factor has provided to the client.⁵³ The client can receive up to 90 per cent of the discounted debts sold upfront, the remaining amount only being given to the client once the debtor has fully repaid their debts.⁵⁴ In addition to this, the factor provides the function of managing the sales ledger and collecting the money on top of providing immediate finance to the business for a fee deducted from the outstanding amount due to the client.⁵⁵ This method requires that the debtor be notified of the agreement between the factor and client.⁵⁶

In essence, 'invoice discounting' is achieved in the same manner. The difference between this and factoring is that, in the former case, notification is not given to the debtor; also 'invoice discounting' does not entail the functions of managing the business' sales ledger and collecting the debts.⁵⁷ Hence, one basis on which to distinguish factoring and invoice discounting is the presence of 'notification' in one case (factoring) and the absence of notification in the other (invoice financing).

⁴⁸ E Lomnicka, 'Financing Devices Involving the Transfer or Retention of Title' in H Beale and others (eds), *The Law of Security and Title-Based Financing* (2nd edn, OUP 2012) 273.

⁴⁹ *ibid.*

⁵⁰ *ibid.*

⁵¹ *ibid.*

⁵² N Ruddy and others, *Salinger on Factoring* (4th edn, Sweet & Maxwell 2006) 2.

⁵³ *ibid* 287.

⁵⁴ UK Government, 'Business Finance Explained' (UK Government) <www.gov.uk/business-finance-explained/invoice-financing> accessed 28 February 2016.

⁵⁵ *ibid.*

⁵⁶ Ruddy and others (n 52) 290.

⁵⁷ *ibid* 51.

Further variants of factoring are factoring ‘with recourse’ and factoring ‘without recourse’ to the client, the former giving the factor the right to shift the risk of non-payment of any debts to the client.⁵⁸ Conversely, factoring ‘without recourse’ is the situation where the factor assumes and takes responsibility for the credit risk of bad debts.⁵⁹ Invoice discounting (in which notification is not given to the debtor) is naturally carried out in conjunction with recourse to the client. Therefore, invoice discounting can be classed as an exclusively financial facility unaccompanied by the additional non-financial elements that factoring entails.

The differences between factoring and invoice discounting are best illustrated with an example. Suppose that a client is owed £40,000 by a debtor. A third party (either a factor or invoice-discounter) purchases the debts at a discounted rate of 3 per cent and the client initially receives only 90 per cent of the debt. For the sake of simplicity, it will be assumed, for the purposes of this example, that the third party does not charge any additional administration fee. Here the client will initially receive £36,000 (90 per cent) and a further £2,800 (7 per cent) once the debt has been paid. If the third party is a factor, he or she will collect the debts and pay the final £2,800 to the client once the debtor has fully paid their outstanding debts. In the case of invoice discounting, however, the client will collect the debts, as they continue to bear the credit risk, and give the full amount of the debt to the factor once the debtor has paid the invoice.’

In its infancy, the use of invoice finance was perceived as a weakness, with academics dismissing it as a last resort for meeting a client’s financing needs.⁶⁰ Woodley elaborates on this criticism with the comment that banks only offered invoice finance once the possibilities of overdraft lending had been exhausted.⁶¹ However, according to Caouette, this financing technique has attracted an increasing amount of approval and respectability in recent years. Due to a reduction in its cost, many financial professionals regard it as a perfectly acceptable financing alternative.⁶² Given this increased acceptance within the financial industry, the evident advice gap must be closed in order effectively to promote the benefits of invoice financing to SMEs.

B. Invoice Financing as a Transaction by Way of Sale

It was submitted earlier that the transaction of invoice financing is seen more as a sale than a secured loan. Whilst that understanding has sometimes been called into question,⁶³ the courts have not yet been persuaded to depart therefrom. In both *Chow*

⁵⁸ *ibid* 18.

⁵⁹ *ibid*.

⁶⁰ Simply Factoring Broker, ‘Factoring: Last Resort or Long Term Support?’ (Simply Factoring Broker 2014) <www.simplyfactoringbroker.co.uk/blog/posts/2014/factoring-last-resort-product-or-long-term-support/> accessed 2 March 2016.

⁶¹ I Woodley, ‘We Live in Strange Times’ *Recovery Magazine* (2014) 25

<www.r3.org.uk/media/documents/publications/recovery/Autumn_2014.pdf> accessed 2 March 2016.

⁶² JB Caouette and others, *Managing Credit Risk: The Great Challenge for the Global Financial Markets* (2nd edn, John Wiley & Sons Inc 2008) 122.

⁶³ Lomnicka (n 48) 287.

*Yoong Hong v Choong Fah Rubber Manufactory*⁶⁴ and *Lloyds & Scottish Finance Ltd v Cyril Lord Carpet Sales Ltd*,⁶⁵ it was accepted that the 'purchase price under a contract of sale may be discounted, reflecting early payment, without the agreement being [viewed] as a disguised loan'.⁶⁶ These cases persistently distinguished a discount charge from a payment of interest on the basis that the price paid at a discount is fixed and paid at once whereas interest accumulates from day to day.⁶⁷ From that point, the notion that receivables financing was a transaction by way of sale was established. On this view, the client is passing on 'ownership' of his or her debts to the factor.⁶⁸

The advantages of the transaction being perceived as a sale are, firstly, in the case of factoring, that such a transaction allows the factor to obtain the debts directly from the debtor.⁶⁹ Furthermore, if understood as sale transactions, factoring agreements do not require to be registered in the companies' charges register.⁷⁰ This is advantageous for clients, who will generally wish to minimise the number of security interests affecting their property.⁷¹ In contrast, this 'sale' characterisation does not operate to the advantage of the factor. If that party is seen to have a registered charge over the receivables, he will be placed in a stronger position regarding disputes as to priority.⁷² This is because a non-assignment clause (discussed below in section D) which prohibits the *sale* of debts, does not necessarily apply to prohibit *charges* created over debts.⁷³ With regard to the latter, the factor is allowed to have a security interest in the debt, thereby allowing him or her to acquire direct rights against the debtor from the client.⁷⁴

Another advantage of the transaction being by way of sale for the client is that, generally, sales are 'off balance sheet', meaning that the invoices are not required to be reported in that way.⁷⁵ This in turn gives the impression of a stronger balance sheet, thereby increasing the client's borrowing capacity from a bank, which may result in more credit, better rates or more lenient loan covenants.

In sum, the advantages of a sale transaction are significantly beneficial, at least for the client.

C. Assignment of Debts

This section explains the law behind the transfer of the debts to the factor. In the context of English law, Ruddy defines debts as belonging to a class of property

⁶⁴ [1962] AC 209 (PC) 217.

⁶⁵ [1992] BCLC 609 (HL).

⁶⁶ Lomnicka (n 48) 287.

⁶⁷ *ibid* 288.

⁶⁸ Ruddy and others (n 52) 147.

⁶⁹ Lomnicka (n 48) 288.

⁷⁰ *ibid*.

⁷¹ *ibid*.

⁷² *ibid* 289.

⁷³ *ibid* (emphasis added).

⁷⁴ *ibid* 276.

⁷⁵ Lomnicka (n 48) 289.

known as ‘choses in action’,⁷⁶ which is defined in *Torkington v Maggee* as personal rights of property which can only be claimed or enforced by action – and not by taking physical possession.⁷⁷ Lomnicka explains that because they are ‘choses in action’, ‘any dealing in them is effected by way of assignment’.⁷⁸ The assignment of the receivables can either be an equitable or statutory assignment, the latter of which can also be referred to as a ‘legal assignment’ under the Law of Property Act 1925.⁷⁹ Regardless of which assignment is given, both have the effect of transferring absolute ‘ownership’ of the debts.⁸⁰

For a statutory assignment to take effect, the assignment must be written by the client of the factor and ‘express notice in writing’ must be given to the debtor.⁸¹ Fulfilment of these requirements transfers to the factor a ‘legal right to the debt, all legal and other remedies for the debt and the power to give a good discharge for the debt without the concurrence of the client’.⁸²

On the other hand, an equitable assignment need not be in writing nor is giving notice to the debtor necessary in order to transfer ‘ownership’.⁸³ The only difference between the two assignments is that the assignee can sue in his or her own name for recovery from the debtor on the basis that it is a statutory assignment, whereas within an equitable assignment, the factor can only enforce an action jointly with the client.⁸⁴

Although giving notice is not obligatory to transfer ‘ownership’ within an equitable assignment, a statutory assignment should be considered as an option by the factor due to its intrinsic advantage of conferring priority on him or her over other factors. This principle was established in *Dearle v Hall*.⁸⁵ Ultimately, as the advantages of giving notice have been established, it is evident that a statutory assignment may be preferable for the factor in terms of affording him or her priority.

D. Non-Assignment Clauses

A successful assignment of the debts does not simply consist of concluding an agreement between the client and the factor. In this regard, a hurdle to the assignment might exist in the form of a ‘non-assignment clause’,⁸⁶ which prohibits or restricts both a statutory and equitable assignment of debts, precluding a factor from acquiring any direct rights against the debtor from the client.⁸⁷ There are a few reasons to include this prohibition for the benefit of the debtor but fundamentally, it is due to the fact that the identity of the debtor’s counterparty is crucial in terms of

⁷⁶ Ruddy and others (n 52) 137.

⁷⁷ [1902] 2 KB 427 [430] (Channell JJ).

⁷⁸ Lomnicka (n 48) 274.

⁷⁹ Law of Property Act 1925, s 136.

⁸⁰ Ruddy and others (n 52) 140.

⁸¹ Law of Property Act (n 79) s 136.

⁸² Ruddy and others (n 52) 139.

⁸³ *ibid.*

⁸⁴ Lomnicka (n 48) 274.

⁸⁵ (1828) 3 Russ 1.

⁸⁶ Lomnicka (n 48) 276.

⁸⁷ G McCormack, ‘Debts and Non-Assignment Clauses’ [2000] JBL 422, 427.

performance and payment.⁸⁸ This relates, at least partly, to expediency. It is more convenient for the debtor to approach the original creditor (the client), who will be familiar to the debtor by virtue of the trading relationship between them. Secondly, the debtor may choose to include a non-assignment clause in relation to his or her set-off rights that would cease to be enforceable against the creditor (the client) had an assignment taken place.⁸⁹

A breach of the assignment clause renders the transfer of rights to the factor ineffective, as was established in the case of *Linden Gardens Trust Ltd v Lenesta Sludge Disposals Ltd*.⁹⁰ However, an obiter dicta comment made by Lord Browne-Wilkinson, expressed that the non-assignment clause does not 'operate to invalidate the contract between the assignor (client) and assignee (factor)'.⁹¹ According to general contractual principles, the contract is valid because the debtor cannot interfere with the contractual obligations between the factor and the client;⁹² as such the factor can still raise an action against the client for breach of contract.⁹³

Ultimately, this contractual clause dilutes the efficacy of invoice financing, impairing the client's freedom to free up cash tied up in invoices. This acts as a barrier in the UK to the use of invoice financing where it has the potential to combat late payment of debts and evade restrictive terms in bank lending. However, as discussed below, there is a way to circumnavigate this contractual obstacle.

E. Scots Law on the Transfer of Ownership of Debts

A concern may arise as to the jurisdiction governing a contract where a company is incorporated in England but the business is carried out in Scotland, resulting in Scottish debtors.⁹⁴ This is because the Scots law on the transfer of 'ownership' of debts is stricter than its English counterpart, particularly as regards the requirement of 'intimation'.⁹⁵

In Scots law, a successful transfer entails a three-stage process. The first stage requires the client to agree to sell (and the factor to agree to buy) the debt. Secondly, an assignation ('assignment' in English law) of the debt must be made by the client to the factor in which actual words of conveyance or words of transfer must be given (in practice, this is done by executing a deed of assignation and delivering it to the assignee).⁹⁶ Thirdly, 'intimation' (notification) to the debtor of the transfer is essential in Scots law.⁹⁷ Assignation confers upon the factor a mere personal right against the client to have the debts transferred; in other words, assignation does not, in itself,

⁸⁸ Lomnicka (n 48) 276.

⁸⁹ *ibid.*

⁹⁰ [1994] 1 AC 85 [108] (Lord Browne-Wilkinson).

⁹¹ *ibid.*

⁹² Lomnicka (n 48) 278.

⁹³ *ibid.*

⁹⁴ Ruddy and others (n 52) 151.

⁹⁵ *ibid.* 153.

⁹⁶ *ibid.*

⁹⁷ *ibid.*

transfer ‘ownership’ of the debts to the factor.⁹⁸ For that to occur, ‘intimation’ is necessary.⁹⁹

The particular stage of intimation attracts significant attention in relation to the stricter rules arising under Scots law. This is primarily due to the fact that intimation determines whether the factor has ‘ownership’ or not, whereas in English law a transfer of ‘ownership’ can occur regardless of whether notice is given or not.¹⁰⁰ With that being said, equitable doctrines relating to the assignation of debts are not recognised in Scots law in the same way that they are in Scots law.¹⁰¹ *Alexander Grigor Allan and Others v Robert Urquhart and Others*¹⁰² accentuates that it is ‘intimation’ that gives the assignee the real right; thus the factor’s position remains weak between the stages of assignation and intimation as they only have a personal right against the assignor.

The strict rules relating to the transfer of debts in Scots law are subject to some uncertainty due to an absence of clear authority concerning intimation. Doubts have been cast upon cases dealing with this issue,¹⁰³ but, as Ruddy explains, textbooks in previous years have stipulated that effective intimation is ‘given in the form of a notice in duplicate, the duplicate copy of which has been acknowledged by the debtor and returned to the assignee or, [alternatively], an old style notarial declaration’.¹⁰⁴ However, in the factoring case of *Liberatas-Kommerz GmbH*, there was an indication that a court may be prepared to allow an informal intimation where that court only sought to ensure that the debtor had been adequately made aware of the transfer, and, if he or she had, that was sufficient to amount to intimation.¹⁰⁵ However, Anderson criticises this case as weak authority due to the test established for determining whether intimation has taken place.¹⁰⁶ He purports that the test fails to fulfil the first function of an intimation which he stipulates as ‘a provision of a certain date of transfer’.¹⁰⁷

The second function, namely to ‘interpel’ the debtor (i.e. to inform him or her that payment must now be made to the assignee’) was the only function that the test addressed.¹⁰⁸ Anderson makes reference to postal delivery, as authorised under section 2 of the Transmission of Moveable Property (Scotland) Act 1862 as a plausible means of providing a certain date of transfer.¹⁰⁹ Under this rule, the assignee sends a certified copy of assignation by post and obtains a written acknowledgment of receipt from the debtor as evidence of intimation having been made.¹¹⁰

⁹⁸ *ibid.*

⁹⁹ *ibid.*

¹⁰⁰ *ibid.*

¹⁰¹ *ibid.*

¹⁰² (1887) 15 R 56.

¹⁰³ *Liberatas-Kommerz GmbH v Johnson* 1978 SLT 222.

¹⁰⁴ Ruddy and others (n 52) 155.

¹⁰⁵ *Liberatas-Kommerz GmbH* (n103) [226] (Lord Kincaig).

¹⁰⁶ R Anderson, ‘Intimation: 1862-2008’ (2008) 12 Edin LR 275, 275-280.

¹⁰⁷ *ibid.*

¹⁰⁸ *ibid.*

¹⁰⁹ Transmission of Moveable Property (Scotland) Act 1862, s 2.

¹¹⁰ *ibid.*

However, a difficulty may arise where the notice or acknowledgment becomes lost in the post. If the client becomes insolvent, it is questionable what the position of the factor will be where an acknowledgment has been sent of the intimation but the factor has failed to receive it due to it becoming lost in the post. The reliance on postage to perfect intimation essentially detracts from the notion that the assignee has control over his or her impending 'ownership' of the debts. In addition to this, where the client has assigned several debts to one factor or various factors, the time, nuisance and cost inherent in notifying the numerous debtors will be onerous for the factor.

A further concern is that the requirement of intimation for a transfer of ownership undermines the use of invoice discounting in Scotland, where notice is not part of the arrangement. Therefore, it may be troublesome when determining whether the factor has 'ownership' of the debts where notice cannot be given within the agreement between the client and the factor. Promoting debt finance alternatives (in particular invoice finance) is crucial in the wake of the recession, but these efforts are inevitably hindered by the rule of 'intimation' in Scotland. Therefore, it seems that businesses in Scotland are restricted to factoring agreements alone, effectively preventing businesses from retaining confidentiality with regard to the sales of the debts.

In sum, the requirement of intimation appears to be burdensome for the factor. He or she cannot even rely on post to transfer the intimation, as there is no guarantee that the notice will safely arrive in the hands of the debtor. It would appear to be much easier to obtain 'ownership' of debts under English law, especially under an equitable assignment where the process of transferring 'ownership' is faster and uncomplicated. In Scots law, the otherwise simple mechanics of invoice financing appear to have been beset by complexity.

F. Law of Trusts

Given the above, a question arises as to how the factor is to protect himself or herself when he or she lacks title, whilst waiting for the impending intimation to be perfected by the debtor acknowledging the notice. More so, the question needs to be answered whether there is a way to implement invoice discounting in Scotland when notice is not given. Ruddy professes that there is one successful method used to circumvent the issue of intimation, which involves the use of a trust.¹¹¹ The famous case of *Tay Valley Joinery v CF Financial Services* approved this principle, mitigating the difficult position in which factors had inevitably been placed.¹¹² The case established that debts may become the subject matter of a trust without the owner divesting himself or herself thereof by means of an intimated assignment.¹¹³

The process operates as follows: once a debt has been sold by the client to the factor and after assignment has been performed, while the factor only holds a personal right vis-à-vis the client to obtain the debts, the client can hold the title to

¹¹¹ Ruddy and others (n 52) 108.

¹¹² 1987 SLT 207 [215] (Lord Dunpark).

¹¹³ *ibid.*

the debt in trust for the factor.¹¹⁴ In practice, an SME declares a trust over its own property (the invoices) in favour of the factor (the beneficiary) making the SME both the trustor and trustee. Hence, like all trustees, they are subject to fiduciary duties, meaning that they must fulfil the purposes of the trust and carry out their powers in a manner that prioritises the interests of the beneficiaries.¹¹⁵ This protects the factor of the property, who has paid for, but not yet received, title. Although there were uncertainties as to the legality of a declaration of oneself as a trustee of one's own property, this was accepted and established in the case of *Allans Trustees v Lord Advocate*.¹¹⁶

The beauty of declaring oneself a trustee is that there is no need to intimate the formation and declaration of the trust to the debtors.¹¹⁷ It makes invoice discounting a practicable, viable option in Scotland. Essentially, the debtor continues to pay their debts to the client. This preserves the attractive aspect of invoice discounting whereby there is no need to give notice to the debtor of the transfer of debts, therefore retaining confidentiality. The debts that he or she holds in trust for the factor will not be attachable by the client's creditors when undergoing insolvency, thereby protecting the factor against the liquidator, administrator or the trustee of the factor's client.¹¹⁸

It is necessary to note at this point that the law of trusts also operates as an invaluable device in relation to non-assignment clauses in English law. The case of *Re Turcan* held that a non-assignment clause does not preclude the client from holding the debts on trust for the factor.¹¹⁹ Lomnicka describes the creation of a trust as granting a proprietary right to the factor in addition to the right of breach of contract that he or she already has against the client.¹²⁰ This proprietary right essentially gives the factor priority should the client become insolvent, producing a similar effect to the Scots law of trusts.

However, a disadvantage to a debtor exists wherever a factor, as a beneficiary under a trust, can sue the debtor directly, in cases where the client does not take steps to enforce performance or bring an action against the debtor.¹²¹ In such cases, the purpose of the non-assignment clause would be undermined. Although a 'limited trust' was proposed by Lightman J as a solution to this problem, whereby rights are granted to the beneficiary (factor) 'vis-à-vis the trustee' but where no rights are conferred on the beneficiary (factor) against the debtor,¹²² the courts were less eager to accept this as a valid principle in the case of *Barbados Trust Co Ltd v Bank of Zambia*.¹²³

In conclusion, although the process of assigning debts comes with problems such as non-assignment clauses, it does not render the factor entirely without rights

¹¹⁴ Ruddy and others (n 52) 155.

¹¹⁵ *Cowan v Scargill* [1985] Ch 270.

¹¹⁶ 1971 SC (HL) 45.

¹¹⁷ *ibid.*

¹¹⁸ Ruddy and others (n 52) 156.

¹¹⁹ (1889) LR 40 Ch D 5.

¹²⁰ Lomnicka (n 48) 278.

¹²¹ *Vandepitte v Preferred Accident Insurance Corporation of New York* [1933] AC 70 (PC).

¹²² *Don King Productions Inc v Warren and Others* [2000] Ch 291.

¹²³ [2007] EWCA Civ 148, [2007] 2 All ER (Comm) 445.

enforceable under the contract. The problem can be addressed via the device of a trust, something which is undeniably beneficial to factors in Scots law who, after assignation but before intimation, hold no title over the debts but have only a personal right thereto against the client. The use of trusts within an assignation (or assignment) of the debts essentially results in three different laws involved in the agreement: property, contract and trusts. This is arguably cumbersome for the parties to the agreement and may negatively influence their decision to use invoice financing. However, so long as legal advice is sought on the matter, the simple process of invoice financing and its benefits should be unequivocally appealing to SMEs.

4. Invoice Financing versus Bank Lending

This article suggests that, while bank lending has its advantages, certain aspects thereof can pose serious risks to SMEs. This section of the article aims to provide an analysis of the supposed benefits of invoice finance. It will explore whether this mode of financing is an intrinsically good alternative to various forms of bank lending and whether the traditionally negative perceptions of this finance technique should be re-assessed.

It was submitted earlier in the paper that SMEs rely on bank lending to obtain working capital with more than half of SMEs in the past three years seeking finance through their main bank.¹²⁴ This is primarily due to SMEs' assumption that bank lending is the most suitable option thanks to its flexibility and the exceedingly easy access to credit that it provided prior to the financial crisis.¹²⁵ Furthermore, research collected by the Forum of Private Business shows that many SMEs remain apprehensive with regard to alternative lending; they express uncertainty as to the long-term feasibility of such options, with one in four stating that they were 'a temporary phenomenon due to current economic conditions'.¹²⁶ In addition, only 1 per cent of SMEs use invoice finance, which highly contrasts with the third of SMEs that use overdrafts and term loans.¹²⁷

Fundamentally, this has contributed to the emergence of an advice gap as regards the financial alternatives available to SMEs; indeed, fewer than 12 per cent were aware of invoice financing.¹²⁸ Essentially, this means that a large percentage of SMEs lack awareness of other suitable options for their business.

The next sections focus on bank lending. Since overdrafts and term loans are the most commonly used facilities by SMEs, where reference is made to 'bank lending' below, consideration will be paid to both.

¹²⁴ BMG Research, 'Small and Medium-sized Enterprise (SME) Journey towards Raising External Finance' (Department for Business Innovation and Skills 2013) 7 <<http://british-business-bank.co.uk/wp-content/uploads/2013/10/SME-Journey-Towards-Raising-Finance.pdf>> accessed 12 March 2016.

¹²⁵ The Royal Bank of Scotland (n 12) 12.

¹²⁶ BMG Research (n 124) 11.

¹²⁷ *ibid.*

¹²⁸ *ibid.*

A. Overdrafts

An overdraft is a form of debt financing intended generally to cover short-term loans.¹²⁹ It is operated through a customer's current account and arises when the business exceeds its available balance.¹³⁰ An overdraft is particularly attractive due to the attendant flexibility of borrowing the exact amount needed at a specific time and it is relatively quick to arrange.¹³¹ This is advantageous for SMEs which operate in a fast-paced environment, where an injection of capital is usually needed immediately to meet daily obligations. The limit of the amount that can be withdrawn is usually agreed beforehand, for which facility (known as an authorised overdraft) the business may be required to pay a commitment fee.¹³² Overdrafts can also be unauthorised, but a higher rate of interest will be imposed on such borrowing and further charges will be incurred by the business; hence it is of benefit to the SME to undertake an authorised facility.¹³³

On the other hand, invoice financing is linked to the client's sales ledger, not the balance sheet, and so the facility is allowed to grow with the business.¹³⁴ Essentially, no debt arises as it is purely a transaction of advancing cash against the unpaid debt, with no consequent increase to balance sheet liabilities. An overdraft, by contrast, is written off as an expense of the business. The cost could even escalate higher than planned if the business overdraws above the overdraft limit or fails to abide by the repayment terms.

Ferran purports that there is a requirement attached to the facility for the debts to be 'repayable on demand', which refers to a loan made where no time repayment is specified or where it is expressly stated to be repayable 'on demand'.¹³⁵ In addition to this, even where a set period has been agreed for the overdraft, the bank can still demand repayment before this time has elapsed.¹³⁶

Arguably, a 'repay on demand' requirement would be ill-suited to the unexpected, often unavoidable, deficits to which SMEs are vulnerable. Where an SME is required to repay on demand, it may not have sufficient funds to repay the money borrowed via the overdraft facility. This is what occurred in *Cripps (Pharmaceutical) Ltd v Wickenden*, where a bank was held still to be within its rights when it appointed a receiver to enforce its rights less than two hours after it had demanded repayment, even when it was evident that the business did not have the capacity to pay at that precise time.¹³⁷ Similarly, a delay of just one hour between

¹²⁹ UK Government (n 54).

¹³⁰ E Ferran and LC Ho, *Principles of Corporate Finance Law* (2nd edn, OUP 2014) 271.

¹³¹ UK Government (n 54).

¹³² Ferran and Ho (n 130) 271.

¹³³ ACCA Global, 'Types of Finance: Overdraft' (ACCA Global 2015)

<www.accaglobal.com/uk/en/business-finance/types-finance/overdraft.html> accessed 3 March 2016.

¹³⁴ Touch Financial, 'Factoring vs. Bank Overdraft' (Accounting Web 2012)

<www.accountingweb.co.uk/blog-post/factoring-vs-bank-overdraft> accessed 25 February 2016.

¹³⁵ Ferran and Ho (n 130) 271.

¹³⁶ TSB, 'Overdrafts' (TSB) <www.tsb.co.uk/current-accounts/faqs/overdrafts/> accessed 5 March 2016.

¹³⁷ [1973] 1 WLR 944.

making the demand and sending in the receivers was held to be justified in the case of *Bank of Baroda v Panessar*.¹³⁸

As the two aforementioned cases make clear, reliance on an overdraft facility could severely prejudice an SME. Where the SME has insufficient funds to repay the overdraft on demand and the receiver is appointed to sell the business' assets to repay the debts to the appointer (bank),¹³⁹ it can be postulated that the business is facing imminent insolvency.

This can be contrasted with a factoring agreement, wherein a proper notice period is required by both sides, reassuring the borrower that the supply of funds will not be cut off at a crucial time during the operation of the business.¹⁴⁰

B. Term Loans

Another lending facility used by banks is the term loan. Here funds are granted to the borrower for a specified period, requiring repayment either at the end of that period,¹⁴¹ or during the period in accordance with the agreed repayment schedule.¹⁴² The primary use of term loans differs from that of an overdraft where the loan is used for start-ups, to fund a project or to acquire assets, usually equipment for the business.¹⁴³ Hence, this form of financing is suitable for the long term.

An interest charge is also levied in addition to the obligation to repay the principal sum.¹⁴⁴ The interest charge is either fixed for the term or variable through LIBOR, the London Inter-Bank Offered Rate, a benchmark giving an indication of the average rate at which a LIBOR contributor bank can obtain unsecured funding in the London inter-bank market for a given period, in a given currency.¹⁴⁵

With regard to repayment of the loan, there are three ways in which the repayment can be executed. An agreement may stipulate that the whole loan is to be repaid at one time, amounting to a 'bullet repayment'.¹⁴⁶ An 'amortized repayment' comprises repayment in instalments over a period of time and, where the final payment is the largest instalment of the repayments made over the period, this is known as a 'balloon repayment'.¹⁴⁷ Additionally, where more money is required, instead of applying for a fresh loan, the business can acquire the funds through a revolving credit facility.¹⁴⁸

¹³⁸ [1986] 3 All ER 751.

¹³⁹ Practical Law, 'Chapter 8 – Secured Creditors' (Practical Law 2009)

<<http://uk.practicallaw.com/book/9781847661487/chapter08>> accessed 3 March 2016.

¹⁴⁰ Ruddy and others (n 52) 262.

¹⁴¹ Ferran and Ho (n 130) 274.

¹⁴² *Danske Bank v Durkan New Homes* [2009] IEHC 278 [22].

¹⁴³ Ferran and Ho (n 130) 274.

¹⁴⁴ *ibid* 277.

¹⁴⁵ *ibid*.

¹⁴⁶ *ibid* 275.

¹⁴⁷ *ibid*.

¹⁴⁸ *ibid* 275.

C. Conditions Precedent

Naturally, a bank's priority in lending money is to ensure that the business is able to repay the loan. As such, prior to lending money to a business, banks will carry out credit assessments. The key factors shaping the bank's decision are: Character, Ability, Margin, Purpose, Amount, Repayment and Insurance (CAMPARI).¹⁴⁹ The process by which CAMPARI operates essentially allows the bank to obtain certain constitutional and financial documents which form part of the 'conditions precedent' stage that must be satisfied before the business can draw down funds.¹⁵⁰

Where the business' operations and financial documents are heavily scrutinised at such an early stage, this constitutes a significant hurdle to borrowing. This assessment stage may harm the prospects of a business seeking to start up because such a business will lack certain attributes required under the assessment.

For example, where the banks look to the 'Ability' element of making a successful business, essentially they are enquiring as to whether the business owner has the relevant sector experience or the skills adequately to control and expand the business.¹⁵¹ The business owner who is starting up may not have previous experience, which may hinder the bank's decision whether to lend.

The element of 'Amount' concerns how much the business is seeking to obtain and how much the business itself is willing to contribute.¹⁵² Where a business is a start-up, the owner may have insufficient or non-existent personal savings, hence relying wholly on a bank loan.

The element of 'Repayment' may be the most difficult requirement to fulfil for both start-ups and established SMEs, as, unsurprisingly, the borrower's ability to repay the money lies at the very heart of the bank's assessment process.¹⁵³ Here, the bank carefully examines the business' annual accounts, management figures and financial projections, thereby scrutinising the business' credit history.¹⁵⁴ Where, as will often be the case with a start-up, the relevant documents display slow progress in its cash flow statement or the business has experienced a significant loss discernible from its profit and loss account, a problem naturally not confined to start-ups, this may act as a warning sign to the bank, suggesting that the business in question is a high-risk borrower. In essence, there is often a vicious circle: bank loans are undeniably important to the survival and growth of businesses; at the same time, a business will struggle to obtain such credit if it cannot produce evidence, based on past borrowing, of its ability to repay the sums borrowed.

The last stage, known as 'Insurance', looks to the form of security that the business can offer.¹⁵⁵ If the sum sought is a relatively small amount, the bank will lend against the business' credit history and impose covenants within the loan to

¹⁴⁹ SME Web, 'How a Bank Will Assess your Business Finance Request' (SME Web)

<www.smeweb.com/finance/features/4124-how-a-bank-will-assess-your-business-finance-request> accessed 29 February 2016.

¹⁵⁰ *ibid.*

¹⁵¹ *ibid.*

¹⁵² *ibid.*

¹⁵³ *ibid.*

¹⁵⁴ *ibid.*

¹⁵⁵ *ibid.*

ensure that the business retains its ability to repay the loan; even so, in such cases, the bank does not usually require a security.¹⁵⁶ However, for larger amounts, the bank will offer a secured loan whereby it will impose either fixed or floating charges upon the assets of the business and use those assets to satisfy the loan in the event of default.¹⁵⁷ Additionally, where the start-up has no or insufficient collateral to serve as security, the bank may agree to an unsecured loan where onerous covenants can hinder the growth of the start-up.

With invoice financing, the factor also examines the business' operations in terms of obtaining status reports, audited accounts and enquiring about the business owner's past experience within the sector.¹⁵⁸ However, in this case, the focus is shifted entirely on to the debtor in terms of their ability to repay the debts as opposed to bank lending where the focus is primarily on the business.¹⁵⁹ Herein lies a disadvantage of this financing technique. Where the majority of the approval is based upon the credit strength of the debtor, a poor 'debt turn' (i.e. the average period of credit taken by debtors)¹⁶⁰ can negatively influence the factor's decision as to whether to enter into the factoring agreement.

However, a factoring agreement does not require security in terms of corporeal moveable or heritable property.¹⁶¹ This renders the business free from the interference of a bank potentially seizing its assets and enables the business to use those assets, for example its plant and machinery, to free up additional cash under the other devices of asset-based financing. Ultimately, then, with regard to the business' assets, SMEs are more protected under invoice financing than they are under bank lending.

Another element of meeting the 'conditions precedent' is the representations and warranties produced by the business.¹⁶² These are statements of facts, (accompanied by promises attesting to the accuracy of those statements) which are subsequently included in the agreement to serve as a basis for protection of the bank throughout the life of the loan.¹⁶³ As Ferran purports, the onus is on the business to repeat the representations and warranties at specified intervals.¹⁶⁴

Clearly, bank lending is an area that still requires the business' involvement even after the loan has been granted, in order for them to be actively aware of changes, big or small, and to have them disclosed promptly to the bank. This is time-consuming and constitutes an additional burden on the owner, above and beyond the running of the business, as an inaccurate representation, even one made in good faith, or failure to notify the bank of the smallest change will entitle the lender to call

¹⁵⁶ *ibid.*

¹⁵⁷ *ibid.*

¹⁵⁸ Ruddy and others (n 52) 111.

¹⁵⁹ *ibid* 110.

¹⁶⁰ *ibid* 117.

¹⁶¹ S Angeles, 'Choosing a Factoring Service for Your Business' (*Business News Daily* 2016) <www.businessnewsdaily.com/7081-factoring-buyers-guide.html> accessed 4 March 2016.

¹⁶² Ferran and Ho (n 130) 281.

¹⁶³ *ibid.*

¹⁶⁴ *ibid.*

for repayment of the whole loan at once and to sue for the amount due if the borrower refuses to pay.¹⁶⁵

D. Covenants

Once the bank has agreed to lend the desired sum or sums, subject to the fulfilment of the conditions precedent, it will continue to ensure protection throughout the life of the loan by including covenants in the agreement.¹⁶⁶ The function of the covenants is to restrict certain business-related conduct to ensure that the credit rating does not decline whilst the loan is outstanding,¹⁶⁷ thereby protecting the bank's primary interest of ensuring full repayment from the business. Where a covenant is breached, this amounts to an event of default allowing the bank to terminate the loan agreement.¹⁶⁸ The covenants included in an agreement will differ from one loan to another, depending on: a) the credit risk of the business; b) whether the loan is to be secured or unsecured; and c) the bargaining strength of the business and the amount of the loan.¹⁶⁹

A majority of the covenants commonly found in agreements place financial constraints on the business such as ensuring that its current assets exceed its current liabilities.¹⁷⁰ The borrower must also ensure that its paid up share capital and reserves exceed a particular threshold.¹⁷¹ Whilst this appears to be beneficial in the long run as the control performed appears to operate to support the survival of the business, the other covenants may restrict the growth thereof, specifically the covenant that restricts the change of the business.¹⁷²

For example, an established SME may produce a viable business idea that fills a gap in a particular market. However, at the bank's discretion, this may be regarded as a 'material adverse change', which is where the change significantly affects the company's ability to perform its obligations under the relevant agreement, resulting in a prohibition of the business from carrying out a commercially successful idea which would lead to increased turnover.¹⁷³ Likewise, with regard to a very young business, the flexibility to change the business plan without consulting the bank may be the determinant of whether it will have a successful start-up or not. Such a covenant will not be found in an invoice financing agreement. The onus on the client is solely to abide by certain financial undertakings that inherently benefit the business and their relationship with the factor, not to restrict its own economic growth.

A certain degree of control can be identified in some factoring agreements. However, in contrast to a bank loan agreement, this type of control is desired by the

¹⁶⁵ *ibid.*

¹⁶⁶ *ibid* 282.

¹⁶⁷ *ibid.*

¹⁶⁸ L Gullifer and J Payne, *Corporate Finance Law* (Hart Publishing Ltd 2011) 168.

¹⁶⁹ *ibid.*

¹⁷⁰ Ferran and Ho (n 130) 282.

¹⁷¹ *ibid.*

¹⁷² *ibid* 288.

¹⁷³ *Grupo Hotelero Urvasco SA v Carey Value Added SL and another* [2013] EWHC 1039 (Comm), [2013] Bus LR D45.

client. He or she does this by choosing to include as part of the factoring agreement the non-financial functions of factoring, namely the administration and collection service.¹⁷⁴ In relation to the sales ledger administration, the factor ensures that the sales ledger is kept on an 'open item' basis, which allows 'all respective invoices, unpaid parts of invoices and unallocated credit notes and payments to be listed in the sales ledger',¹⁷⁵ enabling the factor to make informed credit decisions from the efficient provision of how much is owed by each customer. This subsequently allows the factor to send clear statements and reminders to the debtors in a timely manner.¹⁷⁶

Furthermore, the additional service of credit security (factoring without recourse to the client) is the most important benefit of factoring.¹⁷⁷ It is generally understood that SMEs do not wish to be encumbered by bad debts and late payments as this leads to adverse effects on the business and imposes unnecessary stress on the business owner. As the factor bears the risk of unpaid debts, this renders the client's profit margin fully protected.¹⁷⁸ Hence, these additional services offered are of great benefit to the business where credit expertise is provided by the factor in controlling the sales ledger, freeing up time for the owner to focus on the priorities of the business such as production, sales and planning – therefore paving the way to economic success. Evidently, factoring safeguards the SME from making a loss through bad debts. In contrast, when it comes to bank lending, where the business struggles to meet its obligation of repaying the overdraft or loan, the business faces further charges added on top of the debt. Moreover, a bank will normally include a 'negative pledge covenant' (NPC) in a loan agreement whereby the business is prohibited from creating further security over its assets without the consent of the bank.¹⁷⁹ Issues arise where the business requires more funds in order to implement a project or a purchase. The bank may make inferences that the business is in a financially unstable position and refuse to grant an extension to the loan. In the event that the business seeks finance elsewhere, the NPC stands to preclude this additional funding, thereby leaving the business inadequately financed. Where a breach of a covenant amounts to an event of default,¹⁸⁰ and if the imposed covenants are exceedingly strict in nature, making it more likely that the business will default, this has the potential to trigger the bank to enforce their right to accelerate repayment of the loan before the end of the agreed period. Hence, the purpose of the loan may be undermined where the covenants attached thereto preclude the business from utilising the funds as intended. Furthermore, where some businesses are unable, or choose not, to afford a lawyer to scrutinise the bank's drafted contract, that business may have no choice but to accept the agreement exactly as drafted by the bank. As a consequence, these businesses will be subject to unnecessarily restrictive covenants, hence more likely to fail.

¹⁷⁴ Ruddy and others (n 52) 32.

¹⁷⁵ *ibid.*

¹⁷⁶ *ibid.*

¹⁷⁷ *ibid* 34.

¹⁷⁸ *ibid* 32.

¹⁷⁹ *ibid* 288.

¹⁸⁰ *ibid* 270.

In sum, and in agreement with Wood, it is submitted that there are dangers in having exceedingly restrictive covenants: such constraints could damage a business' potential for economic growth by preventing it from pursuing worthwhile investment and financing opportunities.¹⁸¹ In some cases, onerous covenants could even prevent the rescue of a company that is in serious financial difficulties or hinder the successful start-up of a new business.

E. Cost

Generally, with a bank loan, the repayment consists of the principal sum borrowed together with the interest charge.¹⁸² Notably, where the loan is secured and the business has defaulted, the business risks losing its assets in addition to repaying the principal and interest, which results in a great loss for the business. Further, where the business wishes to repay the loan earlier than the agreed period, extra charges will be incurred. On the other hand, where the loan is unsecured, the business' growth may be constrained through onerous covenants, thereby restricting the attainment of profits.

As expressed by Petri Mäntysaari, invoice financing, in particular factoring, tends to be more expensive than bank lending.¹⁸³ This is primarily due to the inclusion of administration and collection services and credit protection, which are calculated by a fixed percentage over the sales turnover.¹⁸⁴ Another aspect of invoice financing that contributes to the idea that it is an expensive facility is the percentage of discount charged by the factor within the range of 1.5-3 per cent of the invoice.¹⁸⁵ For instance, where the factor purchases an invoice of £10,000 for £9,700, this means that the client has discounted the invoice by 3 per cent of its face value, resulting in a loss in profit. Essentially, a client is placed in a dilemma: either he or she can choose to wait for full payment, which may cause problems with cash flow, or he or she can pay a factor a fee for immediate prepayment of debts, triggering a loss in profits.

The cost of invoice financing may, at first glance, be greater than the cost of bank lending. However, the costs of factoring are easily justified where prepayment of debts allows the business to grasp new business opportunities and meet day-to-day obligations in the time that they would otherwise wait for their invoices to be paid, which, in some instances, can be up to two months.¹⁸⁶ Therefore, it may be a price worth paying where cash flow can be predicted, enabling the client to plan ahead.

¹⁸¹ P Wood, *International Loans, Bonds, Guarantees, Legal Opinion* (2nd edn, Sweet & Maxwell 2007) 282.

¹⁸² Ferran and Ho (n 130) 270.

¹⁸³ P Mäntysaari, *The Law of Corporate Finance: General Principles and EU Law: Volume III* (Springer-Verlag Berlin Heidelberg 2012) 45.

¹⁸⁴ The Royal Bank of Scotland, 'Invoice Finance FAQs' (Royal Bank of Scotland) <www.rbsif.co.uk/invoice-financing/finance-questions> accessed 13 March 2016.

¹⁸⁵ The Startups Team, 'Invoice Discounting and How It Works' (Start Ups 2014) <www.startups.co.uk/invoice-discounting-how-it-works-and-what-it-costs/> accessed 14 March 2016.

¹⁸⁶ Edmonds (n 3).

F. Time

Whilst arranged overdraft applications take as little as three days to be approved, Khaled Soufani notes that bank loan approval can be a lengthy process, especially for new customers with a high credit risk.¹⁸⁷ This can be detrimental to businesses in need of immediate cash and where the negotiation stage itself can take up to three months.¹⁸⁸

Conversely, within invoice financing, setting up a new agreement with the factor will usually take five to ten days where an advance of up to 90 per cent of the debts is funded within twenty-four hours of submitting an invoice.¹⁸⁹ The latter inevitably applies where an established relationship already exists between the factor and the client.

The immediacy of the advance to which invoice financing gives rise is of particular benefit to businesses operating within seasonal industries.¹⁹⁰ Where a business' customers may take months to pay their invoices, the business may struggle to meet obligations that persist throughout the year, such as paying wages and rent, maintaining plant and machinery and investing in raw materials. Consequently, the prepayment schedule that a factor provides allows the business to function continuously even when they are not actively transacting.¹⁹¹

Thus, taking into consideration the nature of SMEs and the fast-paced environment in which they operate, the time taken to approve a bank loan does not suit obligations that must be met in a timely manner. This places a constraint on the cash flow, potentially damaging the owner's reputation and precluding his or her ability to take advantage of any business opportunities arising at short notice.

By contrast, both a loan and overdraft constitute a liability for the business where a debt is created. This debt could be exacerbated where events of default trigger the bank to accelerate payment of the loan or, under secured lending, assets are seized from the business. Additionally, the element of repaying on demand under the overdraft facility can pose some serious risks to a business. Moreover, where flexibility to act upon gaps within a market is paramount to expansion, the mandatory inclusion of covenants within a loan agreement jeopardises the business' economic growth particularly with regard to restrictions on changes to the business.

The significantly faster process of setting up an invoice finance agreement partnered with the prompt advance of up to 90 per cent of the debts cannot be rivalled by the bank lending process, making the former a more convenient and efficient source of funding. Additionally, not only does invoice finance present itself

¹⁸⁷ K Soufani, 'On the Determinants of Factoring as a Financing Choice: Evidence from the UK' (2002) 54 *Journal of Economics and Business* 1, 3
<<https://poseidon01.ssrn.com/delivery.php?ID=523068119001021112012110066083095024063014039074069003068000070124007093060016001121021125026087069030007067078007052000054017067082115070123066023021002086036042048027067091016079124015101074010100079126065002121118126099083123010025086068021&EXT=pdf>> accessed 14 March 2016.

¹⁸⁸ B Coyle, *Bank Finance* (Financial World Publishing 2002) 138.

¹⁸⁹ The Royal Bank of Scotland (n 184).

¹⁹⁰ Ruddy and others (n 52) 41.

¹⁹¹ The Royal Bank of Scotland, 'Asset Based Lending' (The Royal Bank of Scotland) <www.rbsif.co.uk/asset-based-lending/> accessed 25 February 2016.

as an effective financing technique; it also combats one of the biggest problems currently confronting SMEs: that of late payment of debts. The prepayment of the unpaid debt allows a predictable cash flow to be maintained and enables the SME to continue with its customary business obligations. As a result, survival is safeguarded.

5. Conclusion

It has been argued that, for SMEs, invoice financing deserves to be recognised as an acceptable alternative to the conventional overdraft and term loan. However, there is a lack of educational advice and support on this matter. The evident advice gap must be closed in order to promote the advantageous effects of this financing method.

As well as adequately combatting the ongoing issue of late payment of debts, and consequently easing the burden of cash flow problems, it is evident that the intrinsic benefits of invoice finance can extend beyond the scope of temporarily assisting SMEs in the wake of the recession. This is because the technique of invoice financing allows the business to access working capital without creating a debt obligation. Such a facility can be very cost-effective, which is why invoice financing may also provide ongoing financial assistance to SMEs.

Although transferring ownership of debts is straightforward in English law, under Scots law this may prove difficult due to the associated requirement of 'intimation' of the assignation to the debtor. However, the use of trusts can be applied to circumvent this issue. While this can create complexity in theory, in practice the process of invoice financing is uncomplicated for both the SME and the factor. Despite the fact that invoice financing comes with potentially high costs, particularly as regards factoring, this may be preferable to the often much larger losses occasioned by the restrictive covenants associated with bank lending.

Collectively, SMEs make the largest contribution to the UK economy. Any process that improves their performance should therefore be given serious consideration. The immediate prepayment of debt within an invoice financing arrangement bears the potential to allow SMEs to control their cash flow effectively; to manage their operations in a better way; and to take advantage of business opportunities. To SMEs and, by extension, the economy, invoice financing would constitute an invaluable catalyst.

The Market Stability Reserve Strikes Back: A New Hope for the EU's Greenhouse Gas Emissions Trading System

DARIA DRAGANOVA*

Abstract

The European Union's Emissions Trading System (ETS) was implemented in 2005 with the objective of cutting down the EU's greenhouse gas emissions from the highest emitting industry sectors. By putting a cap on the total amount allowed to be emitted and by permitting allowances to be traded, the ETS has put a price on carbon, aiming to stimulate a reduction in emissions and to incentivise investments in low-carbon technologies. Nevertheless, the system has been criticised for its over-allocation of allowances, a move which led to those allowances being priced at lower rates than expected, and hence, arguably, a failure to spur large-scale investments and innovation. In 2015, the decision for the implementation of a Market Stability Reserve was adopted. The Reserve will come into operation in 2019 and is supposed to function by removing or adding a predetermined amount of allowances from the market. This is regarded as a significant improvement, but there have been a number of concerns in relation to the Reserve's ability to address the over-allocation issue. The present paper critically examines the relevant changes and argues that while the Reserve is a step in the right direction, its structure must be altered in order to give the Reserve flexibility to respond to the external changes and to encourage greater investment in long-term decarbonising strategies.

Keywords: Emissions Trading, European Union, Climate Change, Low-Carbon Economy

1. Introduction

The European Union's Emissions Trading System (ETS), implemented in 2005, is concerned with cutting down emissions from the EU industry sectors emitting the highest levels of greenhouse gases (GHG). The ETS was first introduced to ensure that the EU would meet its targets under the Kyoto protocol. However, nowadays it is regarded as the centrepiece of EU policy for reducing emissions by at least 20 per cent by 2020. It functions by putting a cap on the total amount of GHG allowed to be emitted by those companies covered by the system and by permitting so-called 'EU allowances' (EUAs) to be traded, thus permitting the market to determine the price of them. The higher the price of EUAs, the higher the incentive will be to switch to cleaner technologies and to employ new abatement strategies. While in theory emissions trading is one of the most effective tools for GHG reduction, recent history has demonstrated that the reality is a little more complicated. Over-allocation of

* The author graduated from the University of Aberdeen with a First Class LLB (Honours) degree in 2016. She is currently working with the United Nations' Assistance to the Khmer Rouge Trials in Phnom Penh, Cambodia, and will be pursuing a LL.M. in Energy and Climate Law at the University of Groningen, commencing September 2017.

allowances, and the consequent reduction in carbon prices, has resulted in a lack of investment in clean technologies.

The ETS has undergone a number of changes over the years with arguably the most important being the 2015 decision for the implementation of a Market Stability Reserve (MSR). The Reserve will come into operation in 2019 and will function by adding or removing a predetermined amount of allowances to or from the market. It aims to secure a stable price signal to allow companies to anticipate more accurately future fluctuations in prices and to make the ETS more robust in its response to unforeseen changes in allowance demand. This has been regarded as a significant improvement, but there have been a number of concerns in relation to the Reserve's ability to address the over-allocation issues associated with the trading system.

The present paper examines whether the MSR will be able to address the concerns voiced in the literature. Section 2 provides some general background information on emissions trading as well as an overview of the ETS and its history. Section 3 proceeds with an evaluation of the ETS and the current extent to which it can reduce emissions and spur investment in low-carbon technologies. That section proposes that a correlation exists between the over-allocation of allowances and the currently low price of carbon. Thereafter, it is argued that the price of carbon has been insufficiently high to spur investments in low-carbon technologies, thus preventing the EU from decarbonising the economy in a cost-effective way. Section 4 examines whether the implementation of the MSR was necessary; how the system will function; what aim it pursues; and whether it will be able properly to address the issues voiced in the literature. The section will also argue that the Reserve will make the ETS more robust and better able to respond to external shocks in demand, as well as addressing the current problem of over-allocation by placing the back-loaded allowances directly into the Reserve. The triggers, intervention rates, reaction periods and the decision to implement the Reserve as early as 2019 are critically analysed. In essence, this paper argues that the MSR is suitable to address the currently low carbon price and that it represents a positive step towards addressing the flaws associated with the ETS. However, more must be done to the structure of the MSR in order to give it the flexibility needed to respond to changing circumstances and new technologies, and to spur more investment in long-term decarbonising strategies by returning scarcity to the market more quickly.

2. Background

This section examines how emissions trading systems function in theory and what aims they seek to achieve. It also serves as an introduction to the structure and history of the EU ETS, scrutinising the first two trading phases, which were characterised by decentralised rules, and analysing the movement towards greater harmonisation at the beginning of the first phase (phase I).

A. Emissions Trading

Emissions trading systems have been one of the more liberal tools employed by different states to promote emissions reduction, as they allow flexibility for the market to determine the most cost-effective way of achieving a given target. They are considered environmentally successful as they ensure that a certain limit of pollution will never be exceeded; however, if executed inadequately they may result in merely short-term abatement decisions with little effect on long-term investment strategies and decisions to switch to clean technologies. It will be argued that this problem is true of the ETS.

i. Cost-Effective Emissions Reduction

The main objective of emissions trading systems is effectiveness and efficiency: that is, to reduce carbon emissions at the lowest practicable cost.¹ Environmental effectiveness is achieved as the total quantity of allowances is capped.² Companies are free to decide between investing in improving efficiency to lower emissions and buying allowances from others to cover their excess emissions. Ideally, under an emissions trading system, some firms will be more incentivised actively to reduce their emissions than others because of allowances' opportunity costs – the opportunity to sell their excess EUAs.³ Thus, a company will be free to choose whether to reduce emissions or to pay someone else to do so,⁴ an approach which creates strong economic incentives for reducing emissions:

Economic theory suggests that from the moment carbon is priced, industry will react and develop strategies to reduce emissions in order (...) either [to] buy [fewer] allowances, or [to] sell excess allowances resulting from its reduction activities. (...) Thus, the buyer is paying a charge for polluting and the seller is being rewarded for having reduced emissions.⁵

A reasonable GHG-emitter would choose the cheapest option. This is believed to ensure that, 'the market price of carbon is equal to the lowest marginal abatement cost amongst all controlled sources'; it allows regulated companies to identify and choose the most cost-effective means of reducing emissions.⁶

¹ E Woerdman, M Roggenkamp and M Holwerda, *Essential EU Climate Law* (Edward Elgar Pub 2015) 44-45.

² M Bell, 'Improving the EU Emissions Trading Scheme' (Centre Forum 2008) 1 <www.centreforum.org/assets/pubs/ets.pdf> accessed 25 March 2016.

³ Point Carbon, 'EU ETS Phase II – The Potential and Scale of Windfall Profits in the Power Sector' (2008) 7 <www.wwf.org.uk/filelibrary/pdf/ets_windfall_report_0408.pdf> accessed 25 March 2016.

⁴ European Commission, 'EU ETS Handbook' (European Commission) 6 <http://ec.europa.eu/clima/publications/docs/ets_handbook_en.pdf> accessed 25 March 2016.

⁵ C Egenhofer and others, 'The EU Emissions Trading System and Climate Policy towards 2050: Real Incentives to Reduce Emissions and Drive Innovation?' (Centre for European Policy Studies 2011) 8 <http://observgo.quebec.ca/observgo/fichiers/88303_The%20EU%20ETS%20and%20Climate%20Policy%20towards%202050-5.pdf> accessed 25 March 2016.

⁶ *ibid* 2.

ii. Investment and Innovation

Another objective of emissions trading systems is to encourage innovation and investment in low-carbon technologies.⁷ Indeed, one of the key aims of the ETS is to influence decision-making regarding low-carbon technologies.⁸ One means of achieving this goal has been the establishment of a New Entrants' Reserve,⁹ as well as the utilisation of auction revenues for supporting the deployment of new technologies; however, a major driver of low-carbon investments is the price of carbon itself – the higher that is, the greater the incentive for cutting emissions will be.¹⁰

In order for the system to stimulate the switch to low-carbon alternatives it must provide a stable price signal allowing investors to make decisions in the long term.¹¹ However, during 'phase I' and 'phase II' of the operation of the ETS, one of the major concerns expressed related to significant price fluctuation: 'low allowance price combined with periodic instability meant that the ETS has had [a] limited impact on low-carbon investments.'¹² Furthermore, prices below €20 per tonne of CO₂ have been seen as too low to drive investment in low-carbon technologies.¹³ Low prices have the effect of stimulating short-term, rather than long-term, investments, a problem which seems to characterise the ETS. This has also highlighted the need for greater certainty, something which can only be achieved by setting long-term goals to restore investors' confidence and to reduce price volatility.¹⁴

B. The European Union's Emissions Trading System

The EU ETS is the largest greenhouse gas emissions system in operation. It was introduced in 2005 by Directive 2003/87/EC and has so far been running for three phases – 'phase I' from 2005 to 2007, 'phase II' from 2008 to 2012, and 'phase III'

⁷ Directive 2003/87/EC of the European Parliament and of the Council of 13 October 2003 establishing a scheme for greenhouse gas emission allowance trading within the Community and amending Council Directive 96/61/EC, OJ L 275/32, para 20.

⁸ T Laing and others, 'Assessing the Effectiveness of the EU Emissions Trading System' (Centre for Climate Change Economics 2013) para 3.1 <www.lse.ac.uk/GranthamInstitute/wp-content/uploads/2014/02/WP106-effectiveness-eu-emissions-trading-system.pdf> accessed 25 March 2016.

⁹ Directive 2003/87/EC, art 10a(7) as amended by Directive 2009/29/EC.

¹⁰ European Commission, 'The EU Emissions Trading System (EU ETS)' (European Commission 2013) 5 <http://ec.europa.eu/clima/publications/docs/factsheet_ets_en.pdf> accessed 25 March 2016.

¹¹ C Egenhofer, A Marcu and A Georgiev, 'Reviewing the EU ETS Review? Report of the CEPS Task Force on Does the Market Produce the "Right" Price Signal' (Centre for European Policy Studies 2012) 3 <www.ceps.eu/system/files/TF%20Report%20Reviewing%20the%20EU%20ETS_0.pdf> accessed 25 March 2016.

¹² M Grubb, 'Strengthening the EU ETS: Creating a Stable Platform for EU Energy Sector Investment' (Climate Strategies 2012) 14 <<http://climatestrategies.org/wp-content/uploads/2012/03/cs-strengtheningtheuets-fullreport.pdf>> accessed 25 March 2016.

¹³ S Tilford, 'How to make EU Emissions Trading a Success' (Centre for European Reform 2008) 16 <www.cer.org.uk/sites/default/files/publications/attachments/pdf/2011/p_769-275.pdf> accessed 25 March 2016.

¹⁴ Bell (n 2) 4.

from 2013 to 2020.¹⁵ The system was first introduced in order to ensure the achievement of the targets set by the Kyoto protocol, which has been signed by 192 countries to date.¹⁶

The system works by effectively putting a cap on all emissions and reducing that cap over time.¹⁷ Member States are allocated allowances for the emission of certain greenhouse gases, and they can auction those allowances off or distribute them for free between specific companies within their territory. The cap represents the maximum amount of greenhouse gas emissions that are allowed to be emitted by all installations covered by the regime.¹⁸ In 2013, the cap was set to decline by 1.74 per cent annually,¹⁹ and it was recently proposed to increase this rate of reduction to 2.2 per cent.²⁰

Each company applies for a Greenhouse gas permit²¹ and is then allocated allowances, each permitting the emission of one tonne of carbon dioxide or its equivalent for a specified period.²² At the end of each year companies must surrender an amount of allowances equivalent to what they have emitted during that period.²³ If a company is unable to do so it will incur heavy fines, set in a manner which makes them 'effective, proportionate and dissuasive'.²⁴ Since 'phase II', all allowances that are not surrendered at the end of the year, are carried over to the next one, a practice known as 'banking'.²⁵ Each operator of an installation is required to monitor and report their emissions levels to the competent authority every year,²⁶ and Member States must ensure that the reports are verified in accordance with a set of criteria.²⁷ Since the beginning of 'phase III', issued allowances have been held in the Community registry.²⁸ Companies are free to trade allowances between themselves in order to ensure compliance. Any issue, transfer or

¹⁵ European Parliament and Council Directive 2003/87/EC of 13 October 2003 establishing a scheme for greenhouse gas emission allowance trading within the Community and amending Council Directive 96/61/EC [2003] OJ L 275/32; European Commission (n 4) 4.

¹⁶ United Nations Framework Convention on Climate Change, 'Status of Ratification of the Kyoto Protocol' (UN 2014) <http://unfccc.int/kyoto_protocol/status_of_ratification/items/2613.php> accessed 25 March 2016.

¹⁷ Directive 2003/87/EC, art 9.

¹⁸ Woerdman Roggenkamp and Holwerda (n 1) 53.

¹⁹ European Parliament and Council Directive 2009/29/EC amending directive 2003/87/EC so as to improve and extend the greenhouse gas emission allowance trading scheme of the Community [2009] OJ L 140/63, art 9.

²⁰ European Commission, 'Proposal for a Directive of the European Parliament and of the Council amending Directive 2003/87/EC to enhance cost-effective emission reductions and low-carbon investments' COM (2015) 337 final, art 1(3).

²¹ Directive 2003/87/EC, art 5.

²² *ibid* art 3(a).

²³ *ibid* art 6(2)(e).

²⁴ *ibid* art 16.

²⁵ Commission Regulation No 920/2010 of 7 October 2010 for a standardised and secured system of registries pursuant to Directive 2003/87/EC of the European Parliament and of the Council; Decision No 280/2004/EC of the European Parliament and of the Council, art 57.

²⁶ Directive 2003/87/EC, art 14(3).

²⁷ *ibid* art 15.

²⁸ *ibid* art 19.

cancellation of allowances is recorded in a transaction log maintained by a Central Administrator.²⁹

Additionally, under the Kyoto Protocol, Member States are permitted to use project credits under the Clean Development Mechanism (CDM), which are awarded to firms investing in projects in developing countries, or under the Joint Implementation (JI) system, awarded for projects in other developed countries,³⁰ in order to cover their excess emissions.³¹ The beginning of ‘phase III’ imposed a limitation on their use.

C. The 2009 Amendments

The main objective of the EU ETS is to reduce emissions. Thus, ‘it is crucial to understand whether the first two phases have encouraged participating installations to abate.’³² A number of significant changes were introduced by the 2009 Directive, which aimed to address the concerns associated with the system.

i. Cap and Scope

The ETS was divided into a number of trading phases. During ‘phase I’ and ‘phase II’, Member States were to submit National Allocation Plans (NAPs) to justify the, ‘total number of allowances created for the trading period, provide a list of covered installations, and explain how those allowances [were] to be distributed.’³³ The cap effectively represented the total number of allowances from the different Member States combined. This was criticised for being too decentralised and in need of greater harmonisation, since national circumstances would influence the amount of and allocation method for allowances within each jurisdiction.³⁴

Member States were responsible for determining a number of factors such as administrative arrangements in relation to auctioning, monitoring and verification procedures.³⁵ This resulted in different practices being adopted and led to some types of installation being included in some Member States but not in others. These practices included the exclusion of combustion installations which did not produce electricity in the UK during the first two phases.³⁶

²⁹ *ibid* art 20.

³⁰ National Audit Office, ‘European Emissions Trading Scheme’ (National Audit Office 2009) 14 <www.nao.org.uk/wp-content/uploads/2012/10/0809_EU_emissions_trading.pdf> accessed 25 March 2016.

³¹ Directive 2003/87/EC, art 30(3).

³² R Martin, M Muûls and U Wagner, ‘An Evidence Review of the EU Emissions Trading System, Focussing on Effectiveness of the System in Driving Industrial Abatement’ (Department of Energy and Climate Change 2012) 16 <www.gov.uk/government/uploads/system/uploads/attachment_data/file/48449/5725-an-evidence-review-of-the-eu-emissions-trading-sys.pdf> accessed 25 March 2016.

³³ AD Ellerman and PL Joskow, ‘The European Union’s Emissions Trading System in Perspective’ (Pew Centre on Global Climate Change 2008) 5 <www.c2es.org/docUploads/EU-ETS-In-Perspective-Report.pdf> accessed 25 March 2016.

³⁴ *ibid* 5.

³⁵ National Audit Office (n 30) 18.

³⁶ *ibid*.

At the end of 'phase I' it became apparent that the allocation had been too generous, as allowances exceeded verified emissions by almost 7 per cent.³⁷ Thus, the beginning of 'phase II' saw the Commission reducing the NAPs of twelve Member States in order to align them with Kyoto reduction targets.³⁸ The 2009 Directive substituted NAPs with an EU-wide emissions cap, which was to decrease annually by 1.74 per cent.³⁹ The earlier phases of the system have been criticised for 'a general lack of ambition',⁴⁰ with, 'the NAP process prov[ing] to be [a] long, laborious, and unrewarding procedure for all concerned.'⁴¹ Nevertheless, NAPs were a necessary compromise in order to, 'get any system agreed at all', even at the risk of sacrificing its effectiveness at achieving emissions reductions.⁴²

It must also be noted that the amending Directive broadened the scope of covered installations to include, *inter alia*, oil refineries and steel works,⁴³ with aviation to be added to the list from 2012 onwards.⁴⁴ Extending the scope to more sectors is beneficial; it will permit emissions reduction to take place in the contexts where this would be most efficient.⁴⁵ It has been said that, '[t]he broader the scope of a cap and trade scheme, the greater the resources available for reducing emissions and the less[er] the impact on competitiveness.'⁴⁶

ii. Allocation

During 'phase I' and 'phase II' respectively, 95 per cent and 90 per cent of all emissions allowances were allocated for free to all the covered installations,⁴⁷ the aim being to address the initial 'teething problems'.⁴⁸ Free allocation was necessary at the start in order to gain acceptance from the industry sector.⁴⁹ However, it remained

³⁷ European Environment Agency, 'EU Emissions Trading System (ETS) Data Viewer' (European Environment Agency 2016) 6 <www.eea.europa.eu/data-and-maps/data/data-viewers/emissions-trading-viewer> accessed 25 March 2016.

³⁸ M Verdonk and others, 'Evaluation of Policy Options to Reform the EU Emissions Trading System: Effects on Carbon Price, Emissions and the Economy' (PBL Netherlands Environmental Assessment Agency 2013) 8 <www.pbl.nl/sites/default/files/cms/publicaties/PBL_2013_Evaluation-of-policy-options-to-reform-the-EU-ETS_934.pdf> accessed 25 March 2016.

³⁹ Directive 2003/87/EC, art 9 as amended by Directive 2009/29/EC.

⁴⁰ S Bell, D McGillivray and O Pedersen, *Environmental Law* (8th edn, OUP 2013) 553.

⁴¹ AD Ellerman, C Marcantonini and A Zaklan, 'The EU ETS: Eight Years and Counting' (European University Institute 2014) 3 <http://cadmus.eui.eu/bitstream/handle/1814/29517/RSCAS_2014_04.pdf> accessed 25 March 2016.

⁴² Bell, McGillivray and Pedersen (n 40) 553.

⁴³ Woerdman, Roggenkamp and Holwerda (n 1) 54.

⁴⁴ Directive 2008/101/EC of the European Parliament and of the Council of 19 November 2008 amending Directive 2003/87/EC so as to include aviation activities in the scheme for greenhouse gas emission allowance trading within the Community [2009] OJ L 8/3.

⁴⁵ Bell (n 2) 8-9.

⁴⁶ Tilford (n 13) 12.

⁴⁷ Point Carbon (n 3) 6.

⁴⁸ European Commission, 'Report on the functioning of the European carbon market Accompanying the document Report from the Commission to the European Parliament and the Council: Climate action progress report, including the report on the functioning of the European carbon market and the report on the review of Directive 2009/31/EC on the geological storage of carbon dioxide' COM (2015) 576 final, 6.

⁴⁹ Woerdman, Roggenkamp and Holwerda (n 1) 56.

doubtful whether it had managed to secure that aim. Furthermore, ‘grandfathering’ (i.e. allocation based on historical emission levels) was the preferred method for allocation during the first two phases.⁵⁰ It was criticised as favouring ‘dirtier’, less-efficient firms by allocating more allowances to those that emit a higher amount of greenhouse gases. It has also been submitted that this method of allocation encouraged firms to increase their emissions in order to receive a higher number of allowances in the future,⁵¹ which arguably led to the, ‘perverse effect of encouraging investment in carbon heavy power generators over low-carbon alternatives’.⁵² ‘Grandfathering’ was criticised for incentivising firms to keep heavily polluting installations in operation for longer in order to retain the associated allowances.⁵³

From 2013 onwards, companies were allocated free allowances based on established ‘benchmarks’ determined on the, ‘average performance of the 10 per cent most efficient installations in a sector (...) in the years 2007-2008.’⁵⁴ This has been seen to encourage the use of the most efficient technologies.⁵⁵ Furthermore, in creating a level playing field for producers in every country, this method of allocation arguably achieves a greater degree of harmonisation.⁵⁶ While this method poses its own difficulties, for example disagreements on uniform product benchmarks,⁵⁷ it should be welcomed as a positive step towards the implementation of the ‘polluter pays’ principle and towards encouraging the deployment of cleaner technology.

Since 2013, all allowances that have not been allocated for free are to be auctioned.⁵⁸ Higher emphasis was placed on auctioning in ‘phase III’, making it the default method of allocation.⁵⁹ At least 50 per cent of allowances are to be auctioned,⁶⁰ and Member States are to use at least 50 per cent of the revenue raised to combat climate change by such means as reducing both greenhouse gas emissions and deforestation, and by supporting renewable energy development, et cetera.⁶¹ Auctioning is believed to be more efficient as companies will only buy what they need in order to cover their emissions. Furthermore, it will reduce the cost of running the system as well as the bureaucracy associated with the allocation process. Lastly, it will require polluters to pay by internalising the cost of their pollution. These changes are to be welcomed as a positive movement which rewards cleaner companies; stimulates innovation; and scores higher on efficiency. The risk of carbon

⁵⁰ S de Bruyn and others, ‘Does the Energy Intensive Industry Obtain Windfall Profits through the EU ETS? An Econometric Analysis for Products from the Refineries, Iron and Steel and Chemical Sectors’ (CE Delft 2010) 13

<www.ce.nl/publicatie/does_the_energy_intensive_industry_obtain_windfall_profits_through_the_eu_ets/1038> accessed 25 March 2016.

⁵¹ National Audit Office (n 30) 25.

⁵² Bell (n 2) 5.

⁵³ National Audit Office (n 30) 26.

⁵⁴ Directive 2003/87/EC, art 10a(2) as amended by Directive 2009/29/EC.

⁵⁵ National Audit Office (n 30) 26.

⁵⁶ Ellerman and Joskow (n 33) 37-39.

⁵⁷ *ibid* 36.

⁵⁸ Directive 2003/87/EC, art 10(1) as amended by Directive 2009/29/EC.

⁵⁹ Woerdman, Roggenkamp and Holwerda (n 1) 56.

⁶⁰ Bell, McGillivray and Pedersen (n 40) 552.

⁶¹ Directive 2003/87/EC, art 10(3) as amended by Directive 2009/29/EC.

leakage (discussed further below) undeniably exists.⁶² Nevertheless, firms that have been identified as 'at risk' of carbon leakage would still receive a part of their allowances for free, thus minimising that risk.⁶³ This is an acceptable compromise between ensuring a cost-optimal distribution of allowances and addressing the concerns associated with shifting production abroad.

iii. Carbon Leakage

Member States had to take into consideration the effect that a carbon price would have on the international competitiveness of EU companies. Higher production costs would incentivise companies to shift their, 'productive capacity from one country to another as a result of [a] differential emissions pricing policy',⁶⁴ commonly referred to as carbon leakage.⁶⁵

Economic theory suggests that firms would pass through the cost of the allowances obtained for free if that enhanced their profitability and outweighed the additional costs.⁶⁶ Evidence has been put forward suggesting that the ETS has not impeded the competitiveness of firms and sectors that can simply pass on the costs for EUAs to the final consumer.⁶⁷ However, many firms may not have the opportunity to pass on the carbon costs without any negative consequences such as losing market shares – something which would undermine 'the environmental effectiveness and political acceptability' of the system.⁶⁸ Free allocation has a lesser impact on companies;⁶⁹ however, free allocation would not be sufficient on its own to dissuade industries in global competition from relocating their businesses. Other factors such as high levels of regulation, higher minimum wages etc, have added a cost to EU businesses, influencing their decision to relocate.⁷⁰ Thus the extent to which free allocation has prevented carbon leakage should be regarded as limited.⁷¹

3. Evaluation of the EU ETS

The system was first introduced to ensure fulfilment of the emissions targets under the Kyoto protocol, and would later grow to become one of the central elements of EU environmental policy and legislation.⁷² Its primary objective is to, 'promote reduction of greenhouse gas emissions in a cost-effective and economically efficient

⁶² Bruyn and others (n 50) 15.

⁶³ Directive 2003/87/EC, art 10a as amended by Directive 2009/29/EC.

⁶⁴ Laing and others (n 8) para 4.

⁶⁵ Verdonk and others (n 38) 41.

⁶⁶ Bruyn and others (n 50) 7.

⁶⁷ Martin, Muûls and Wagner (n 32) 27.

⁶⁸ *ibid* 24.

⁶⁹ Bruyn and others (n 50) 14.

⁷⁰ Bell (n 2) 6-7.

⁷¹ *ibid* 7.

⁷² European Commission, 'Impact Assessment Accompanying the document Proposal for a Decision of the European Parliament and of the Council concerning the establishment and operation of a market stability reserve for the EU greenhouse gas emission trading scheme and amending Directive 2003/87/EC' SWD (2014) 17 final, 11.

manner.⁷³ The EU has a goal to reduce emissions by at least 20 per cent by 2020 compared with 1990 levels,⁷⁴ and emissions trading is considered to be the most cost-efficient means of meeting that target.⁷⁵

The EU ETS has a secondary aim: the stimulation of investment in low-carbon technologies. In order to understand why the 2015 reform was necessary and whether the changes are fit to remedy the flaws of the current system, one must first identify those shortfalls. The following section examines the question whether the ETS has been successful at achieving carbon emission cuts in a cost-effective way and whether it has contributed to investment in technological innovation.

A. Over-allocation

Over-allocation is problematic for two reasons: it fails to encourage emissions reduction, and it fails to incentivise investment in low-carbon technologies. During ‘phase I’, verified emissions were lower than allocated allowances by almost 7 per cent.⁷⁶ The lack of consistent historical data prior to 2005 was identified as the main reason behind the EU’s failure to set sufficiently tight emissions caps.⁷⁷ States allocated allowances generously in order to safeguard the competitiveness of their industries.⁷⁸ Allocations were largely based on industries’ predictions of how much they would need;⁷⁹ predictions which, in hindsight, seem to have been overly optimistic.⁸⁰ However, the over-allocation in ‘phase I’ did not have a major effect afterwards, seeing as unused allowances could not be banked for ‘phase II’ compliance purposes.

‘Phase II’ saw the Commission reducing NAPs by an average of 10 per cent.⁸¹ However, the economic crisis in 2008-2009 resulted in a lowered production output and hence reduction in emissions, which led to a fall in demand for EUAs.⁸² The financial crisis had a noticeable effect on both GDP and industrial output:⁸³ ‘[t]he combined impact of [the] recession, [the] response to the carbon price in 2008-2011,

⁷³ Directive 2003/87/EC, art 1.

⁷⁴ European Commission ‘20 20 by 2020: Europe’s Climate Change Opportunity’ (Brussels 23.01.2008) COM (2008) 30 final.

⁷⁵ Woerdman, Roggenkamp and Holwerda (n 1) 44.

⁷⁶ National Audit Office (n 30) 19.

⁷⁷ Tilford (n 13) 17.

⁷⁸ Woerdman, Roggenkamp and Holwerda (n 1) 65.

⁷⁹ Egenhofer and others (n 5) 5.

⁸⁰ Woerdman, Roggenkamp and Holwerda (n 1) 65.

⁸¹ C Kettner, D Kletzan-Slamanić and A Köppl, ‘The Challenge of the European Carbon Market: Emission Trading, Carbon Leakage and Instruments to Stabilise the CO₂ Price. The EU Emission Trading Scheme – Sectoral Allocation Patterns and the Effects of the Economic Crisis’ (Austrian Institute of Economic Research 2012) 3 <[www.researchgate.net/publication/254457030_ETCLIP_-_The_Challenge_of_the_European_Carbon_Market_Emission_Trading_Carbon_Leakage_and_Instruments_to_Stabilise_the_CO₂_Price_The_EU_Emission_Trading_Scheme_Sectoral_Allocation_Patterns_and_the_Effe](http://www.researchgate.net/publication/254457030_ETCLIP_-_The_Challenge_of_the_European_Carbon_Market_Emission_Trading_Carbon_Leakage_and_Instruments_to_Stabilise_the_CO2_Price_The_EU_Emission_Trading_Scheme_Sectoral_Allocation_Patterns_and_the_Effe)> accessed 25 March 2016.

⁸² E Van der Werf and others, ‘Quantifying the Effects of Reforming the EU Emissions Trading System. A Computable General Equilibrium Analysis’ (Green Growth Knowledge Platform 2014) 5 <www.greengrowthknowledge.org/sites/default/files/VanderWerf_Quantifying_the_effects_of_reforming_the_EU_Emissions_Trading_System.pdf> accessed 25 March 2016.

⁸³ Ellerman, Marcantonini and Zaklan (n 41), 8.

and complementary measures, ha[s] led to a surplus of allowances that will last out to 2020'.⁸⁴ Additionally, the sale of the 300 million unused allowances held in the New Entrants' Reserve contributed to the lack of scarcity in the market.⁸⁵

Lower economic activity reduces the demand for allowances. However, while in a natural market this would result in the supply side reacting to that change as well,⁸⁶ in the ETS the supply is fixed.⁸⁷ In 'phase II' the total supply exceeded the total demand by 2,095 credits, which were banked for 'phase III',⁸⁸ carrying forward a total of about 2 billion tonnes.⁸⁹ EUA banking helped to stabilise the price and resulted in diminished price volatility;⁹⁰ however, it was regarded as potentially having an adverse impact on the effectiveness of the 'phase III' cap.⁹¹

Furthermore, the use of international credits under the CDM and JI have further deepened the issue.⁹² On the one hand, the use of CDM and JI would promote the transfer of investments and technology in poorer countries and thus would increase the cost-effectiveness of the system by allowing companies to reduce emissions where this can be done cheaply.⁹³ Moreover, a limit on the total number of credits which could be used would ensure a certain level of domestic reduction.⁹⁴ On the other hand, the allowable emissions exceed the cap since they also include allowable project credits.⁹⁵ Installations which have sufficient EUAs to cover their emissions might choose to sell them and buy cheaper Certified Emissions Reductions (CERs) or European Reduction Units (ERUs).⁹⁶ This is problematic since:

[o]ffset programs [*sic*] award credits based on reductions from a projected business-as-usual scenario, thereby creating perverse incentives to exaggerate pre-project emissions in order to 'earn' additional credits, even while absolute emissions increase.⁹⁷

Moreover, interaction with other policy instruments, such as the EU's target to make 20 per cent of energy consumed renewable and to improve energy efficiency by 20

⁸⁴ Grubb (n 12), 5.

⁸⁵ Verdonk, and others (n 38) 14.

⁸⁶ Ellerman and Joskov (n 33) 5.

⁸⁷ Egenhofer, Marcu and Georgiev (n 11) 4.

⁸⁸ Van der Werf and others (n 82) 5.

⁸⁹ European Commission (n 72) 8.

⁹⁰ Environmental Defence Fund, 'European Union: An Emissions Trading Case Study' (Environmental Defence Fund and others 2015) 14

<www.ieta.org/resources/Resources/Case_Studies_Worlds_Carbon_Markets/euets_case_study_may2015.pdf> accessed 25 March 2016.

⁹¹ National Audit Office (n 30) 40.

⁹² *ibid* 10.

⁹³ Tilford (n 13) 23.

⁹⁴ Egenhofer and others (n 5) 17.

⁹⁵ National Audit Office (n 30) 19.

⁹⁶ *ibid* 24.

⁹⁷ LM Brown, A Hanafi and A Peterson, 'The EU Emissions Trading System: Results and Lessons Learned' (Environmental Defence Fund 2012) 23

<www.edf.org/sites/default/files/EU_ETS_Lessons_Learned_Report_EDF.pdf> accessed 25 March 2016.

per cent by 2020, has contributed to the issue.⁹⁸ The success of other policies has led to further abatement by companies covered by the ETS, which additionally reduces the demand for EUAs.⁹⁹ The EUA price fell ‘significantly’ after the approval of the Energy Efficiency Directive.¹⁰⁰ It is, additionally, worth noting that individual Member States can adopt domestic climate change and energy policies, which would also have an effect on emissions.¹⁰¹ It is submitted that the existence of these other instruments creates problems. Because the ETS is unable to respond to their success the price of EUAs will be vulnerable to further reduction.

Emissions reduction is the central objective of the ETS.¹⁰² If the supply of EUAs exceeds demand, the ETS will not incentivise emissions-reduction seeing as ‘business-as-usual’ emissions will fall below the cap. Some commentators regard cap-and-trade markets as a ‘technologically neutral instrument of price discovery’.¹⁰³ The ETS is a quantity-based instrument, which means that the carbon market is free to determine the price, which would normally be the least costly option.¹⁰⁴ Thus, as long as emissions do not exceed the cap, there will be no basis for intervention. However, the driving idea behind the ETS was to reduce emissions below ‘business-as-usual’ levels; otherwise the cap would not be regarded as binding:¹⁰⁵ ‘[t]he stringency of the emission[s] cap can be interpreted as [an] indication of the environmental effectiveness of the emissions trading system.’¹⁰⁶ Furthermore, over-allocation creates a problem because it leads to low prices, which, in turn, would fail to spur investment in low-carbon technologies: ‘[a] large surplus (...) strongly confounds the signal for investments, which are necessary for the transition into a low-carbon economy’.¹⁰⁷ The effects that low and fluctuating prices have on abatement and investment strategies are examined in greater detail further below.

Some commentators have argued that the initial over-allocation has become ‘less of a problem’, since now caps are determined with regard to past verified emissions and the linear reduction factor has aided the establishment of an appropriate cap, based on all the data collected in previous years.¹⁰⁸ However, fears

⁹⁸ European Commission (n 74).

⁹⁹ P del Rio, ‘Analysing the Interactions of Renewable Energy Promotion and Energy Efficiency Support Schemes: The Impact of Different Instruments and Design Elements’ (Institute for Public Policies and Goods 2009) <www.aee.at/2009-IAEE/uploads/fullpaper_iaee09/P_117_DEL%20RIO_PABLO_25-Aug-2009,%208:57.pdf> accessed 25 March 2016.

¹⁰⁰ Egenhofer, Marcu and Georgiev (n 11) 7-8.

¹⁰¹ Ellerman, Marcantonini and Zaklan (n 41) 6.

¹⁰² Directive 2003/87/EC, art 1.

¹⁰³ A Marcu, ‘The Market Stability Reserve in Perspective’ (Centre for European Policy Studies 2014) 2 <www.ceps.eu/publications/market-stability-reserve-perspective> accessed 25 March 2016.

¹⁰⁴ O Sartor, ‘The EU ETS Carbon Price: To Intervene, or Not to Intervene?’ (Climate Brief 2012) 1 <www.cdclimat.com/IMG/pdf/12-02_climate_brief_12_-_the_eu_ets_carbon_price_-_to_intervene_or_not_to_intervene.pdf> accessed 25 March 2016.

¹⁰⁵ C Kettner, D Kletzan-Slamanig and A Köppl, ‘The EU Emissions Trading Scheme – Sectoral Allocation Patterns and Factors Determining Emissions Changes’ (WIFO 2013) 5-6 <www.wifo.ac.at/jart/prj3/wifo/resources/person_dokument/person_dokument.jart?publikationsid=46299&mime_type=application/pdf> accessed 25 March 2016.

¹⁰⁶ Kettner, Kletzan-Slamanig and Köppl (n 81) 5.

¹⁰⁷ European Commission (n 72) 9.

¹⁰⁸ Environmental Defence Fund and others (n 90) 11-13.

that the oversupply of allowances will continue past 2020 have been expressed in the literature.¹⁰⁹ The total surplus of allowances was estimated to have reached 2.7 billion from 2013-2014,¹¹⁰ and, assuming a scenario where no further measures are adopted, the surplus would be expected only to decline to 1.5 billion by 2028.¹¹¹

B. The Carbon Price

The start of the system marked prices around €5-€10, which quickly rose and stayed above €20. Nevertheless, in April 2006 when verified emissions reports were published and it became apparent that 'phase I' emissions would not exceed the cap, the prices plummeted to almost €0 in late 2007. The beginning of 'phase II' saw prices above €20 until the 2008 economic crisis slowed production and lowered CO₂ emissions, which led to prices of about €10. 'Phase III' marked no major fluctuations; however, prices remained at levels lower than expected – below €8.¹¹² The price of carbon is a major driver for innovation; hence, it is important to evaluate the effect that price fluctuation has had on driving investment in low-carbon technologies and to determine whether the current price level is able to promote investments in innovation.

The ETS is not a naturally developing market in the sense that there is no correlation between supply and demand for allowances. There is always a degree of government intervention in setting the cap and allocating allowances.¹¹³ Economic theory suggests that when the supply of a certain commodity remains constant while the demand therefor drops, the price consequentially drops as well. In the ETS, the supply is set at a certain level – the cap – while the demand fluctuates due to a number of economic factors, including, 'changes in economic activity, weather events, fuel prices, and technology developments.'¹¹⁴ In fact, the ETS allowance prices have been less volatile than those of other commodities such as fossil fuels and even oranges and coffee.¹¹⁵ Some commentators have pointed out that price volatility is part of the regular functioning of the market and that this should not be taken as a sign of market failure.¹¹⁶ Others have suggested that the lower prices following the recession have, 'softened the impact of pollution regulations on

¹⁰⁹ Egenhofer, Marcu and Georgiev (n 11) 1.

¹¹⁰ Karsten Neuhoff and others, 'Banking of Surplus Emissions Allowances: Does the Volume Matter?' (DIW Berlin 2012) 5 <www.diw.de/documents/publikationen/73/diw_01.c.394484.de/dp1196.pdf> accessed 25 March 2016.

¹¹¹ European Commission (n 72) 8.

¹¹² EEX, 'European Emission Allowances Auction (EUA) | Global Environmental Exchange' (EEX 2016) <www.eex.com/en/market-data/emission-allowances/auction-market/european-emission-allowances-auction#!/2016/02/16> accessed 25 March 2016.

¹¹³ Egenhofer and others (n 5) para 2.2.

¹¹⁴ H Fell and RD Morgenstern, 'Alternative Approaches to Cost Containment in a Cap-and-Trade System' (Resources for the Future 2009) 9-14 <rff.org/RFF/Documents/RFF-DP-09-14.pdf> accessed 16 February 2016.

¹¹⁵ Brown, Hanafi and Petsonk (n 97) 16.

¹¹⁶ *ibid* 15.

businesses during the difficult economic times'; however, subsequent economic growth has not been accompanied by an increase in prices.¹¹⁷

The fall in the price of EUAs at the end of 2007 could be attributed to the lack of 'bankability' of those allowances, meaning that firms which finished the year with a surplus of 'phase I' allowances could not use them for compliance purposes in 'phase II'.¹¹⁸ Therefore, banking of allowances is seen as promoting early reduction and as saving the excess EUAs in anticipation of future higher prices and more stringent emissions caps.¹¹⁹

The main reason why market participants bank allowances is to meet future carbon needs, thus lessening exposure to adverse variations in price – also known as hedging.¹²⁰ Other motives include arbitrage (buying allowances and signing future contracts) and speculation (carrying the risk in anticipation of a future rise in prices).¹²¹ Such practices as hedging are common amongst participants in a number of markets as a means of minimising exposure to price fluctuations.¹²² Hedging in order to safeguard against future uncertainties is the reason why the carbon price did not fall to €0 in 'phase III'.¹²³ Nevertheless, allowing producers to hedge allowances has provided them with the option of delaying low-carbon investments.¹²⁴ While banking allowances prevented prices from plummeting at the end of 'phase II', it increased the cap in 'phase III', which led to the current low price level.¹²⁵

Price volatility in the ETS, resulting from the inflexibility of supply, is greater than in natural markets and it has had a major impact on investment decisions. On the one hand, some commentators claim that low EUA prices have managed to spur investment in low-carbon technologies since, 'the expectation of a carbon price five, ten and twenty years from now is far more important in influencing the long-term investment that is essential for low-carbon development.'¹²⁶ However, the expectation that fossil fuels will become more expensive and eventually scarce would also act as an incentive in much the same manner as a future high carbon price. It has, consequently, been stated that, 'additional tools are needed to accelerate the natural rate of investment and technological change'.¹²⁷

¹¹⁷ L Taschini, S Kollenberg and C Duffyat, 'System Responsiveness and the European Union Emissions Trading System: Policy Paper' (Centre for Climate Change Economics and Policy 2014) 9-10 <www.lse.ac.uk/GranthamInstitute/wp-content/uploads/2014/01/System-responsiveness-and-the-EU-ETS.pdf> accessed 25 March 2016.

¹¹⁸ Ellerman, Marcantonini and Zaklan (n 41) 11-12.

¹¹⁹ Brown, Hanafi and Petsonk (n 97) 17-19.

¹²⁰ A Schopp and K Neuhoff, 'The Role of Hedging in Carbon Markets' (Deutsches Institut für Wirtschaftsforschung 2013) 2

<www.diw.de/documents/publikationen/73/diw_01.c.416987.de/dp1271.pdf> accessed 25 March 2016.

¹²¹ Neuhoff (n 110) 5.

¹²² Ellerman and Joskov (n 33) 42.

¹²³ Verdonk and others (n 38) 20.

¹²⁴ *ibid* 22.

¹²⁵ Van der Werf and others (n 82) 6.

¹²⁶ Brown, Hanafi and Petsonk (n 97) 15.

¹²⁷ Egenhofer and others (n 5) v.

On the other hand, a number of studies have concluded that price fluctuations and regulatory instability have stifled investment.¹²⁸ One observes that:

The knowledge that prices will remain high over a relatively short time horizon is enough to sway a decision in favour of gas and away from coal, or to justify investment in greater energy efficiency. But it is not sufficient to persuade companies to invest in new technologies, such as CCS (Carbon Capture and Storage). The time horizons for such investments are very long (...) so companies need to be confident that carbon will remain expensive throughout the term of the investment.¹²⁹

The price of carbon has become a driver for small-scale investment decisions.¹³⁰ There is a need for a clear price signal – lower or fluctuating price levels, combined with regulatory uncertainty, have contributed to the failure of the system to incentivise major investments in low-carbon technologies.¹³¹ In order for the ETS to be efficient, it must provide incentives for innovation as well as emissions abatement.¹³² Investments will only be made if they are considered sound by their investors, and so, ‘the carbon price will only influence the type of investment by making high-carbon investments less profitable.’¹³³

The low carbon prices have been criticised for being unable to spur innovation, which in turn may lead to companies investing in carbon-intensive technologies, thus making investment ‘stranded’.¹³⁴ If companies expect prices to remain low, they may choose to invest in carbon-intensive capital stock, making it extremely difficult to switch to low-carbon technologies in the future.¹³⁵ Furthermore, ‘the learning curve [associated] with new low-carbon technologies may become too steep to keep ambitious emissions reductions by mid-century within [cost-effective] reach’.¹³⁶ In other words, the earlier low-carbon technologies are implemented, the more experience industries will gain. While those arguments are not invulnerable to criticism, especially since they tend to disregard factors such as capital costs and investment lead-times,¹³⁷ there has been a consensus that in order to drive investments in low-carbon technologies, a credible price signal is necessary.¹³⁸ The level of price required to incentivise investments varies between sectors; however, ‘a price of €10/tCO₂ or lower would mean that the ETS has little impact on investment decisions, and might also lead to a wider lack of confidence in

¹²⁸ Laing and others (n 8) para 3.1.

¹²⁹ Tilford and others (n 32) 20.

¹³⁰ VH Hoffman, ‘EU ETS and Investment Decisions: The Case of the German Electricity Industry’ as cited in Laing and others (n 8) para 3.1.

¹³¹ Laing and others (n 8) para 3.1.

¹³² Martin, Muûls and Wagner (n 32) 48.

¹³³ Egenhofer, Marcu and Georgiev (n 11) 3.

¹³⁴ *ibid* 5.

¹³⁵ Sartor (n 104) 3.

¹³⁶ *ibid* 4.

¹³⁷ *ibid*.

¹³⁸ European Commission (n 72) 9.

EU climate policies'.¹³⁹ Generally, prices below €20 have been accepted as being much too low to incentivise investments.¹⁴⁰

A low carbon price would not, in itself, hinder the reduction of greenhouse gas emissions, as long as the cap were not exceeded.¹⁴¹ However, incentivising investment in low-carbon technology is necessary to keep the EU on the cost-optimal pathway to decarbonising the economy. The lower the EUA prices are, the less incentivised companies will feel to invest therein at present, and thus the further the EU will shift from the cost-effective path to decarbonisation.

C. The Impact of the EU ETS

Despite the number of shortfalls associated with the EU ETS, it can still be viewed as a partial success. It is necessary to examine the effect that the system has had on emissions abatement so far in order to determine whether the 2015 reforms were necessary.

A number of studies have attempted to establish a link between the ETS and emissions abatement, but there are some inherent difficulties in relation thereto. It has been submitted that if emissions are below the cap, this should not be taken as evidence of the success of the system since, especially during 'phase I' and 'phase II', allowance supply was determined by reference to 'business-as-usual' forecasts, which could prove inaccurate and misleading for a variety of reasons.¹⁴² Firstly, it is difficult to make accurate 'business-as-usual' predictions, given that they, 'involve the construction of counterfactual estimates of what emissions would have been in the absence of the ETS (...) taking into account actual economic growth, energy prices, and weather'.¹⁴³ Secondly, the effects of different policy instruments could be difficult to disentangle.¹⁴⁴ Lastly, the information gap and the difficulties of obtaining commercially sensitive information make an assessment of the system's effectiveness even more problematic.¹⁴⁵

'Phase I' saw a 'best estimate' of reductions around 200 MtCO₂ in 2005-2006,¹⁴⁶ and an overall estimate of up to 3 per cent across all sectors and countries for the entire period,¹⁴⁷ regardless of the over-allocation and lack of banking provisions. However, many view 'phase I' as a success since its purpose was to establish a functioning market for carbon emissions, which was achieved:¹⁴⁸ '82 per cent of respondents had traded carbon allowances and overall 40 per cent considered the Scheme had had some or significant impact.'¹⁴⁹

¹³⁹ Grubb (n 12) 15.

¹⁴⁰ Tilford (n 13) 16.

¹⁴¹ Van der Werf and others (n 82) 2.

¹⁴² National Audit Office (n 30) 6.

¹⁴³ Ellerman and Joskov (n 33) 34.

¹⁴⁴ National Audit Office (n 30) 33.

¹⁴⁵ *ibid.*

¹⁴⁶ Laing and others (n 8) 2.1.

¹⁴⁷ Martin, Muûls and Wagner (n 32) 20.

¹⁴⁸ Bell (n 2) 12.

¹⁴⁹ National Audit Office (n 30) 8.

In 'phase II', while emissions rose to a 'business-as-usual' level in 2007, the effect of the economic crisis has been identified as the, 'main driver of the concurrent fall in emissions.'¹⁵⁰ Furthermore, while the 'bulk of abatement' has been in the form of fuel switching, which could be attributed to the effect of the ETS,¹⁵¹ oil price shocks could have had such a profound effect on companies' behaviour as to render business-as-usual projections unreliable.¹⁵²

One of the system's objectives, namely to reduce the investment in carbon-intensive assets, has been achieved, even if there are doubts as to how much of this success can be exclusively attributed to the implementation of the system.¹⁵³ Nevertheless, only a small number of surveyed firms considered the ETS to have had a significant impact on their emissions.¹⁵⁴ Furthermore, where a company's emissions had decreased, a number of other factors were listed as having a greater impact on the board's decisions.¹⁵⁵

Although it has been considered not possible to attribute falls in emissions to the ETS,¹⁵⁶ the system has had the effect of changing company behaviour. A 2011 study of six countries found that a large number of the eight hundred firms considered were pursuing some measures to reduce their emissions and were investing in research and development aimed at curbing emissions or achieving energy efficiency.¹⁵⁷ The study concludes that there has been some impact from the ETS on investment and innovation; however, it also notes that, 'there is no strong evidence that ETS firms in general differ in their innovativeness from non-ETS firms', finding certainty rather than price to be more relevant for research and development decisions.¹⁵⁸

In conclusion, the carbon price has been factored into decision-making, but it has been unable to impact long-term capital projects; something which is necessary to effect the transition from a carbon-intensive to a low-carbon economy. This is largely due to oversupply of allowances and a low carbon price.¹⁵⁹

4. The Market Stability Reserve

On 6 October 2015, Decision (EU) 2015/1814 on the establishment and operation of a Market Stability Reserve for the EU ETS was adopted by the European Parliament

¹⁵⁰ Martin, Muûls and Wagner (n 32) 22.

¹⁵¹ AD Ellerman, FJ Convery and C de Perthuis, *Pricing Carbon: The European Union Emissions Trading Scheme* (CUP 2010) as cited in Egenhofer and others (n 5) 9.

¹⁵² Egenhofer and others (n 5) 14.

¹⁵³ K Neuhoff, 'Carbon Pricing for Low Carbon Investment' (Climate Policy Initiative/DIW Berlin 2011) 4 <<http://climatepolicyinitiative.org/wp-content/uploads/2011/12/Carbon-Pricing-Exec-Summary.pdf>> accessed 25 March 2016.

¹⁵⁴ National Audit Office (n 30) 41.

¹⁵⁵ Neuhoff (n 153) 8; National Audit Office (n 30) 42.

¹⁵⁶ Kettner, Kletzan-Slamanig and Köppl (n 105) 15.

¹⁵⁷ Martin, Muûls and Wagner (n 32) 4.

¹⁵⁸ *ibid* 20-21.

¹⁵⁹ Neuhoff (n 153).

and the Council. The Reserve will become operational on 1 January 2019,¹⁶⁰ and it will function as a supply-demand adjusting mechanism. The Reserve seeks to address the current oversupply of allowances and, ‘to improve the system’s resilience to major shocks.’¹⁶¹ The following section examines whether the Reserve will be able properly to address those issues and to deliver a stable and reliable carbon price signal in order to stimulate investment in low-carbon technologies and improve the overall functioning of the system.

A. Is it Necessary to Intervene?

Before examining the contribution that the Reserve can make towards the functioning of the ETS, one must first consider whether structural intervention is justified.

Some commentators have argued that the ETS is a ‘technologically neutral instrument’, which will be successful as long as the cap is observed.¹⁶² A low carbon price is a reflection of the fundamental principle of cost-effectiveness behind the ETS.¹⁶³ The MSR is deemed to ‘interfere with the market’ in an unnecessary way,¹⁶⁴ exacerbating the lack of predictability for market players. Indeed, regulatory uncertainty has been one of the major criticisms associated with the ETS’ inability to drive investment in low-carbon technologies: ‘the speed and scale at which carbon prices can drive the switch in innovation and investment depends on the strength of the price signal created, both in terms of magnitude and long-term credibility.’¹⁶⁵ In general, investors are facing a number of uncertainties regarding such general considerations as energy prices, but also regarding the question whether and, if so, how the ETS could be revised in the future.¹⁶⁶ Uncertainty surrounding the long-term supply of allowances and the design of the system makes it difficult for market participants to form long-term expectations as to price.¹⁶⁷ Commentators have consequently pointed out that one of the aims of any measure that intervenes in the normal functioning of the market should be the avoidance of an ‘unacceptable level of uncertainty’;¹⁶⁸ otherwise the impact of market failures will be exacerbated.¹⁶⁹ The Commission has consistently emphasised the central importance of the system being

¹⁶⁰ Decision (EU) 2015/1814 of the European Parliament and of the Council of 6 October 2015 concerning the establishment and operation of a market stability reserve for the EU greenhouse gas emissions trading scheme and amending Directive 2003/87/EC, art 1(1).

¹⁶¹ European Commission, ‘2030 Climate and Energy Goals for a Competitive, Secure and Low-carbon economy’ Press Release IP/14/54, 2.

¹⁶² Marcu (n 103) 2.

¹⁶³ Verdonk (n 38) 20.

¹⁶⁴ Grubb (n 12) 32.

¹⁶⁵ Laing and others (n 8) para 3.1.

¹⁶⁶ Verdonk and others (n 38) 20.

¹⁶⁷ Sartor (n 104) 2.

¹⁶⁸ *ibid* 6.

¹⁶⁹ K Neuhoff and others, ‘Is a Market Stability Reserve Likely to Improve the Functioning of the EU ETS? Evidence from a Model Comparison Exercise’ (Climate Strategies 2015) iii <<http://climatestrategies.org/wp-content/uploads/2015/02/Climate-Strategies-MSR-Report-Final1.pdf>> accessed 25 March 2016.

cost-efficient.¹⁷⁰ Current EUA prices are much lower than the level required to incentivise fuel switching.¹⁷¹ Firms have been focusing primarily on short-term abatement with the carbon market failing to influence long-term investment decisions.¹⁷² From the above it follows that the need to intervene outweighs the attendant perils.

Furthermore, as this article already highlighted above, the ETS is not robust; it does not have, 'the ability to respond to persistent changes in economic circumstances, technological development and overlapping policies', owing to the inelastic features of the supply side.¹⁷³ It has been said that, 'An emissions market requires both scarcity of emission allowances to create the price signal, and also long-term clarity and predictability of rules and targets.'¹⁷⁴ The economic crisis highlighted this shortfall perfectly by demonstrating the effect that 'rapid and dramatic changes in economic output' can have on the price of carbon.¹⁷⁵ Arguably, the MSR would provide the flexibility needed to respond to unforeseen events by revising and adjusting the EUA supply correspondingly.¹⁷⁶

B. Purpose and Functioning of the Reserve

The MSR will become operational in January 2019.¹⁷⁷ The Reserve will automatically regulate the supply of allowances each year by removing 12 per cent of the number of allowances in circulation for a period of 12 months beginning 1 September of that year with a minimum of 100 million allowances,¹⁷⁸ if the pre-defined 833 million EUAs threshold is exceeded.¹⁷⁹ It will re-inject 100 million EUAs, or all EUAs if there are fewer than 100 million EUAs in the Reserve, when the total number of allowances in circulation falls below 400 million.¹⁸⁰ Furthermore, 100 million allowances shall be released when the average allowance price for the period of six

¹⁷⁰ European Commission (n 72) 9.

¹⁷¹ A Gilbert and others, 'Assessing Design Options for a Market Stability Reserve in the EU ETS' (ECOFYS 2014) 4 <www.ecofys.com/files/files/ecofys-2015-assessing-design-options-for-market-stability-reserve-euets.pdf> accessed 25 March 2016.

¹⁷² *ibid* 6.

¹⁷³ *ibid* 5-6.

¹⁷⁴ International Emissions Trading Association, 'IETA's Reaction to the Proposed Market Stability Reserve (MSR)' (IETA 2014) 1 <https://ieta.memberclicks.net/assets/PositionPapers/ieta_position_paper_msr_13.06.2014.pdf> accessed 25 March 2016.

¹⁷⁵ Egenhofer and others (n 5) 23.

¹⁷⁶ W Acworth, 'Can the Market Stability Reserve Stabilise the EU ETS: Commentators Hedge Their Bets' (Department of Climate Policy at DIW Berlin 2014) 3 <www.diw.de/documents/publikationen/73/diw_01.c.465929.de/diw_roundup_23_en.pdf> accessed 25 March 2016.

¹⁷⁷ Decision (EU) 2015/1814, art 1(1).

¹⁷⁸ *ibid* art 1(5).

¹⁷⁹ European Commission, 'Proposal for a Decision of the European Parliament and the Council concerning the establishment and operation of a market stability reserve for the EU's greenhouse gas emission trading scheme and amending Directive 2003/87/EC' COM (2014) 20/2, 3.

¹⁸⁰ Decision (EU) 2015/1814, art 1(6).

consecutive months is over three times more than the average price of EUAs in the two preceding years pursuant to Article 29a.¹⁸¹

Each year, the Commission shall publish the total number of allowances in circulation by 15 May of the subsequent year. This shall include the total sum of all allowances issued and used project credits minus verified emissions.¹⁸² The adjustment of allowance supply should happen, 'without undue delay following the publication of the total number of allowances in circulation[.] The adjustment (...) should be spread over a period of 12 months following a change to the relevant auctioning calendar.'¹⁸³

In order to address the large surplus of allowances at the end of 'phase III', 900 million allowances are to be removed from auctioning and placed in the Reserve.¹⁸⁴ Furthermore, EUAs not allocated to installation because they have ceased operation,¹⁸⁵ shall also be placed in the Reserve.¹⁸⁶

The Commission will monitor the performance of the reserve, and within three years from the start of its operation, and once every five years thereafter, will analyse the Reserve and submit a proposal if necessary.¹⁸⁷

As mentioned already, one problem associated with the EU ETS is its lack of robustness and the fact that it is not sufficiently 'flexible' to, 'adapt to unforeseen events.'¹⁸⁸ The main purpose of the Reserve is to make the supply of allowances more flexible and sensitive to demand,¹⁸⁹ thus making the ETS more resilient to future supply-demand shocks¹⁹⁰ and unforeseen developments.¹⁹¹ It aims to address the inflexibility of supply in the short term without affecting the total long-term supply in place.¹⁹² The MSR should, 'ease costs for businesses when times are hard (...) but require that businesses pay more for carbon when times are good'.¹⁹³ The Reserve has no effect on the overall cap; thus robustness is the driving principle objective behind the MSR.¹⁹⁴

The EU has recognised the large surplus of allowances, which, if no action is taken, is expected to reach 2.6 billion EUAs by 2020.¹⁹⁵ It has been stated that:

While the environmental objective is guaranteed by the cap, the presence of a large surplus reduces the incentives for low-carbon investment and thereby negatively affects the cost-efficiency of the system.¹⁹⁶

¹⁸¹ *ibid* art 1(7).

¹⁸² *ibid* art 1(4).

¹⁸³ *ibid* art 1(6).

¹⁸⁴ *ibid* art 1(2).

¹⁸⁵ Directive 2003/87/EC, art 10a(19) and 10a(20), as amended by Directive 2009/29/EC.

¹⁸⁶ Decision (EU) 2015/1814, art 1(3).

¹⁸⁷ *ibid* art 3.

¹⁸⁸ Verdonk and others (n 38) 53.

¹⁸⁹ European Commission (n 179) para 2.

¹⁹⁰ *ibid* para 3.

¹⁹¹ Verdonk and others (n 38) 20.

¹⁹² European Commission (n 72) 14.

¹⁹³ Sartor (n 104) 5.

¹⁹⁴ Acworth (n 176) 3-4.

¹⁹⁵ G Erbach, 'Reform of the EU Carbon Market: From Backloading to the Market Stability Reserve' (PE 538.951, European Parliamentary Research Service 2014) 5.

The MSR would ensure a stable and sufficiently high carbon price in order to incentivise investments in low-carbon technologies. Furthermore, 'a flatter carbon price growth trajectory' would allow market participants to form price expectations and thus make long-term investment decisions.¹⁹⁷ The Reserve would, 'address the concerns that investment decisions were being made against the background of an oversupply of allowances, resulting in a less than economically efficient way of reaching the ambitious mid-to-long term EU greenhouse gas reduction objectives.'¹⁹⁸ An appropriate carbon price signal is necessary in order to reach the EU's long-term decarbonisation target in a cost-effective manner.¹⁹⁹ Periodic instability combined with low EUA prices have been highlighted as the reasons for the ETS's failure to incentivise low-carbon investment, potentially resulting in a carbon lock-in.²⁰⁰ Furthermore, the Reserve will reduce the uncertainty that permeates international negotiations.²⁰¹ It will send a clear political message about the EU's commitment to reducing greenhouse gas emissions, showing that the ETS is the primary tool for achieving that aim in a cost-efficient way.²⁰²

C. Is the MSR Fit for Purpose?

There are at least four indicators that could determine whether the MSR's design will be effective: (1) it must ensure cost-efficient emissions reduction; (2) it must create a stable price and price trajectory consistent with the expectations of policy-makers and market actors; (3) it must restore the ETS to its long-term decarbonising trajectory; (4) it must make the ETS more robust and better able to respond to external shocks.²⁰³ This paper submits that the design of the MSR would make the ETS better suited to respond to new information. It is also argued that, at the same time, the MSR would boost the price of allowances and ensure that participants can form reasonable future expectations and thus make long-term investment decisions. Nevertheless, a number of concerns have been identified in relation to its structure, which will be examined for the remainder of the section.

i. The Triggers

Quantity-Based Triggers

The MSR will be triggered when the number of allowances reaches certain levels. By adopting a predominantly quantity-based mechanism of adjustment, policy-makers have chosen certainty as to the number of EUAs over their price.²⁰⁴ At first sight this does not seem to address the criticism of EUAs' prices being too low to stimulate

¹⁹⁶ European Commission (n 179) para 1.

¹⁹⁷ Neuhoff (n 169) 13.

¹⁹⁸ Marcu (n 103) 2.

¹⁹⁹ European Commission (n 72) 9.

²⁰⁰ Grubb (n 12) 14.

²⁰¹ European Commission (n 72) 15.

²⁰² International Emissions Trading Association (n 174) 5.

²⁰³ Neuhoff and others (n 169) iii-iv.

²⁰⁴ Sartor (n 104) 1.

innovation. Article 29a is also of little help since it is only concerned with prices being too high.²⁰⁵

Nevertheless, ‘volumetric triggers’ serve as an ‘indirect price shaper’: economic theory suggests that a shortage of allowances would drive the price up, and while there are some concerns about their ability, ‘accurately [to] capture changes in external conditions’, they are perceived as purer from a market perspective and necessary for reasons of simplicity.²⁰⁶ A quantity-based mechanism ensures that the ETS remains ‘cap neutral’.²⁰⁷ As the inflexibility of the system to respond to changes in demand has been identified as the ‘root cause’²⁰⁸ of the price collapse, the triggers should be set with reference to the desired target price.²⁰⁹ If the triggers are determined in line with the hedging demands of the industry the Reserve will deliver ‘scarcity pricing’.²¹⁰ Clear indications must be demonstrated as to policy makers’ determination only to tighten carbon regulations in the future; thus investors will be convinced of the merits of investing in low-carbon technologies regardless of the short-term allowance prices.²¹¹ If the rules are sufficiently transparent and clear, participants in the market will be able to form clear expectations about the short-term adjustments.²¹²

Some concerns have been raised regarding the implementation of the Reserve. First, the current triggers have been determined based on hedging assumptions.²¹³ This has been criticised as being ‘accident prone’, as those assumptions could prove to be wrong or might change in the future.²¹⁴ Many study respondents believe the energy companies’ hedging demands to exceed 900 million; thus the upper trigger has been viewed as appropriate. However, uncertainty exists in relation to the lower threshold.²¹⁵ The need for careful review has been highlighted since the hedging demands of the energy sector are prone to differ over time.²¹⁶ Setting quantity-based triggers inappropriately risks removing too many allowances, leading to, ‘expensive forms of short-term demand reduction’ or the opposite – not removing enough in order to stimulate steady investment.²¹⁷ Underestimating hedging demands and setting triggers too high could result in lower-than-desired prices, while thresholds

²⁰⁵ Decision (EU) 2015/1814, art 1(7).

²⁰⁶ Marcu (n 103) 10.

²⁰⁷ Acworth (n 176) 4.

²⁰⁸ K Kankaanpaa and C Jacazio, ‘Emission Allowance Supply Management in the EU ETS: A Discussion Paper’ (Fortum 2013) 4

<http://fortum.com/SiteCollectionDocuments/Public_affairs/130717%20White%20paper_Automatic%20supply%20adjustment%20mechanism%20in%20the%20EU%20ETS.pdf> accessed 25 March 2016.

²⁰⁹ *ibid* 6.

²¹⁰ Acworth (n 176) 4.

²¹¹ Sartor (n 104) 4.

²¹² Acworth (n 176) 4.

²¹³ *ibid*.

²¹⁴ Sandbag, ‘Structural Reform of the EU ETS: Response to the Commission’s Proposal for a Market Stability Reserve’ (Sandbag 2014) 7

<https://sandbag.org.uk/site_media/pdfs/reports/Policy_Briefing_on_the_Market_Stability_Reserve.pdf> accessed 25 March 2016.

²¹⁵ Gilbert and others (n 171) 56.

²¹⁶ International Emissions Trading Association (n 174) 8.

²¹⁷ Sandbag (n 214) 7.

that are too low may result in an, 'unexpectedly high upward price spike in the case of rigid hedging/banking demand'.²¹⁸ Moreover, it has been argued that an upper limit should be placed on the amount of allowances that could be held in the Reserve in order to prevent the weakening of future climate action.²¹⁹ That proposal can be rejected since that would mean that the Reserve would have an effect on the overall cap and it could reduce its ability to respond to future demand spikes.

Some concerns have been voiced over the speed at which the MSR would be able to rectify the current over-allocation. A higher rate of extraction would return the scarcity faster.²²⁰ Many studies have suggested that the 12 per cent removal rate per annum would mean that balance would not be restored to the market for the continuation of 'phase IV'.²²¹ Therefore, commentators have called for a reduction in the sensitivity of the triggers by making them, '33% of the difference between the supply and the nearest supply trigger', which would result in larger adjustments the further the supply moves from the threshold.²²² Hence, if the total surplus of allowances reached 2 billion, the MSR would remove only 240 million while the 33 per cent factor would result in over 385 million EUAs being placed in the Reserve. Such was the French proposal, which suggested a 33 per cent difference adjustment, combined with an extension of the corridor - 800 million to 1.3 billion.²²³ This would allow for a, 'more gradual approach towards the threshold'.²²⁴ The present paper supports this proposal since the 12 per cent rate, 'risks taking unnecessarily long to get the supply down'.²²⁵ Furthermore, once the triggers are reached a higher percentage would ensure that emissions stay closer to the desired corridor.

Another suggestion is to allow the triggers to decline predictably over time, thus adapting to lower hedging demands by the energy sector,²²⁶ which is expected to adopt more low-carbon technologies.²²⁷ This paper supports the adoption of a lower percentage from the start of the MSR, which would avoid drastic interventions. However, a higher percentage, for example 50 per cent, should be set

²¹⁸ Deutsche Emissionshandelsstelle 'Strengthening Emissions Trading: Discussion Paper on Designing the Market Stability reserve (MSR)' (German Emissions Trading Authority DEHSt 2014) 13 <www.dehst.de/SharedDocs/Downloads/EN/Publications/MSR-Paper.pdf?__blob=publicationFile> accessed 25 March 2016.

²¹⁹ Climate Action Network Europe and others, 'CAN Europe's Position on the Market Stability Reserve' (Climate Action Network Europe 2014) 3 <www.caneurope.org/docman/position-papers-and-research/eu-ets-2/2486-can-europe-s-position-on-the-market-stability-reserve/file> accessed 25 March 2016.

²²⁰ European Federation of Energy Traders, 'EFET Calls for Early Implementation of the EC Proposal for a Market Stability Reserve to Help Restore Confidence in the EU ETS' (2014) <http://efet.org/Cms_Data/Contents/EFET/Folders/Documents/PressRoom/PressStatements/2006Today/~contents/PFJ26DQBTM7YW3R4/EFET_PR_87_14.pdf> accessed 25 March 2016.

²²¹ Gilbert and others (n 171) 57.

²²² Sandbag (n 214) 8.

²²³ Neuhoff and others (n 169) 23.

²²⁴ Deutsche Emissionshandelsstelle (n 218) 16.

²²⁵ D Morris, 'Slaying the Dragon - Vanquish the Surplus and Rescue the ETS. The Environmental Outlook for the EU Emissions Trading Scheme' (Sandbag 2014) 69 <https://sandbag.org.uk/site_media/pdfs/reports/Sandbag-ETS2014-SlayingTheDragon.pdf> accessed 25 March 2016.

²²⁶ Sandbag (n 214) 7-10.

²²⁷ Climate Action Network Europe and others (n 219) 2.

as the quantity of EUAs in circulation approach the upper threshold since this would ensure better confinement within the identified desired corridor.²²⁸ Alternatively, some commentators have suggested using a percentage of the cap as a trigger since that would allow it to change in tandem with changes to the cap and hedging demands.²²⁹ This may result in a more flexible system, better equipped to respond to changes in demand, and thus lead to a more stable price signal.

Lastly, the injection of 100 million EUAs into the market has been criticised for being too inflexible, and calls have been made to adopt symmetrical criteria for return and withdrawal of allowances to and from the market.²³⁰ A more symmetrical approach is favoured by market participants.²³¹ A fall in EUAs below the lower threshold would result in a stronger reaction by market actors. Consequently, the re-injection quantity should not be set in ‘absolute terms’:²³² ‘[w]hile setting an absolute intervention level provides predictability to the market, a percentage would allow the corrective action to reflect the scale of imbalance.’²³³ A ‘discontinuity problem’ is raised by the requirement that a minimum of 100 million allowances be removed, especially in situations where the EUAs in circulation are even just a little over the threshold. This could be problematic in the long-term as it is regarded as a ‘relatively strong intervention in the market’.²³⁴

The MSR has been praised for being transparent and predictable and thus for respecting the ‘principle of automaticity’.²³⁵ The risk of market manipulation by companies is lower since, in order to trigger the release of EUAs, an increase of emissions is required and thus companies will use their own allowances.²³⁶ Furthermore, by boosting current prices the Reserve can prevent companies from delaying abatement efforts for future years, which would have led to higher prices in the future.²³⁷ Therefore, it is likely that the Reserve will guide the market towards the most cost-optimal pathway.²³⁸ It has been remarked that:

[A] quantity based MSR (...) provides additional flexibility through publicly banking surplus allowances, thus incentivizing higher abatement levels early on and shifting firms’ abatement profiles closer to the optimal abatement cost pathway.²³⁹

²²⁸ Deutsche Emissionshandelsstelle (n 218) 19.

²²⁹ Gilbert and others (n 171) 57.

²³⁰ International Emissions Trading Association (n 174) 8.

²³¹ *ibid* 7.

²³² Deutsche Emissionshandelsstelle (n 218) 17-18.

²³³ Gilbert and others (n 171) 57.

²³⁴ Deutsche Emissionshandelsstelle (n 218) 14.

²³⁵ International Emissions Trading Association (n 174) 2.

²³⁶ Gilbert and others (n 171) 35.

²³⁷ A Schopp and others, ‘Modelling a Market Stability Reserve in Carbon Markets’ (Deutsches Institut für Wirtschaftsforschung 2015) 12
<www.diw.de/documents/publikationen/73/diw_01.c.505253.de/dp1483.pdf> accessed 25 March 2016.

²³⁸ Gilbert and others (n 171) xiii.

²³⁹ Schopp and others (n 237) 3.

To summarise, this paper calls for a higher percentage to be adopted for both withdrawal and reinjection of EUAs in order to bring scarcity to the market faster and to allow for that percentage to grow to respond to the ebb and flow of behaviour, hedging demands, and deployment of new technologies.

Article 29A and Price-Based Triggers

If the price is regarded as one of the main concerns driving the structural reform, a price-based trigger would seem to be a better option since this would reduce price volatility and limit uncertainty for investors:²⁴⁰ '[a] price trend over a given period would be the simplest, most transparent and least easily manipulated trigger.'²⁴¹ It would reduce the incentive for firms to distort investment decisions in order to trigger the subsequent release or withdrawal of EUAs.²⁴² Some authors have suggested the adoption of an article similar to Article 29A in relation to withdrawal of allowances based on price trends that are lower than expected.²⁴³ However, one has to bear in mind that this would result in higher complexity; it would upset the nature of the ETS; and it would reduce flexibility for market participants.

Article 29A's main function is to provide a safeguard against the possibility of volume-triggers being set inappropriately.²⁴⁴ It aims at addressing, 'abrupt price strikes[,] without being specific on the cause of the price movement' but it does not respond to depressed prices, for which it has been criticised.²⁴⁵ It provides a price element which upsets the nature of the ETS as a quantity mechanism and could cause firms to deviate from cost-optimal pathways.²⁴⁶ It is seen as, 'too complex' and is considered to add, 'uncertainty to the system' as it could be, 'heavily influenced by financial players'.²⁴⁷ Moreover, the one-sided adjustment in response to higher prices and its effect on price after the injection of 100 million EUAs back into the market remains unclear.²⁴⁸ There is no justifiable objection to the introduction of a low-price trend trigger,²⁴⁹ which would serve as a 'secondary check'.²⁵⁰ However, a number of difficulties can be identified in relation to creating a price-trend corridor which fall outwith the scope of the current article.

The proposal for reforming the ETS into a hybrid mechanism, influenced by both price and quantity, was not supported during stakeholder consultations.²⁵¹ Price-based triggers would raise the issue of effectively determining the desired 'price'.²⁵² Reduced price volatility comes at the cost of an increase in emissions volatility.²⁵³ A supply-based trigger, 'seems to be a sensible political compromise';²⁵⁴

²⁴⁰ Gilbert and others (n 171) 32.

²⁴¹ Acworth (n 176) 4.

²⁴² Taschini (n 117) 14.

²⁴³ *ibid* 11.

²⁴⁴ European Commission (n 72) 21.

²⁴⁵ Marcu (n 103) 8.

²⁴⁶ Gilbert and others (n 171) 42.

²⁴⁷ *ibid* 43.

²⁴⁸ International Emissions Trading Association (n 174) 9.

²⁴⁹ European Commission (n 72) 22.

²⁵⁰ Taschini (n 117) 11-14.

²⁵¹ Van der Werf and others (n 82) 8.

²⁵² Sartor (n 104) 5.

²⁵³ Gilbert and others (n 171) 11.

however, to adopt a price-based mechanism would also mean changing the quantity-based market nature of the ETS.²⁵⁵ Therefore, implementing volume-based triggers respects the, ‘market-based nature of the EU ETS and ensures continuity of a system that allows price discovery’.²⁵⁶ Thus, the application of Article 29A has been criticised as confusing the objective and nature of the ETS and the MSR.²⁵⁷ Lastly, quantity-based triggers would improve the ETS functioning across key performance criteria by pushing it closer to the efficient abatement pathways.²⁵⁸ Therefore, the decision to adopt a primarily volume-based trigger is the most appropriate method of adjusting the supply of allowances to the carbon market.

ii. The Back-Loaded Allowances

Although permanent cancellation of the back-loaded allowances would reduce the flexibility of the Reserve and possibly burden European companies in periods of economic growth, thus increasing the risk of carbon leakage, it would nevertheless help achieve the identified desired level of allowances quicker and return the EU to the cost-efficient decarbonisation path.

Back-loading the 900 million allowances²⁵⁹ into the Reserve is necessary in order to address the current over-supply and the criticism that the 12 per cent extraction rate is insufficient to restore scarcity in a timely manner.²⁶⁰ Nevertheless, back-loading alone would be insufficient to address the accumulated allowance quantity. Additional measures, such as raising the linear reduction factor to 2.2 per cent, have been proposed to tackle the over-allocation issue further.²⁶¹ The supply of allowances for the period 2008-2020 has been, ‘set on the assumption of much better economic circumstance[s]’.²⁶² Placing the 900 million allowances in the MSR has been criticised as having little effect on mid-to-long-term price expectations as market actors would predict their future reinjection.²⁶³ Therefore, some commentators have called for the cancellation of the 900 million EUAs.²⁶⁴

On the one hand, cancellation could reduce the Reserve’s robustness and ability to address shifts in demand since, in times of major economic growth, when the need for CO₂ allowances increases together with emissions, there will be a need for, ‘sufficient liquidity for an efficient market’.²⁶⁵ Thus the Reserve should hold a sufficient amount of EUAs to meet demand: ‘the carbon market must function in line with the principles of sustainable development, in that the environmental and

²⁵⁴ *ibid* xiii.

²⁵⁵ Kankaanpaa and Jacazio (n 208) 6.

²⁵⁶ International Emissions Trading Association (n 174) 2.

²⁵⁷ *ibid* 9.

²⁵⁸ Neuhoff and others (n 169) 26.

²⁵⁹ Decision No 1359/2013/EU of the European Parliament and of the Council of 17 December 2013 amending Directive 2003/87/EC clarifying provisions on the timing of auctions of greenhouse gas allowances [2013] OJ L 343/1.

²⁶⁰ Deutsche Emissionshandelsstelle (n 218) 10.

²⁶¹ European Commission (n 179) art 1(3).

²⁶² Sartor (n 104) 2.

²⁶³ International Emissions Trading Association (n 174) 4.

²⁶⁴ Sandbag (n 214) 5.

²⁶⁵ Kankaanpaa and Jacazio (n 208) 6.

economic aspects must be balanced'.²⁶⁶ An argument exists for allowing flexibility for the Reserve to adjust supply levels to novel circumstances and to ensure the proper functioning of the Reserve. Indeed, room should be left to ease businesses in decoupling economic growth from carbon intensity. On the other hand, cancellation would directly address the oversupply problem, which it perceives as the core issue associated with the ETS,²⁶⁷ and would restore the balance as intended at the time of adopting the cap.²⁶⁸ It could rectify the 'political mistakes of the past' and prevent the watering down of future targets occasioned by the return of those allowances to the market.²⁶⁹ Clearly, cancellation would lead to decarbonisation and abatement in a manner that is cost-efficient. The best way of addressing the evident oversupply is to cancel some of the excess amount. This would also showcase the EU's determination to decarbonise the economy. While carbon leakage might be an issue, it could be addressed through other means,²⁷⁰ the details of which fall outwith the scope of the current article.

iii. Early Start and Reaction Times

Two of the most important changes introduced by the MSR decision compared with the Commission's proposal concerned the earlier implementation of the reserve and quicker response times for supply adjustment.

The majority of commentators have pointed to the importance of implementing the MSR as soon as possible in order to achieve an, 'earlier restoration of the effectiveness of the EU ETS'.²⁷¹ This would ensure that the, 'surplus reaches the hedging corridor earlier', thus increasing early abatement efforts.²⁷² Some commentators have called for the MSR to become operational as soon as 2016.²⁷³ Earlier implementation would reduce later abatement costs since the increase in experience would reduce the mitigation costs.²⁷⁴ An earlier start would ensure a price signal more in line with the supply scarcity that a natural market would provide; however, it may affect regulatory confidence as it would have an impact within 'phase III' and would be inadvisable with regard to carbon leakage uncertainty.²⁷⁵ Consequently, the chosen date seems adequate.

²⁶⁶ Marcu (n 103) 3.

²⁶⁷ Acworth (n 176) 5.

²⁶⁸ Grubb (n 12) 6.

²⁶⁹ Deutsche Emissionshandelsstelle (n 218) 19.

²⁷⁰ G Erbach, 'Post-2020 Reform of the EU Emissions Trading System' (PE 568.334, European Parliamentary Research Service 2015); see also A Marcu and others, 'Carbon Leakage: Options for the EU' (Centre for European Policy Studies 2014)

<www.ceps.eu/system/files/CEPS%20Special%20Report%20No%2083%20Carbon%20Leakage%20Options.pdf> accessed 25 March 2016; A Marcu and others, 'Carbon Leakage: An Overview' (Centre for European Policy Studies 2013)

<http://aei.pitt.edu/46163/1/Special_Report_No_79_Carbon_Leakage.pdf> accessed 25 March 2016.

²⁷¹ International Emissions Trading Association (n 174) 6.

²⁷² Schopp and others (n 237) 14.

²⁷³ Sandbag (n 214) 6.

²⁷⁴ Schopp and others (n 237) 18.

²⁷⁵ Marcu (n 103) 6.

The Commission's initial proposal suggested a time lag of two years between the year when the allowance circulation is estimated and the adjustment year.²⁷⁶ However, Decision 2015/1814 provides for adjustment 'without undue delay' after the publication of allowances in circulation in May the following year.²⁷⁷ This would improve the effectiveness of the MSR and would reduce price volatility,²⁷⁸ since, the longer the period allowed to elapse between announcement of plans to intervene and actual intervention, the greater the level of volatility that will be experienced: '[i]ncreasing the response rate as well as the speed of the reserve to react to unforeseen shocks within one instead of two years can increase robustness and dynamic efficiency.'²⁷⁹ Nevertheless, sellers would want to sell before the additional allowances were reinjected, while buyers would prefer to wait: 'since the injection would still go ahead independently of how the market responds to the announcement (...) the supply in the market would increase even further'.²⁸⁰ The surplus of allowances can be determined only once a year and thus reaction times have been criticised for running the risk of intervening at a time when the market has balanced itself.²⁸¹

The MSR is deemed to provide 'confidence and predictability', which allows market participants to anticipate supply adjustments.²⁸² Nevertheless, the scheduled review every five years and potential changes could lead to additional uncertainty and have an impact on the long-term credibility of the Reserve.²⁸³ Therefore, the review criteria should be published as soon as possible in order for the parameters to be monitored systematically over the time period.²⁸⁴

Although some concerns have been identified, the MSR is believed to increase the robustness of the system to respond to external shocks.²⁸⁵ It will allow for a continued price discovery by the market and will make the ETS able to respond to the success of energy efficiency, renewable energy, and other policies.²⁸⁶ The reform should thus be embraced, albeit with an emphasis on the importance of effective monitoring in order both to determine the MSR's effectiveness and to respond adequately to changes in knowledge, technology and participants' behaviour.

5. Conclusion

The EU ETS effectively managed to put a price on carbon, which led to a reduction in emissions and incentivised investment in low-carbon technologies. Nevertheless, the system has been characterised by an over-allocation of allowances, estimated to

²⁷⁶ European Commission (n 179) art 1(3).

²⁷⁷ Decision (EU) 2015/1814, art 1(6).

²⁷⁸ Acworth (n 176) 4.

²⁷⁹ Schopp and others (n 237) 3.

²⁸⁰ Gilbert and others (n 171) 32.

²⁸¹ *ibid* 36.

²⁸² Kankaanpaa and Jacazio (n 208) 8.

²⁸³ Acworth (n 176) 5.

²⁸⁴ International Emissions Trading Association (n 174) 8.

²⁸⁵ European Commission (n 63) 14.

²⁸⁶ *ibid* 15.

number two billion by the beginning of 'phase III'. The economic crisis, the influx of international credits and the success of other policy instruments were amongst the factors deepening the gap between the supply of and demand for allowances. This has been something of a problem: the oversupply led to a fall in allowance prices, which failed to incentivise the switch to low-carbon technologies. This paper has argued that, as a consequence of this fall in prices, intervention was necessary in order to achieve a balance between supply and demand. Creating shortage would result in a higher and more stable price, thus allowing market actors to form long-term investment expectations and returning the EU to the cost-optimal path towards decarbonisation.

The MSR aims to correct the inelasticity of the supply side. By making the system more robust in its response to unanticipated changes, the Reserve can achieve a stable carbon price determined by a shortage of allowances. A high and sufficiently stable price is desirable in order for market participants to be able to make long-term investment decisions. Nevertheless, some concerns have been voiced regarding the structure of the Reserve and whether it will be able properly to address the current problem of over-supply.

Four criteria were identified by which the MSR's effectiveness could be judged: cost-efficient reduction; stable price and trajectory; restoring the EU to its long-term decarbonising trajectory; and improving robustness. This article has demonstrated that the decision to implement the MSR earlier than 2019, as well as a shorter response time to avoid an unreasonable period between identifying the need to intervene and actual intervention, have been praised as positive steps towards price stability. As regards the placement of back-loaded allowances directly into the Reserve, rather than permanently cancelling them, the article has made the following argument: while cancellation would potentially limit the Reserve's ability to respond to fast economic growth in the future, it would also serve to address the problem of over-supply directly and to ensure that the EU decarbonises the economy in a faster way. The use of volumetric triggers as an indirect price shaper is welcomed as it ensures that emissions reduction is cost-effective. The paper has compared volumetric triggers to a price-based threshold, arguing why the former is preferable to the latter, and has questioned the decision to implement a price trend threshold under Article 29a. It is submitted that the thresholds should be made more flexible by allowing them to grow in line with the annual reduction factor, allowing for greater robustness and making them more responsive to changes in energy companies' hedging demands, economic conditions, technology developments and novel information. Furthermore, this paper endorses the adoption of higher percentage reduction rates in order to restore market balance in a timely manner.

In summary, the MSR can be regarded as an appropriate measure to correct the ETS's market failures and to ensure a flatter carbon price growth trajectory for long-term investment certainty. Nevertheless, in so far as the structure aims to facilitate a *cost-efficient* transition to a low-carbon economy, there is certainly room for improvement.

The Lost Symbol: A Semiotic Analysis of the Psychoactive Substances Act 2016

NICHOLAS BURGESS¹

Abstract

Due to the resource-intensive and time-consuming requirement to provide evidence of a drug's harms before it can be controlled under the Misuse of Drugs Act 1971, and the unprecedented rate at which New Psychoactive Substances (NPS), colloquially known as 'legal highs', have been appearing, the Psychoactive Substances Act 2016 has been created to make the sale, production, importation, and, in some cases, possession, of psychoactive substances illegal, subject only to certain exemptions. However, in its attempt to curb the flow of NPS, the present Government seems to have created a law that is ambiguous and contradictory, and which symbolises an uncompromising stance on the prohibition of drugs. By using Bart van Klink's recent work to assist evaluation, this article offers a critical analysis of the semiotics of the new statute, concluding that it is a flawed piece of legislation, even in symbolic terms.

Keywords: Psychoactive Substances Act 2016, New Psychoactive Substances, Drugs, Symbolic Law

1. Introduction

Intuition tells us that products labelled 'plant fertiliser' and 'bath salts' are probably not for human consumption. Even so, people who are prepared to consume products thus labelled should, for the sake of their health, be discouraged from doing so. Over the past decade, a rapidly increasing proportion of the worldwide population has been ingesting products labelled in this way for recreational intoxication, typically buying them online or from high-street retailers.² These products can collectively be referred to as 'New Psychoactive Substances' (NPS), i.e. drugs which are used for their recreational effect and which are not prohibited by any UN Convention or (in the UK) by the Misuse of Drugs Act 1971 (MDA).³ Typically, these substances are designed to mimic the effects of controlled drugs: e.g. 'synthetic cannabinoids' have been developed to be similar in their psychoactive effect to cannabis.⁴ The

¹ The author graduated from the University of Aberdeen with a First-Class LLB (Honours) in 2016.

² European Monitoring Centre for Drugs and Drug Addiction, 'Legal Approaches to Controlling New Psychoactive Substances' (Publications Office of the European Union 2015) 1

<www.emcdda.europa.eu/attachements.cfm/att_211257_EN_EMCDDEA_POD_2013_Controlling%20new%20psychoactive%20substances.pdf> accessed 12 February 2016.

³ United Nations Office on Drugs and Crime, 'New Psychoactive Substances' (UNODC 2013) 2

<www.unodc.org/documents/drugs/printmaterials2013/NPS_leaflet/WDC13_NPS_leaflet_EN_LORES.pdf> accessed 19 January 2016.

⁴ O Bowden-Jones and others, 'One New Drug a Week: Why Novel Psychoactive Substances and Club Drugs Need a Different Response from UK Treatment Providers' (Faculty Report FR/ AP/02, Royal College of Psychiatrists 2014) 7 <www.rcpsych.ac.uk/pdf/FR%20AP%2002_Sept2014.pdf> accessed 27 February 2016.

descriptions attached to these drugs, such as ‘plant fertiliser’, are, of course, entirely false, but this is not to mislead the customer: the purpose of such mislabelling is to allow entrepreneurs to sell drugs with impunity from trading standards and medicine laws.⁵ In just a few years, a very successful, growing and essentially unregulated market has emerged,⁶ but very little is known about the dangers of the substances sold on this market and, in many cases, even what chemicals they contain.⁷ Following a number of deaths attributed to these drugs,⁸ the present Conservative Government has reacted by creating a new framework in order to combat the problem of NPS. This framework is the Psychoactive Substances Act 2016 (hereinafter referred to as ‘the Act’ and/or ‘the PSA’), which received Royal Assent on 28 January 2016 and came into force on 26 May 2016.⁹ In contrast to previous attempts at prohibitive legislation – which allowed everything that was not expressly forbidden¹⁰ – the Act employs a new, blanket ban approach, whereby the production and supply of all psychoactive substances is illegal by default, with exemptions made for certain substances.

Politicians’ desire to be seen as perpetually tough on crime, and the astronomical output of legislation on criminal law in recent times,¹¹ suggests that a key intention behind many modern statutes is to serve as a symbol demonstrating, *inter alia*, that something is being done about the problem. The Expert Panel created to inform the drafting of the Act set out a number of ‘guiding principles’ when evaluating various potential approaches, which (in addition to the professed need for a proportionate, evidence-based response to the problem of NPS, and for a limitation of the involvement of organised crime in the illicit drug market) included numerous references to communicating a message that NPS use is unacceptable.¹² Such statements suggest that the Act is intended to be (at least in part) symbolic. The primary focus of this article is to explore the argument that the PSA serves an important and worthwhile symbolic function. By using Bart van Klink’s useful

⁵ European Monitoring Centre for Drugs and Drug Addiction, ‘Legal Approaches to Controlling New Psychoactive Substances’ (Publications Office of the European Union 2015) 2
<www.emcdda.europa.eu/attachements.cfm/att_211257_EN_EMCCDDA_POD_2013_Controlling%20new%20psychoactive%20substances.pdf> accessed 12 February 2016.

⁶ EMCDDA, ‘New Psychoactive Substances in Europe: An Update from the EU Early Warning System’ (Publications Office of the European Union 2015) 6
<www.emcdda.europa.eu/attachements.cfm/att_235958_EN_TD0415135ENN.pdf> accessed 15 January 2017.

⁷ *ibid* 5.

⁸ 52 deaths in England and Wales in 2012: ‘New Psychoactive Substances Review: Report of the Expert Panel’ (Home Office 2014) 12
<www.gov.uk/government/uploads/system/uploads/attachment_data/file/368583/NPSexpertReviewPanelReport.pdf> accessed 17 March 2016.

⁹ See generally ‘Trade in So-called “Legal Highs” Now Illegal’ (Home Office 2016)
<www.gov.uk/government/news/trade-in-so-called-legal-highs-now-illegal> accessed 24 November 2016.

¹⁰ Misuse of Drugs Act 1971, s 2(1)(a).

¹¹ See generally J Chalmers and F Leverick, ‘Tracking the Creation of Criminal Offences’ [2013] Crim LR 543.

¹² ‘New Psychoactive Substances Review: Report of the Expert Panel’ (Home Office 2014) 5
<www.gov.uk/government/uploads/system/uploads/attachment_data/file/368583/NPSexpertReviewPanelReport.pdf> accessed 17 March 2016.

‘positive’ and ‘negative’ conceptions of symbolic laws as a toolkit for evaluation,¹³ it will be demonstrated that the PSA is a questionable symbolic law, as well as being an inadequately drafted statute as a whole.

2. Is the PSA Symbolic?

Symbolic legislation is characterised by, ‘a layered structure of meaning: on the primary or literal layer of meaning, we find the conceptual content of the substantive provisions (rules of behaviour) and the provisions to secure (...) compliance (...), whereas the secondary or symbolic layer contains immaterial values that are attached to this conceptual content’.¹⁴ There are a number of features common to all symbolic legislation, which are useful in determining whether a particular statute is indeed symbolic. The following section seeks to demonstrate how the PSA conforms to this description.

A. Typical Aspects of Symbolic Laws

According to Bart van Klink, who has recently developed a very useful tool for analysing and evaluating the well-established concept of legislative symbolism, the criteria of obscurity and vagueness are facets of all symbolic laws.¹⁵ This is where the text of the law is incomprehensible and/or ambiguous, presenting serious interpretive challenges both to the public and to those acting in a professional capacity. The Act’s attempt to define a ‘psychoactive substance’ to the express inclusion of NPS and the express exclusion of substances which are objectively legitimate,¹⁶ is vague, obscure and lacking in legal certainty. As previously noted, the Act follows the ‘general prohibition’ or ‘blanket ban’ approach. Section 2 thereof attempts to define a psychoactive substance as:¹⁷

(...) any substance which (...) is capable of producing a psychoactive effect in a person who consumes it [by allowing the substance, or fumes given off by the substance, to enter the person’s body in any way] (...) if, by stimulating or depressing the person’s central nervous system, it affects the person’s mental functioning or emotional state

¹³ See generally B van Klink, ‘Symbolic Legislation: An Essentially Political Concept’ in B van Klink, B van Beers and L Poort (eds), *Symbolic Legislation Theory and New Developments in Biolaw* (Springer 2015).

¹⁴ *ibid* 22.

¹⁵ *ibid*.

¹⁶ By ‘objectively legitimate’ is meant substances which most people would not regard as worthy of criminalisation/categorisation along with the NPS that are the target of the Act, such as nutmeg or ornamental flower seeds. Such substances are discussed in detail below.

¹⁷ Psychoactive Substances Act 2016, s 2.

This, though, is subject to the exemptions contained in schedule 1, which exist for: controlled drugs, medicinal products, ethyl alcohol products, nicotine and tobacco products, caffeine, and food and drink.¹⁸

Significantly, the words 'new' and 'synthetic' are missing from the definition of a psychoactive substance, which would have gone some way to limiting the scope of the ban to the NPS that are the target of the legislation.¹⁹ Instead, a potentially endless list of substances could be prohibited under the Act. For example, the seeds of some species of common ornamental flowers would be covered by the above definition due to the presence of psychoactive substances within them. *Ipomoea Tricolor* (morning glory) seeds contain both ergonovine, an essential medicine used in obstetrics, and ergine, a hallucinogen and precursor to the Class A²⁰ drug LSD.²¹ In short, the definition of a 'psychoactive substance' includes the NPS that are the target of the legislation, but also, arbitrarily, a multitude of legitimate substances, with no guidance on how to distinguish the legal from the illicit.

The exemption for food and drink also creates a degree of uncertainty in the law. While the given definition of food and drink as, 'any substance which (...) is ordinarily consumed as food, and does not contain (...) any psychoactive substance (...) which is not naturally occurring in the substance, and (...) the use of which in or on food is not authorised by an EU instrument',²² is largely unproblematic, the spice *Myristica Fragrans* (nutmeg) presents interpretive difficulties. Nutmeg is commonly used in food, but naturally contains the deleriant Myristicin,²³ and so is sometimes used for its psychoactive, as opposed to its culinary, properties.²⁴ The question arises whether such use is captured by the 'ordinarily consumed' definition. It is difficult to envisage criminal convictions being imposed for the sale of a common spice, but if the nutmeg is sold for recreational, psychoactive use, and if the criminal courts adopt the 'mischief' approach to statutory interpretation – which would militate in favour of conviction, as such behaviour is precisely the 'mischief' that the Act was drafted to remedy – this will at least be a possibility. Thus, due to the numerous issues in defining a psychoactive substance to include NPS, but to exclude legitimate substances such as ornamental flower seeds and nutmeg, the above hypotheticals demonstrate that the PSA fulfils van Klink's criteria of vagueness and obscurity.

Another feature common to all symbolic laws is what Van Klink calls the 'criterion of discrepancy', i.e. where the legislative provisions lack effective enforcement mechanisms.²⁵ The Act does give the police stop and search powers,

¹⁸ *ibid* sch 1.

¹⁹ S Barber, 'The Psychoactive Substances Bill 2015' (House of Commons Briefing Paper CBP 7334, 16 October 2015) 35 <<http://researchbriefings.parliament.uk/ResearchBriefing/Summary/CBP-7334#fullreport>> accessed 10 February 2016.

²⁰ Misuse of Drugs Act 1971, sch 2.

²¹ See generally Erowid, 'Morning Glory Chemistry' (Erowid, 29 April 2000) <www.erowid.org/plants/morning_glory/morning_glory_chemistry.shtml> accessed 17 March 2016.

²² Psychoactive Substances Act 2016, sch 1, para 7.

²³ H Hallström and A Thuvander, 'Toxicological Evaluation of Myristicin' (1997) 5 *Nat Toxins* 186, 186.

²⁴ MK Abernethy and LB Becker, 'Acute Nutmeg Intoxication' (1992) 10 *Am J Emerg Med* 429, 429.

²⁵ B van Klink, 'Symbolic Legislation: An Essentially Political Concept' in B van Klink, B van Beers and L Poort (eds), *Symbolic Legislation Theory and New Developments in Biolaw* (Springer 2015) 24.

and also allows for premises and prohibition notices, whereby the police can require someone to take reasonable steps to prevent prohibited activities occurring at a designated place.²⁶ Also, websites selling NPS can be forced to cease trading under the PSA.²⁷ However, these powers are also found in the Irish equivalent²⁸ of the Act (on which the PSA was largely based) and/or the MDA,²⁹ which have both struggled to eliminate drug production, supply and use. Although the sample size was small (consisting of a mere five hundred interviewees), a European Commission Eurobarometer survey has indicated that NPS use in Ireland increased from 16 per cent to 22 per cent between 2011 and 2014.³⁰ Additionally, the survey showed that NPS use in Ireland is the highest in the EU amongst the 16-24 age group.³¹ Therefore, the enforcement mechanisms in the PSA might not be as effective as intended. Additionally, convictions may be precluded by the various definitional issues in the Act noted above. Indeed, the ambiguity in the PSA might make establishing a prosecution case difficult in many situations: the Government's Advisory Council on the Misuse of Drugs (ACMD) has also warned that, 'psychoactivity cannot be unequivocally proven' without human trials.³² Human trials of this nature would raise serious ethical issues, and would be inconsistent with the intended purpose of the Act, i.e. to protect humans from becoming intoxicated by NPS ingestion.

The fulfilment of all these criteria (obscurity, vagueness and discrepancy) strongly indicates that the Act can, and should, be regarded as an example of symbolic legislation. The 'umbrella' category of symbolic legislation is further divided by van Klink into two sub-categories: 'positive' and 'negative' symbolic laws, both of which are defined and discussed in the next two sections. These are helpful concepts for further analysing and evaluating the PSA.

B. The PSA: A Positive Symbolic Law?

Positive symbolic laws seek to create a 'communicative framework' between the Government and the public, 'in order to de-automatise current patterns of thinking'.³³ In other words, positive symbolic laws seek to encourage the co-production of laws between the public and the state, based on society's widely recognised fundamental values.

If it is possible to frame the PSA in this positive sense, then the argument that the Act serves an important symbolic purpose will be of some weight. This is

²⁶ Psychoactive Substances Act 2016, ss 13, 14 and 36.

²⁷ *ibid* s 5.

²⁸ Criminal Justice (Psychoactive Substances) Act 2010 (IE), ss 7-14.

²⁹ Misuse of Drugs Act 1971, s 23.

³⁰ TNS Political and Social, 'Flash Eurobarometer 401: Young People and Drugs' (European Commission 2014) 8 <http://ec.europa.eu/public_opinion/flash/fl_401_en.pdf> accessed 17 January 2016.

³¹ *ibid* 9.

³² 'Letter from the Advisory Council on the Misuse of Drugs to the Home Secretary' (2 July 2015) 2 <www.gov.uk/government/uploads/system/uploads/attachment_data/file/441400/2-7-15-ACMD_advice_on_PS_Bill.pdf> accessed 16 January 2016.

³³ B van Klink, 'Symbolic Legislation: An Essentially Political Concept' in B van Klink, B van Beers and L Poort (eds), *Symbolic Legislation Theory and New Developments in Biolaw* (Springer 2015) 27.

because positive symbolic laws can play a valuable role in society. They may serve an 'epistemic function', i.e., '[where] a symbolic law offers a vocabulary that affects the way in which legal and political actors perceive reality [as] accessed through the concepts and distinctions provided by the law'.³⁴ A prime example of this is the concept of 'human dignity' found in the UN Universal Declaration on Human Rights.³⁵ The term 'human dignity' is symbolic, as it conforms to the criteria of obscurity and vagueness due to its abstract nature, and also to the criterion of discrepancy as it does not in itself provide a mechanism for enforcement. 'Human dignity' is *positively* symbolic as it offers legal actors (such as judges) an interpretive vocabulary which can be used to influence the reality of what constitutes a human rights violation. This is achieved via a 'communicative framework' which, instead of (in this case, supra-national) law-makers issuing concrete parameters on what might lead to a breach of human rights, allows for a more adaptable law which might better reflect society's expectations of justice. In this way, a symbolic concept can benefit society, which is what proponents of the PSA would hope to achieve.

Another way in which positive symbolic laws can be of benefit is by redistributing status in society. One example might be the UK legislation forbidding employment discrimination on the grounds of race, sex, disability et cetera.³⁶ Enforcing crimes under this legislation is difficult due to the difficulties in proving that the discrimination was a product of racism, sexism, ableism etc.,³⁷ but such laws seek to change attitudes where they cannot offer any immediate solutions. As what follows will demonstrate, the PSA does not fulfil the criteria for a positive symbolic law.

The communication criterion requires that, for a law to be positively symbolic, it must play a central role in the public debate on the matter regulated by the legislation.³⁸ In other words, 'a structured procedure [of public consultation] enhances the chances for a high quality statute'.³⁹ It is submitted that the Act fails to fulfil the communication criterion, as it actually does the opposite, silencing the national debate before it has even had the chance to begin. This is evident in the manner in which the Government created the Act, dismissing the opinions of various organisations, and drafting the provisions before even consulting their own experts, the ACMD.⁴⁰ The ACMD expressed various concerns with the Psychoactive Substances Bill on a number of occasions,⁴¹ which included its definition of

³⁴ *ibid* 25.

³⁵ Universal Declaration on Human Rights (adopted 10 December 1948) UNGA Res 217 A(III) (UDHR) Preamble.

³⁶ E.g. the Equality Act 2010.

³⁷ LH Krieger, 'The Content of Our Categories: A Cognitive Bias Approach to Discrimination and Equal Employment Opportunity' (1995) 47 *Stan L Rev* 1161, 1163.

³⁸ B van Klink, 'Symbolic Legislation: An Essentially Political Concept' in B van Klink, B van Beers and L Poort (eds), *Symbolic Legislation Theory and New Developments in Biolaw* (Springer 2015) 24.

³⁹ A Siehr, 'Symbolic Legislation and the Need for Legislative Jurisprudence: The Example of the Federal Republic of Germany' (2008) 2 *Legisprudence* 271, 277.

⁴⁰ Release and Transform, 'Parliamentary Briefing: Psychoactive Substances Bill (2nd Reading House of Lords)' (9 June 2015) 1 <www.tdpf.org.uk/sites/default/files/Psychoactive-Substances-Bill-briefing-2015.pdf> accessed 22 March 2016.

⁴¹ See generally L Iversen, 'ACMD's Final Advice on Definitions for Psychoactive Substances Bill' (23 October 2015)

psychoactive substances in lay (rather than scientific) terms, and the consequent lack of a reference to pharmacology or neurochemistry. Such criticisms were met with the reply that many of the ACMD's recommendations simply could not be implemented for the token reason that they were not the Government's intentions.⁴² The recommendations of two pressure groups on drug laws, *Release* and *Transform*, were similarly ignored, as demonstrated by the fact that their concerns regarding legal certainty,⁴³ including the aforementioned 'nutmeg' problem, were not addressed.⁴⁴ Additionally, despite the fact that the World Health Organisation has recommended that people not be criminalised for possession of drugs for personal use,⁴⁵ the PSA makes it illegal to possess a psychoactive substance in a custodial institution,⁴⁶ and also criminalises purchasing NPS for personal use 'through the back door' in cases where the substance was bought from a foreign website.⁴⁷ This is not an illustration of constructive debate, or a 'framework of communication', whereby the views of experts are taken on board. In the face of recent international efforts to end criminal sanctions for drug use, supply and production, such as the decriminalisation of all drugs in Portugal⁴⁸ and the legalised regulation of cannabis in Uruguay and some North American states,⁴⁹ the UK Government's selective approach to the evidence appears all the more striking. Whatever the merits of these alternatives to criminalisation, the fact that they were not even considered in the drafting of the Act is telling in itself. Viewed against this comparative backdrop, the blanket ban does not symbolise the Government engaging in discussion; far from it, it suggests that the Government is acting on a pre-formed view and ignoring any contrary evidence.

Finally, the criterion of 'symbolic working' requires that positively symbolic laws, 'succee[d] gradually in achieving the goals intended through communication and interaction'.⁵⁰ As at the time of writing, the PSA has been in force for little more than half a year; hence concrete statements cannot be made about whether the statute will conform to this criterion. However, based on the fact that people still use drugs despite the MDA having been in operation for over forty-five years, it is at least conceivable that the PSA will fail to serve its symbolic purpose of persuading people not to take drugs. This point is extensively expanded on below in section 3 of this article.

<www.gov.uk/government/uploads/system/uploads/attachment_data/file/470421/ACMD_definitions_advice_final-23-October-2015.pdf> accessed 12 February 2016.

⁴² *ibid* 1.

⁴³ *Release and Transform*, 'Joint Submission to the Public Bill Committee into the Psychoactive Substances Bill' (29 October 2015) 4.

⁴⁴ See above: text to nn 22-24.

⁴⁵ World Health Organisation, *Consolidated Guidelines on HIV Prevention, Diagnosis, Treatment and Care for Key Populations* (2014 WHO) 86-9.

⁴⁶ Psychoactive Substances Act 2016, s 9.

⁴⁷ *ibid* s 8.

⁴⁸ S Jones, *Criminology* (2013 OUP) 14.

⁴⁹ A Travis, 'UN Drugs Body Warns US States and Uruguay Over Cannabis Legalisation' *The Guardian* (London, 3 March 2015) <www.theguardian.com/society/2015/mar/03/un-drugs-body-warns-us-states-and-uruguay-over-cannabis-legalisation> accessed 6 January 2016.

⁵⁰ B van Klink, 'Symbolic Legislation: An Essentially Political Concept' in B van Klink, B van Beers and L Poort (eds), *Symbolic Legislation Theory and New Developments in Biolaw* (Springer 2015) 25.

C. The PSA: A Negative Symbolic Law?

In contrast to positive symbolic laws, negative symbolic laws can be defined as laws whose 'main purpose is to give expression to values in the political sphere'.⁵¹ In this understanding, they refer to instances of the legislature attempting to transmit to the public an uncompromising message about what constitutes legitimate behaviour, based on what politicians themselves subjectively see as right and wrong. It will be demonstrated that it is to this category of symbolic legislation that the PSA is most akin.

These symbolic laws are typically drafted in response to societal circumstances which demand an immediate Government reaction: what van Klink terms 'the crisis criterion'.⁵² The applicability of this criterion to the Act is illustrated by the rapidly increasing prevalence of NPS, both worldwide and in the UK.⁵³ The fear that this increasing prevalence has generated, of unknown drugs with unknown risks being sold to teenagers and others on the high street and the internet, has incentivised the Government to be seen to be acting with swift conviction.

A further aspect of negative symbolic laws is that they are made to, 'enunciate and reinforce what is understood as right and wrong',⁵⁴ and that the relevant legislative actor views, 'the enactment of the law as a moral victory'.⁵⁵ A prohibitive stance on drugs has long been attached to conceptions of moral superiority.⁵⁶ It is also worth noting the pride evident in such Government proclamations as, 'we have already banned 350 substances and have been quicker to respond to this challenge than most other countries'.⁵⁷ Arguably, the current Government views prohibition not only as the correct approach to the NPS problem but also as evidence of some perceived ethical high ground.

In sum, negative symbolic laws can be understood to be a product of the legislature identifying a controversial societal problem and choosing with firm (and possibly misguided) conviction a side of the fence on which to land. For the reasons given immediately above, the PSA clearly conforms to this negative, rather than a positive, conception of symbolic legislation. What exactly the Act symbolises will now be discussed and assessed in order to determine the value, or otherwise, of creating a negative symbolic law in response to the problem of NPS.

⁵¹ *ibid* 22.

⁵² *ibid*.

⁵³ European Monitoring Centre for Drugs and Drug Addiction, 'Legal Approaches to Controlling New Psychoactive Substances' (Publications Office of the European Union 2015) 1 <www.emcdda.europa.eu/attachements.cfm/att_211257_EN_EMCDDEA_POD_2013_Controlling%20new%20psychoactive%20substances.pdf> accessed 12 February 2016.

⁵⁴ K Daly, 'Aims of the Criminal Justice System' in M Marmo, W de Lint and D Palmer (eds), *Crime and Justice: A Guide to Criminology* (4th edn, Lawbook Co 2012) para 17.05.

⁵⁵ B van Klink, 'Symbolic Legislation: An Essentially Political Concept' in B van Klink, B van Beers and L Poort (eds), *Symbolic Legislation Theory and New Developments in Biolaw* (Springer 2015) 22.

⁵⁶ HS Becker, *Outsiders* (The Free Press 1963) 148.

⁵⁷ 'Government Response to the New Psychoactive Substances Review Expert Panel Report' (October 2014) 2

<www.gov.uk/government/uploads/system/uploads/attachment_data/file/368498/ResponseNPS_reviewExpertPanelReport.pdf> accessed 10 January 2016.

3. Evaluating the Negative Symbolism in the PSA: Discouraging Drug Use

Proponents of the Act argued that a blanket ban on all psychoactive substances would symbolise that the production, supply and use of recreational drugs is unacceptable behaviour.⁵⁸ There is a good reason for sending such a message: certain drugs do cause some people serious harm,⁵⁹ and people should be encouraged to pursue more productive activities, hobbies or employment in order to increase their individual and collective quality of life and well-being. What the following section calls into question is not so much the merits of such an ‘anti-drugs’ message, but rather the effectiveness with which that message has been conveyed by the provisions of the PSA.

A. The Blanket Ban is an Unclear Symbol

The objective of the negative symbolism of the PSA, i.e. public recognition of the notion that involvement in recreational drugs in any capacity is wrong, may be frustrated by the sheer vagueness of the symbol used to communicate that message. In other words, the superficial symbolism of a blanket ban loses its value when the breadth of the definition of a ‘psychoactive substance’ under the PSA, and the absence of any distinction between varying levels of harm, is taken into account. Just as person A may conclude that a blanket ban is a clear indication of unacceptable conduct, person B may conclude that, due to the mechanics of the blanket ban, the PSA does not clearly delineate legitimate and illegitimate behaviour: if the recreational use of all psychoactive substances is to be condemned, should person B refrain from purchasing nutmeg or certain species of ornamental flowers, or even from drinking alcohol, a substance which would have fallen within the Act’s definition of a ‘psychoactive substance’ had an exemption not been made?⁶⁰ Many would deem that an absurd suggestion. However, hypotheticals such as these show that the message symbolised by the blanket ban may not be as precise as it appears at first glance. In other words, if someone like person B feels the Act to be internally incoherent, they will fail to grasp the message that the Government is attempting to convey. It could be argued that the public at large is unlikely to be attuned to the definitional inconsistencies contained in the PSA, so will grasp the superficial message, rather than taking person B’s viewpoint. However, this can be questioned for a number of reasons. Firstly, one need look no further than the coverage of the PSA in the UK’s tabloid press, which slated the Act,⁶¹ to show why the public may reach the conclusion of person B. Secondly, this argument does not address the Act’s deficiencies (such as its lack of legal certainty) but rather pretends that they do not

⁵⁸ ‘New Psychoactive Substances Review: Report of the Expert Panel’ (Home Office 2014) 5 <www.gov.uk/government/uploads/system/uploads/attachment_data/file/368583/NPSexpertReviewPanelReport.pdf> accessed 17 March 2016.

⁵⁹ See generally NHS, ‘The Effects of Drugs’ (NHS Choices, 13 May 2015) <www.nhs.uk/Livewell/drugs/Pages/Drugsoverview.aspx> accessed 22 March 2016.

⁶⁰ PSA, s 3 and sch 1, para 3.

⁶¹ See, e.g., D Bloom, ‘Legal Highs Bill: Theresa May is Silencing Her Own Drugs Experts, Warns Former Top Cop’ *The Mirror* (London, 22 June 2015) <www.mirror.co.uk/news/uk-news/legal-highs-Act-theresa-silencing-5929408> accessed 10 March 2016.

exist, which is at best a dubious rationale for creating a law. At worst, it is an indictment of the selective-evidence policy which has been the backbone of the PSA since its inception.

B. The Blanket Ban Symbolises that the Law is Asinine

Law and order considerations also militate against the symbolism of a blanket ban. It has been argued that if society does not demonstrate solidarity in denouncing certain actions as intolerable, the inclination to abide by its rules will decrease.⁶² However, it could be argued that when ineffectual laws such as the MDA are perpetuated and expanded upon – as is the case with the PSA – this symbolises, especially for the target group consisting of those who continually break these laws, that the law is an ‘ass’.⁶³ This is especially pertinent due to the PSA being a negative symbolic law, where the submissions of those who represent this section of the populace have been largely ignored, rather than having been taken on board. If the PSA is to be regarded as a symbol that people should change their behaviour and abstain from involvement in psychoactive substances, then alienating those people who are the target of this Act is counter-productive, and this brings the whole purpose of legislative symbolism into disrepute.

4. Evaluating the Negative Symbolism in the PSA: Putting Government Interests First

A. Government is in Control

In conforming to the crisis criterion, negative symbolic laws are an attempt by the Government to symbolise that the crisis, in this case the proliferation of NPS, is under control. In this regard, the PSA could be seen as beneficial to society as it might communicate that psychoactive substances are not an unregulated ‘free for all’, and this may discourage those who produce and supply them. Additionally, it could convey to those affected by NPS that the Government is taking a proactive approach to resolving the problem. However, this view is potentially misconceived. Arguably, in its symbolic display of control, the Government is misleading the population at large. Firstly, the blanket ban will allow the demand for NPS to be supplied by criminal organisations (who care not for, e.g., imposing age restrictions on their products, and are unable to resort to legitimate legal channels to resolve disputes),⁶⁴ thus leaving the Government unable to control NPS any better than they have managed to control ‘conventional’ recreational drugs under the MDA. Ironically, this is in contrast to the pre-PSA position, whereby tax-paying

⁶² JR Gusfield, ‘On Legislating Morals: The Symbolic Process of Designating Deviance’ (1968) 56 CLR 54, 55.

⁶³ A notion popularised by the novel, *Oliver Twist*. See C Dickens, *Oliver Twist* (1st edn 1837-8, rev edn, Penguin Books 2003) 436: ‘If the law supposes that (...) the law is [an] ass – [an] idiot.’

⁶⁴ P Goldstein, ‘The Drugs/Violence Nexus: A Tripartite Conceptual Framework’ (1985) 39 Journal of Drug Issues 493, 497.

‘headshops’ (high street retailers of NPS and drug paraphernalia) could ensure responsible retail practices by not selling products to minors or intoxicated individuals.⁶⁵ Thus, an unregulated ‘free for all’ is precisely what the PSA will create. Secondly, the weaknesses and inadequacies of the legislation caused by dismissing expert opinions⁶⁶ – and the indications from Ireland that NPS use has increased since the passage of the Irish equivalent of the PSA⁶⁷ – indicate the opposite of a proactive approach.

B. Political Expediency?

Arguably, there are political points to be scored by implementing the PSA other than giving the appearance of control. At the very least, it could be hazarded that the PSA represents the path of least resistance when it comes to tackling the problem of NPS. By alluding to the supposedly devastating health harms, societal problems and immorality of NPS, the Government, ‘reap[s] the political benefits of voting for [health]’, but ‘sidesteps’ difficult policy and regulation questions.⁶⁸ Drug prohibition was a feature of the Conservative Party manifesto in 2015,⁶⁹ and by furthering the prohibition of drugs by introducing a blanket ban on psychoactive substances, rather than trialling radical reforms such as those in Portugal, there was arguably little danger of alienating voters with a traditionalist outlook on the drugs problem: persons who might be referred to broadly as conservatives with a small ‘c’. This move also made particular strategic sense in the 2015 General election. It allowed the Conservative party to distance itself from the Liberal Democrats, who (*inter alia*) adopted a non-prohibitionist stance on drugs near the end of the Conservative/Liberal Democrat coalition Government.⁷⁰ The impression that these circumstances create, namely that the Act was potentially the product of political expediency, further taints its value as a symbol. This is even the case if, contrary to appearances, the Act was motivated purely by a desire to tackle the NPS problem in an effective manner. This article is considering the effectiveness of the NPS in as much as it constitutes a symbol, and a symbol, by definition, is only as effective as what it *appears* to be. However noble or well-meaning the drafters’ actual intentions might have been, the Act *seems* to have been the product of political expediency, if only because of the aforementioned circumstances surrounding its enactment. Thus, as a piece of symbolic legislation purporting to adopt a clear, uncompromising stance on the NPS issue, the PSA is of somewhat questionable worth.

⁶⁵ ‘New Psychoactive Substances Review: Report of the Expert Panel’ (Home Office 2014) 14 <www.gov.uk/government/uploads/system/uploads/attachment_data/file/368583/NPSexpertReviewPanelReport.pdf> accessed 17 March 2016.

⁶⁶ As discussed in Section 2B.

⁶⁷ As discussed in Section 2A.

⁶⁸ JP Dwyer, ‘The Pathology of Symbolic Legislation’ (1990) 17 *Ecology LQ* 233, 233 and 316.

⁶⁹ ‘The Conservative Party Manifesto 2015’ (Conservative Party 2015) 9 <<https://s3-eu-west-1.amazonaws.com/manifesto2015/ConservativeManifesto2015.pdf>> accessed 22 March 2016.

⁷⁰ ‘The Liberal Democrats Manifesto 2015’ (Liberal Democrats 2015) 125 <https://d3n8a8pro7vhmx.cloudfront.net/libdems/pages/8907/attachments/original/1429028133/Liberal_Democrat_General_Election_Manifesto_2015.pdf?1429028133> accessed 22 March 2016.

5. Conclusion

The mere fact that a law is symbolic is not necessarily a criticism, as symbolic laws can and do exist in a positive sense. Thus, the PSA, and the arguments of those who support it on the basis that drug use must be seen to be discouraged, might have had merit despite the fact that it is unlikely to achieve its stated aims of reducing demand for NPS, and limiting the involvement of organised crime in the illicit drug market et cetera. However, in the pursuance of being seen not to condone drug use, the knee-jerk reaction to NPS by the present Government actually sends a negative message. It has been shown why it is unlikely that the PSA will serve as an effective symbol for discouraging drug use, and that instead it symbolises political self-preservation. As other countries start to question the merits of prohibition, the UK Government has chosen to widen the prohibition of drugs, creating a severely deficient law, and ignoring evidence which is inconsistent with their own agenda.

The Classification of Murder and Slaughter in the Justiciary Court from 1625-1650: Malice, Intent and Premeditation - Food 'Forethought'?

STEPHANIE A DROPULJIC¹

Abstract

The prosecution of homicide in Scotland during the early seventeenth century is all too relevant to a wider appreciation of the law, government, legal profession and society of the time. It is, consequently, surprising that this subject has hitherto received little in the way of scholarly attention. By re-examining various historical records (both archival and printed) this article aims to answer important questions about the classification of the law of homicide, with a specific focus on the prosecution of murder and slaughter. It considers the key differences between these crimes with regard to various issues, most notably: (1) the role of 'malice' in the prosecution of slaughter and (2) the term 'forethought' and how that influenced the prosecution of murder.

Keywords: Evidence, Procedure, Seventeenth-Century, Homicide, Scots Criminal Law, Archival Material, Malice, Forethought

1. Introduction

This article explores the classification of homicide in seventeenth century Scotland. It is important first to understand the distinction between murder and slaughter in that period; thereafter, the further classifications within these categories are examined.

In the first half of the seventeenth century, acts of homicide could be placed into one of two broad categories: 'slaughter' and 'murder'. As shown below in Table 1, the overwhelming majority of homicide cases heard before the Justiciary Court from 1625-1650 were for the crime of 'slaughter' (more than 76 per cent).² The next

¹ PhD candidate, University of Aberdeen. This article is part of a thesis that was presented for the degree of Masters of Law (by Research) at the University of Aberdeen in September 2016.

² There are two principal sets of sources from which the Justiciary Court cases were taken. Firstly, there were three volumes contained in the National Records of Scotland: JC2/6 High Court Book of Adjournal - Old Series (7 Oct 1619 - 24 Mar 1631) (hereinafter, 'JC2/6'); JC2/7 High Court Book of Adjournal - Old Series (29 Mar 1631 - 15 Nov 1637) (hereinafter, 'JC2/7'); and JC2/8 High Court Book of Adjournal - Old Series (17 Nov 1637 - 17 Jul 1650) (hereinafter, 'JC2/8'). Secondly, the cases can be found in printed editions of SA Gillon (ed), *Selected Justiciary Cases: 1624-1650, Vol 1* (Stair Society 1953) (hereinafter, 'SJC (Vol 1)'); JI Smith (ed), *Selected Justiciary Cases: 1624-1650, Vol 2* (Stair Society 1972) (hereinafter, 'SJC (Vol 2)'); JI Smith, *Selected Justiciary Cases: 1624-1650, Vol 3* (Stair Society 1974) (hereinafter, 'SJC (Vol 3)'). Given the nature of the archival material it is often difficult to provide pinpoint paragraph references. The references made throughout this article are to the modern pagination provided by the National Records of Scotland, which can be found on the bottom left-hand corner of each page. The manuscripts, at least those contained in the three Justiciary Court volumes used here, do not include paragraph numbers. In some of the 'case law' citations below,

most prevalent accusation to feature in this period was that of 'murder and slaughter', with just over 12 per cent of the cases accounting for that double charge. As explored in greater detail below, this double charge might have been used by pursuers who, despite wishing to libel murder, were not confident of fulfilling the burdens of proof associated with that crime.

Crimes	Number of Cases	Percentage*
'Slaughter'	126	76.8%
'Murder and slaughter'	21	12.8%
'Murder'	14	8.5%
'Murder under trust'	1	0.6%
'Treasonable murder'	1	0.6%
'Accidental slaughter'	1	0.6%
Total Cases	164	100%

~TABLE 1: SHOWING THE NUMBER OF HOMICIDE CASES FROM 1625 - 1650 THAT CAME BEFORE THE JUSTICIARY COURT (*PERCENTAGES ROUNDED TO ONE DECIMAL PLACE).

There is very little secondary literature on the topic of the classification of murder and slaughter in the seventeenth century. This was recognised by David Sellar, who remarked that, apart from writings on the payment of assythment, blood-feuds in early Scottish society and violence more broadly, little has been written about the development of the law of homicide in Scotland.³ His study on 'forethought felony' is the most comprehensive examination of the classification of murder and slaughter, and so is examined here in detail. Sellar suggested that the notion of 'forethought felony' was part of Scots law for more than five hundred years.⁴ He traced the classification of murder starting with the earliest known classification of homicide in Scots law to *Regiam Majestatem*, in the early-fourteenth century. That source specifically associates the term 'murder' with a secret killing.⁵

more than one page number will be mentioned. This has been done because the case in question appears more than once in the volume, the aim being to provide the reader with as much information as possible.

³ D Sellar, 'Forethocht Felony, Malice Aforethought and the Classification of Homicide' in WM Gordon and TD Fergus (eds), *Legal History in the Making: Proceedings of the Ninth British Legal History Conference Glasgow 1989* (Hambledon Press 1991) 47. With regard to this perceived lack of scholarship Sellar cites the following works: R Black, 'Historical Survey of Delictual Liability in Scotland for Personal Injuries and Death - Part I, Early History' (1975) 8 *Comp & Intl LJ S Afr* 46; CHW Gane, 'The Effect of a Pardon in Scots Law' [1980] *Jur Rev* 18; J Wormald, 'Bloodfeud, Kindred and Government in Early Modern Scotland' (1980) 87 *Past and Present* 54; and KM Brown, *Bloodfeud in Scotland, 1573-1625* (John Donald 1986). Little else has been written following these sources on the classification of homicide in the seventeenth century.

⁴ Sellar, 'Forethought Felony' (n 3) 45.

⁵ *ibid* 48. Sellar cites the Stair Society edition of *Regiam Majestatem*: J Skene, *Regiam Majestatem and Quoniam Attachiamenta* (TM Cooper ed, Stair Society 1947) 4,5. The edition used for this article was J Skene, *Regiam Majestatem: The Auld Lawes and Constitutions of Scotland* (Thomas Finlason 1609). See generally 4,5.

The law of homicide was further developed by legislation. A statute of 1370 enacted that the King should not grant a remission for homicide until an inquest had determined whether the killing had been committed, *'per murthyr vel per praecogitatam malitiam'* ('by murder or malicious forethought').⁶ Two years later an Act required that an assize or inquest determine whether the accused had killed another, *'ex certo et deliverato proposito vel per forthouchfelony sive murthir vel ex calore iracundiae viz chaudemellee'* ('according to the certain and deliberated proposition, either by forethought felony or murder, or from heat of passion, namely *chaudemellee*').⁷ The 1372 statute mentioned the assize adjudicating on the possible existence of 'forethocht felony' or 'chaudemella'.⁸ According to Sellar, at least two fifteenth-century Scottish statutes mentioned 'forethocht felony' and contrasted it with actions on a 'suddante' or 'chaudemella'.⁹ By the eighteenth century, the term 'murder' seemed to have moved from its older, restricted meaning of secret killing to cover all killing done with forethought felony.¹⁰ However, Sellar said little of the law in the seventeenth century.

This topic has also been addressed in the work of Alexander Grant, who discusses the classification of the law of homicide with reference to *Regiam Majestatem*,¹¹ as well as Skene's *De Verborum Significatione*,¹² and the importance of the classifications contained in these works to the context of feuding in medieval Scotland.¹³ Grant suggests that secret killing was central to the concept of murder and retained this meaning until at least the end of the fifteenth century.¹⁴ He quotes Skene's definition of murder (or *murthurum*), which is provided in *De Verborum Significatione*:

Whereof of some is called private, that is manslaughter, whereof the author is unknown, whereof the inquisition belongs to the crowner; as where a person is found slain, or drowned, in any place or water. Other is public committed by forethought felony. And murder is committed by forthought felony and not by suddenty.¹⁵

⁶ APS i 509, as cited in Sellar, 'Forethought Felony' (n 3) 48. See also '1370/2/12' in KM Brown and others (eds), 'The Records of the Parliaments of Scotland to 1707' (St Andrews, 2007-2016) (hereinafter, 'RPS') <www.rps.ac.uk/trans/1370/2/12> accessed 17 February 2017. The version of this statute consulted by Sellar seems to differ in its wording from the 'RPS' version thereof.

⁷ APS I, 547-48, as cited in Sellar, 'Forethought Felony' (n 3) 48; this statute can also be found in the Records of the Parliaments of Scotland': RPS 1372/3/6.

⁸ Sellar, 'Forethought Felony' (n 3) 48. See RPS 1372/3/6. This seems to be the only statute in 1372 that refers to murder and *chaudemella*.

⁹ Sellar, 'Forethought Felony' (n 3) 49. In fact, there seem to have been at least three such statutes. See APS II, 9, c 7, RPS 1426/10; APS II, 21, c 7, RPS 1432/3/8; and APS II, 95, c 11, RPS 1469/25.

¹⁰ Sellar, 'Forethought Felony' (n 3) 49.

¹¹ (n 5).

¹² Skene, *De Verborum Significatione: The Exposition of the termes and difficull words* (EG 1641).

¹³ A Grant, 'Murder Will Out: Kingship, Kinship and Killing in Medieval Scotland' in S Boardman and Julian Goodare (eds), *Kings, Lords and Men in Scotland and Britain, 1300-1625: Essays in Honour of Jenny Wormald* (Edinburgh University Press 2014) 222.

¹⁴ *ibid* 224.

¹⁵ See both Skene, *De Verborum Significatione* (n 12) 69 and Grant, 'Murder Will Out' (n 13) 224.

Grant's reading of Skene's definition is that murder came under 'forethought felony' and that it was still 'non-public, with unknown perpetrators', meaning that it had essentially retained its clandestine aspect.¹⁶

The main influence on Scots criminal law was undoubtedly Civilian. This can, for example, be observed in Mackenzie's division of homicide into four categories borrowed from the Civilians: homicide committed casually, in defence, culpably and wilfully.¹⁷ More particularly, according to Sellar, this division could ultimately be traced to the Canonist source, *Decretals*.¹⁸ For instance, the term '*de homicidio voluntario vel casuali*'¹⁹ can be taken to prefigure the distinction, later recognised by Mackenzie,²⁰ between homicide committed casually and homicide committed wilfully.²¹

Sellar argues that the development of homicide and the terms associated therewith, 'points to the consistent and uninterrupted use of the term malice aforethought to describe a premeditated, rather than a merely deliberate homicide, from at least the later fourteenth until the eighteenth century'.²² However, exploration of the Justiciary Court records reveals some, admittedly minor, variations in terminological classification. The term 'aforethought felony' was successfully prosecuted once²³ and mentioned only one other time.²⁴ The term 'malice aforethought' was not in frequent use during the early seventeenth century, which is not to say that Sellar is incorrect; at worst, his observations simply belie the terminology, as against the concepts, used by the courts during the early seventeenth century. The use of these terms is explored more fully below in the section on slaughter. As this article shows, the terminology relating to homicide was not used consistently in practice; rather, the pursuer would allege slaughter, murder or both in reference to particular circumstances from which 'malice' or 'intention'²⁵ on the part of the 'panel' (i.e. the accused) could be inferred. At times, evidence would also be produced in court, to disprove the murder or slaughter charge.²⁶ For instance, the case of *Thomas Crombie and others* (1625) featured the averment of medical evidence. However, reliance on such expert evidence was notable for its relative rarity in the seventeenth century; it seems, in other words, to have been the exception rather than the rule.

¹⁶ *ibid.*

¹⁷ Sellar, 'Forethought Felony' (n 3) 58. Mackenzie's classification of homicide can be found in Sir George Mackenzie: *The Laws and Customs of Scotland in Matters Criminal* (OF Robison ed, Stair Society 2012) 1,11.

¹⁸ *Decretals* (Compiled by St Raymund of Penafort for Pope Gregory IX, 1234) V,xii. This source was discussed in Sellar, 'Forethought Felony' (n 3) 58. It was not directly consulted for the purposes of this article.

¹⁹ This translates roughly as 'homicide wilful or accidental'.

²⁰ Sir George Mackenzie, *Matters Criminal* (n 17) 1,11,6-8.

²¹ Sellar, 'Forethought Felony' (n 3) 58.

²² *ibid* 56.

²³ In the case of *Robert Walker and others* (1642), SJC (Vol 3) 536.

²⁴ In the case of *Johnne Bell* (1644), JC2/8, 374.

²⁵ By 'intent' is meant 'thinking in advance of an act' or 'planning in advance of an act'; it should not be understood in the looser, more modern sense of the word. A further discussion is provided in section 3 below.

²⁶ For example, see the case of *Thomas Crombie and others* (1625), SJC (Vol 1) 18 and JC2/6, 324.

From the foregoing it would appear that, in order truly to understand the law pertaining to homicide in seventeenth century Scotland, it will be necessary to re-examine the classification of murder and slaughter as distinct forms of homicide in the light of the seventeenth century evidence. By reference to the Justiciary Court records, this article examines the crimes of slaughter, murder, and thereafter the double charge of 'murder and slaughter'. To that end, specific reference is made to the procedure of the Justiciary Court and to the pleadings brought before it. The ultimate aim of this historical survey is to furnish a richer understanding of court practice in seventeenth century homicide cases and, more generally, a greater appreciation of the concept of homicide in Scottish legal history.

2. The Crime of Slaughter

A. Overview of Findings

This section explores the prosecution of 'slaughter' before the Justiciary Court from 1625-1650. Re-examining the prosecution of this crime allows for interesting conclusions to be drawn regarding the classification of homicide, for instance the determination that, in contrast to 'murder', 'slaughter' was never libelled on the grounds of express 'intent' or 'deliberate' killing. It is also possible that 'slaughter' was used, in the early seventeenth century, as a general term for homicide. 'Slaughter' was libelled in three principal ways: 'slaughter *chaudemella*', 'slaughter with *precogitat malice*', and 'accidental slaughter'.

As many of the sources considered below make clear, 'slaughter with *precogitat malice*' was the killing of another maliciously by a deadly wound. The malice in these cases concerned the conduct or verbal words exchanged between the 'panel' (i.e. the accused) and the 'defunct' (i.e. the alleged victim) prior to the latter's death. The *precogitata* (i.e. thinking in advance) did not appear to relate to the actual act of killing; rather 'slaughter with *precogitat malice*' seemed to concern the scenario of killing preceded by a more general, non-homicidal sentiment of malice and ill-will. The notion of 'accidental slaughter' can be discerned from the single case so termed,²⁷ and also from the cases in which this term was invoked as a defence against an allegation of 'slaughter with *precogitat malice*'.²⁸ In other words, the panel might plead that, although there had been a homicide, this had been committed without *precogitat malice*. This defence, explored more fully below, was often accepted by the court. The practice before the Justiciary Court indicates that 'slaughter *chaudemella*' was, like 'accidental slaughter', different from 'slaughter with *precogitat malice*', though the confines of that classification are not made explicitly clear. There was specific reference to statutory material, explored in the sub-section

²⁷ *James Mathiesone* (1640), SJC (Vol 2) 395.

²⁸ See for example the cases of: *Johnne Young* (1630) SJC (Vol 2) 313 and JC2/6, 667; *William Jamesoun and another* (1632) SJC (Vol 1) 201 and JC2/7 102 & 103 & 105; and *James Heart* (1637) SJC (Vol 2) 332. The varying page numbers listed for the case of *William Jamesoun* have been provided so that consultation of the case itself is easier.

below, that provides an indication of the meaning of *chaudemella* in the period under consideration.

It is sometimes difficult to pinpoint the meaning of the term 'malice' in this period: the pursuers tended simply to libel slaughter committed with 'malice' without any precise explanation as to the meaning of the latter term and, based on the content of the criminal letters²⁹ and evidence presented before it, the court would conclude whether said 'malice' had existed. This lack of definitional clarity could be taken to suggest that the lawyers of the period had not reached a complete consensus as to the word's meaning.

The present section focuses on homicide in connection with the different classifications that emerge from the Justiciary Court records: 'slaughter *chaudemella*', 'slaughter with *precogitat* malice' and 'accidental slaughter'. Thereafter, the evidential implications of the allegations of slaughter through the nature of the wound are considered. The Justiciary Court records indicate an increasing tendency on the part of pursuers to include 'deadly wound' within the libel or to mention expressly the location of the wound and the effect that it had on the body of the defunct. Prior to the 1630s, libels had not always mentioned 'deadly wound'.³⁰ The subsequent tendency to mention this concept in the 'libel' could be attributed to a change in practice. The need to prove that the wound was deadly, or, in other words, that the defunct had died therefrom, suggests that there existed a requirement of causation in all but name. The notion of 'deadly wound' and the requisite severity of the wound to secure a conviction are explored further below.

B. Malice

The records show that 'malice' had an important connection to the charge of 'slaughter', simply libelled as such, in the Justiciary Court from 1625-1650. The idea of malice in connection to slaughter can be seen in the panel's defences to the crime. It seems that the panel would allege, as a defence to slaughter, that there had been no malice in connection with the homicide committed and therefore that a charge of slaughter could not be sustained. The following section considers the different forms of 'slaughter' in descending order of gravity.

i. Slaughter with Precogitat Malice

It seems that malice was an important element in the successful prosecution of most cases involving a libel of 'slaughter'. In such cases, this was referred to as '*precogitat* malice', which consisted of malice borne by the panel to the defunct prior to the

²⁹ The criminal letters were a central part of Scottish criminal procedure during the seventeenth century. The records suggest that the criminal letters consisted of a formal written accusation drawn up on behalf of the pursuers and served on the panel before the trial. Generally speaking, the criminal letters identified the panel(s) directly by name(s) and made reference to various key details: for instance, where and when the alleged crime took place, the circumstances of the death, and when the defunct died. Sometimes there was specific reference to the 'dittay', which seems to have been the substantial legal point contained within the criminal letters. However, the criminal letters would contain a substantial legal point even where there was no reference to the 'dittay'.

³⁰ For examples of cases that do not mention the concept of a 'deadly wound', see *Robert Buchannane and others* (1626) SJC (Vol 1) 54 and *Robert Anderson* (1627) SJC (Vol 1) 69.

latter's death. This could find expression through, for instance, hateful words. Crucially, however, this did not entail homicidal premeditation. The presence of 'malice' might signify a desire on the panel's part to injure the defunct or to do him or her an ill turn, but it did not, in itself, bespeak a desire to *kill* the defunct.

In 'slaughter' cases (i.e. those cases where the term 'slaughter' was specifically used in court or the criminal letters) the pursuer had to show that some form of malice had occurred in connection with the homicide. This can be seen in the court pleadings of *Thomas Bryce and another* (1639). The accused in that case were charged with the cruel slaughter of the defunct.³¹ Contained within the criminal letters was the allegation that the panels had openly vowed an 'evil turn' on the defunct and to 'wind him ane pirne' him.³² Given the context it seems that 'wind him ane pirne [him]' signified a threat to do the defunct a physical injury. The King's Advocate, who appeared as a pursuer alongside the defunct's mother and sisters, submitted to the court that the malice only applied to Thomas Bryce (the father who was then art and part panel with his son) and not to Robert Liddell (the other co-accused).³³ Liddell and his advocate submitted to the court that since only one strike had been given to the defunct, as alleged in the criminal letters, the charge of slaughter was not relevant to the said Robert Liddell. The defenders and pursuers then debated at great length the law of accomplice, namely whether Liddell's harbouring of Bryce meant that he was a party to the crime and could be charged as such.³⁴ The trial was 'continued'³⁵ and does not reappear in the records for the selected period. However, the averment of circumstances suggestive of '*precogitat malice*' in the criminal letters reflects the possibility that '*precogitat malice*' was, at least as a general rule, a requirement for the successful pursuit of a slaughter charge.

The element of '*precogitat malice*' (i.e. malicious forethought) in connection to slaughter was further developed in the form of a defence invoked by panels in court. This is particularly apparent in the cases of *Johnne Young* (1630)³⁶ and *William Jamesoun and another* (1632).³⁷ In these two cases the only defences provided were that the slaughters had not been committed with '*precogitat malice*'; the panels in both cases were acquitted. These two cases are now explored in detail to identify what, precisely, 'no malice' signified in connection to slaughter from 1625-1650 in the Justiciary Court.

The trial of *Johnne Young* (1630) for the slaughter of Archibald Reid contained a particularly lengthy court discussion about the lack of malice.³⁸ To, 'clear his innocence' the panel presented to the court and the assize a testimonial subscribed by the minister of Curmonnoche and in the name of the elders of the said Kirk and

³¹ SJC (Vol 1) 299.

³² *ibid.*

³³ *ibid* 301.

³⁴ SJC (Vol 1) 299.

³⁵ The 'continuation' of a trial meant that it had been postponed. This was done by the Justice-Depute, usually without giving an explicit reason therefor. One reason for 'continuing' a case was to allow more time for witnesses to be called (*William Anderson* (1641), SJC (Vol 2) 409 and JC2/8, 122). 'Continued' cases could, but did not invariably, reappear at some later date.

³⁶ SJC (Vol 2) 313 and JC2/6, 667.

³⁷ SJC (Vol 1) 201 and JC2/7, 102 & 103 & 105.

³⁸ SJC (Vol 2) 313 and JC2/6, 667.

Session.³⁹ The ministers and elders testified that there was never malice or discord between the panel and the defunct.⁴⁰ The panel had never uttered a word against, or borne any malice towards, the defunct for any cause whatsoever before his death or thereafter.⁴¹ The pursuers objected to the evidence submitted by the ministers and elders.⁴² Even so, the Justice-Depute allowed the evidence and the matter went before the assize.⁴³ The assize returned a verdict of 'innocent' and acquitted Young of the charge of slaughter. It is unclear whether this evidence was the basis on which the assize acquitted the panel, but it could be argued that the absence of malice was a contributing factor to that outcome.

Likewise, in *William Jamesoun and another* (1632), which concerned two men charged with slaughter, the panels used as exculpatory evidence the fact that they had never borne any '*precogitat malice*' towards the defunct or committed a previous felony against him and the fact that the strike inflicted upon the defunct had actually proceeded upon a 'sudden passion and *chaudemella*' by the father of Johnne Andersone.⁴⁴ The case is slightly convoluted as the original entry charged William Jamesoun and Johnne Andersone as 'art and part' of this slaughter. It was then alleged that the father of Johnne Andersone was the one who had committed the strike and that William Jamesoun had only been there to witness the slaughter, which led the court to conclude that William Jamesoun (and also Johnne Andersone) could not have been art and part thereof. Clearly, in this case, 'art and part' liability for slaughter was seen to involve physical participation in the homicide. This seventeenth century understanding can be contrasted with later, more 'modern'⁴⁵ conceptions of the term, according to which 'art and part' liability did not require each party involved to deliver a fatal blow or strike; rather, simply being present whilst the strike or fatal blow was delivered would have sufficed to render those present 'art and part'. The key point to note, for present purposes, is that the assize in *William Jamesoun and another* acquitted William Jamesoun and Johnne Andersone. This could have been on the basis that they did not bear any malice towards the defunct. On the other hand, the verdict may have resulted from the allegation, made after the criminal letters had been drawn up, that the father of Andersone was the one who had committed the crime and that he had not been charged in the criminal letters. Thus, there are at least two possible explanations for the assize's decision. Regardless, it can at least be said that the panel's defence placed great emphasis on the absence of malice. In addition, this case can be seen to reinforce the distinction between '*precogitat malice*', on the one hand, and a sudden passion (or *chaudemella*), on the other. The latter term is discussed below in relation to 'slaughter *chaudemella*'.

³⁹ *ibid.*

⁴⁰ *ibid.*

⁴¹ *ibid.*

⁴² *ibid.*

⁴³ *ibid.*

⁴⁴ SJC (Vol 1), 201 and JC2/7 102 & 103.

⁴⁵ This reference to 'modern' 'art and part' liability is to Sir Baron Hume's treatise: *Commentaries on the Law of Scotland*. For a fuller discussion of 'art and part' liability, see D Hume, *Commentaries on the Law of Scotland* (2nd edn, Bell and Bradfute 1819), i, 273. For a more current understanding of 'art and part' see the Criminal Procedure (Scotland) Act 1995, s 293.

Later in the period under review, panels began to invoke not only the defence of a lack of malice but also that of the lack of a deadly wound. This was the case in *James Heart* (1637),⁴⁶ in which the panel was charged with slaughter for giving a ‘cruel and unmerciful’ deadly strike to the defunct.⁴⁷ What follows in the court records is a rather elaborate discussion between the advocates and the Justices about the nature of the wound and the alleged ‘*precogitat malice*’ on the part of the panel. The panel contended that there had been no malicious discord and argued that, as such, he could not be accused of slaughter; further, he argued that he had not inflicted a mortal wound upon the defunct.⁴⁸ The Justice, having read and considered the ‘exceptions and duplyis’ made by the panel and the pursuer’s answers, found the criminal letters relevant and remitted them to the trial of an assize. The assize returned a verdict of ‘clean’ and ‘innocent’ and ‘acquit[ted]’ Heart of the slaughter. The assize would have been present in court when the panel, advocates and pursuers were making their arguments. Given the assize’s reliance on the criminal letters,⁴⁹ it seems likely that the defence put forward by the panel, to the effect that there had been no ‘*precogitat malice*’ in connection to the slaughter, was a key reason for the acquittal.

This indicates that ‘slaughter’, when specifically libelled, tended to involve ‘*precogitat malice*’ between the parties. If the panel could put to the court, in defence, that he or she had not harboured such malicious sentiments towards the defunct, there would likely be an acquittal. The above cases indicate a focus on a ‘*precogitat malice*’ in the sense of prior discord between the panel and the defunct, as opposed to forethought or planning of the actual killing. This is important to the internal classification of slaughter but even more important to the distinction between murder and slaughter. What appears to be the case, in connection to the libel of ‘slaughter’, is that ‘*precogitat malice*’ consisted of hostile behavior or words towards a person; whether this had to occur *immediately* prior to the violent act leading to death is unclear. This can be compared with murder, explored in more detail below, whereby a specific circumstance would be libelled in connection to the killing from which premeditated *homicidal* intent could be inferred.

ii. Slaughter Chaudemella

‘Slaughter *chaudemella*’ was a different libel from ‘slaughter with *precogitat malice*’. In *John Bell* (1643),⁵⁰ which concerned a charge of slaughter, the advocate stated that ‘our auld law and practik’ distinguish between slaughter that is ‘*chaudemella*’ and murder ‘per *precogitatam malitiam*’, and that the crime of slaughter ‘without forthought felony’ was of a different nature from slaughter ‘per *precogitatam malitiam*’.⁵¹ This case follows the classifications that were put forward in *William Jamesoun and another* (1632),⁵² which suggest that there were two principal ways in

⁴⁶ SJC (Vol 2) 332.

⁴⁷ *ibid.*

⁴⁸ SJC (Vol 2) 339.

⁴⁹ See n 29.

⁵⁰ SJC (Vol 3) 582.

⁵¹ SJC (Vol 3) 586.

⁵² SJC (Vol 1) 201 and JC2/7, 102 & 103.

which slaughter could be libelled: that which was *chaudemella* and that which was *precogitatam malitiam* (characterised by prior malice). The former term is not defined in the Justiciary Court records. However, the Act of 1372 defines 'chaudemella' as 'the heat of anger', and 'foirthought felony or murder' as a 'certain and deliberate purpose'.⁵³ From the court discussion, and with reference to the statutory material, it seems that 'slaughter *chaudemella*' involved a culpable killing but one that lacked '*precogitat malice*'. In other words, the term *chaudemella* seems to have referred to a killing committed in 'hot blood'.

iii. 'Accidental Slaughter'

Towards the end of the period under consideration, certain panels alleged that there had been no malice and that the killing had, in fact, been accidental. This classification seemed to indicate a special, exceptional category of 'slaughter', distinguishable from 'slaughter with *precogitat malice*'. *Williame Watstone* (1639) concerned the alleged slaughter of Niniane Calderwoid. The panel and his advocates stated that if the defunct had been slain, as was alleged in the criminal letters, then it had been done ignorantly and without any 'foirthought felonie or *precogitat malice borne towards*' the said defunct.⁵⁴ The Justices continued the diet until 3 June. They ordered letters to be directed for the summoning of witnesses who would be able to attest to, 'the veritie and the form and manner of the said slaughter'.⁵⁵ Interestingly, the Justices accepted the declaration made by the panel and his advocates concerning the 'maner and sudane accident'. It is unclear whether the above statement by the accused was meant to imply that he was only guilty of 'accidental slaughter', or if he was attempting to exculpate himself completely. Unfortunately, the case was ultimately deserted because of the Justices' absence.

There is only one case in which 'accidental slaughter' was specifically libelled. The pursuers in *James Mathiesone* (1640) alleged the 'rashe and accidental' slaughter of the defunct by causing stones to roll off of Edinburgh castle.⁵⁶ The defunct was playing at the foot of the hill and was killed by the falling rocks. Mathiesone was convicted and banished. *Mathiesone* can be contrasted with *Watstone*, as the pursuer in the former case averred that the slaughter occurred because of accidental circumstances; the criminal letters only alleged 'accidental' slaughter. In *Mathiesone*, it seemed that there was genuinely no malice on the part of the panel. Still, he was not blameless in the eyes of the court, for the 'rashe' conduct libelled in connection to the slaughter ultimately led to his conviction and banishment. It is quite likely that the term 'rashe' here referred to the carelessness and negligence of his actions.

The above cases illustrate the importance of distinguishing between libels and defences in 'slaughter' cases. The libel in *Mathiesone* expressly stated that the slaughter was 'accidental'; the libel in the first case considered (*Watstone*) was simply one of slaughter, the panel attempting to prove that it was accidental by way of defence. As such, the case of *Watstone* is particularly illustrative of the distinction between 'slaughter with *precogitat malice*' and 'slaughter without *precogitat malice*'

⁵³ RPS, 1372/3/6 (n 7).

⁵⁴ SJC (Vol 1) 305 and JC2/8, 76.

⁵⁵ *ibid.*

⁵⁶ SJC (Vol 2) 395.

(or ‘accidental slaughter’). It demonstrates that ‘slaughter’ was generally understood to be ‘slaughter with *precogitat malice*’; hence a successful conviction of slaughter usually required the pursuer to prove that there had been a *precogitat* malicious relationship. A panel’s argument, by way of defence, that the slaughter had been committed accidentally meant that he or she was attempting to obtain an acquittal, as in *Watson*, or at least a more lenient sentence, as in *Mathiesone*.

iv. Conclusions

The above discussion on malice allows for certain conclusions to be drawn regarding the classification of slaughter. The first concerns the central importance of the criminal letters in the classification of the crime. It seems that the criminal letters needed to be very specific to the circumstances surrounding the crime of slaughter. One such circumstance that tended to be alleged by pursuers was the mental state of ‘*precogitat malice*’. This thinking in advance seemed to have been connected to specific conduct, for instance hostile behaviour or words between the panel and defunct. Besides the central case of ‘slaughter with *precogitat malice*’, there appear to have been two other ways in which slaughter could be labelled: ‘slaughter *chaudemella*’ and ‘accidental slaughter’.

It could be that slaughter, more generally, was used as a term for homicide. This was alleged by Grant in his study of the sixteenth century.⁵⁷ On that broad view, murder was simply a form of ‘slaughter’. Still, it remains to be considered what distinguished murder from ‘slaughter’ in the narrower sense of the word.

As explored in section 3 (below), murder, like slaughter, entailed an element of thinking in advance. The important conceptual distinction between slaughter and murder was the fact that, in the latter case, the forethought concerned the act of killing itself, whereas the concept of forethought associated with slaughter (*precogitat malice*) related to a more general, non-homicidal manifestation of hostility. However, it would potentially be an overstatement to argue that *all* cases of murder and slaughter in the early seventeenth century were distinguishable in that manner.

The other types of slaughter besides ‘slaughter with *precogitat malice*’ can be referred to collectively as ‘slaughter without *precogitat malice*’. This category of slaughter sub-divides into ‘slaughter *chaudemella*’, which probably concerned a spontaneous, emotionally-fuelled killing not in any way prefigured by feelings of malice, and ‘accidental slaughter’, which, as the name suggests, encompassed cases in which there had not been any homicidal intent at all, premeditated or otherwise.

C. Nature of the Wound

The nature of the wound was an important aspect of the evidential burden associated with the prosecution of slaughter from 1625-1650. This section of the article examines three terms used to describe the injury inflicted by the defunct: namely, ‘diverse strikes’, ‘cruel strikes’ and ‘deadly wound’. It considers to what

⁵⁷ Grant, ‘Murder Will Out’ (n 13) 218.

extent, if any, these features of the pleadings bore upon the classification of slaughter.

i. 'Cruel' Strikes in Connection to 'Slaughter with Precogitat Malice'

As discussed above, the main type of slaughter was that committed by '*precogitat malice*'. It could be that the way in which malice was inferred was through the use of the word 'cruel'. In *Jon Sagery* (1625), the earliest case in the period selection, the panel was charged with the slaughter of Helene Thomestune by 'severall crewall strikes'.⁵⁸ The case was 'continued', and does not appear again in the period under review. The use of the term 'cruel' in connection to the strike can, however, be explored in a later case.

In *Margaret Dannills and another* (1627),⁵⁹ the panels were charged with the 'cruel slaughter' of the defunct.⁶⁰ It was alleged in the criminal letters that the slaughter was committed by means of strikes to the defunct's head, back and side and one cruel strike inflicted by *one* of the panels.⁶¹ The panels argued that the criminal letters were not relevant because of the way in which the strikes were libelled.⁶² The criminal letters did not give a clear impression as to *who* had given the cruel strike that caused the death of the defunct. They simply stated that, 'one cruel strike was given by the panel', causing confusion as to which of the panels had committed that specific act. This detail seems to have influenced the assize in its decision to acquit both Dannills and Bell.

It is suggested here that the use of the term 'cruel' in relation to the injury inflicted was indicative of '*precogitat malice*' towards a person; if the pursuer libelled a 'cruel' strike it would be possible for the court to infer from that word the allegation of malice. This follows from Armstrong's suggestion, made in his study of medieval Justiciary Ayre cases, that the use of the word 'cruel' was meant to signify 'wickedness or mal-intent'.⁶³

ii. Diverse Strikes

There are two cases in which the pursuers alleged that so-called 'diverse strikes' had been sustained by the defunct. 'Diverse strikes' were strikes given to several parts of the body rather than one part. The case of *Peter Balmanno and others* (1626) records the slaughter of the defunct by, 'dyuers straikis ane grevous wound (...) in his breist, shoulderis, heid, bak, bellie and dyuers utheris partis of his body and thairby bruiseing and breaking his intrall and noble paris uithin him':⁶⁴ a rather detailed description of the location of the wound and the effect that it had on the body. This focus on the frequency of the strikes can also be seen in the case of *Thomas Hunter and another* (1627), in which the panels were charged with the slaughter given by

⁵⁸ JC2/6, 316.

⁵⁹ JC2/6, 433.

⁶⁰ *ibid.*

⁶¹ *ibid.*

⁶² *ibid.*

⁶³ JW Armstrong, 'The Justice Ayre in the Border Sherifffdoms, 1493-1498' (2013) 92 *Scottish Historical Review* 1, 26.

⁶⁴ SJC (Vol 1) 37 and JC2/6, 355.

'duyers deidlie strikies'.⁶⁵ This aspect of 'diverse strikes' may have featured prominently in the pleadings but it does not seem to have been a requirement for a successful conviction of 'slaughter'.

iii. Deadliness of Wound

In slaughter cases, one aspect of the pursuer's evidential burden was to prove that the strike(s) inflicted upon the defunct had been 'deadly' or, in other words, that they had caused the latter's death. In *Thomas Crombie and others* (1625),⁶⁶ it was alleged that the panel had slaughtered the defunct by striking him on his right arm. The court examined the nature of the wound libelled and the implications that this had on the panel's culpability. The defunct had been struck with a short sword on his right arm and had died the following day.⁶⁷ The panel refuted the libel, stating that the criminal letters were not relevant because they had not libelled the injury as a deadly wound. A doctor appeared before the court to give evidence that the wound was not lethal. Counsel for the pursuers, however, submitted to the court that, although the wound had not been deadly, the libel was still relevant because the panels gave the defunct a 'crucial' strike, resulting in a wound that *became* deadly. Given the context of the pursuers' allegations it seems that they were attempting to illustrate that the 'crucial' strike was fundamental in causing the death of the defunct. The panels were found innocent of the charge of slaughter.⁶⁸ It is possible that the initially non-deadly nature of the wound was what convinced the assize of their innocence. Clearly, the failure to libel, and subsequently prove in court, a deadly strike would have cast into doubt whether the panel had brought about the defunct's death.

The 'deadly wound' issue that featured in *Crombie* was one of causation in all but name. Interestingly, the facts of this case seem to suggest that the strike libelled could, under the modern 'but for' test,⁶⁹ be deemed to have caused the defunct's death. In other words, but for the strike, the defunct would not have died of a wound that *became* deadly; indeed, there would not have been a wound in the first place. This raises the question as to what was needed for a wound to be deemed causal in the seventeenth century. Seemingly, the wound would have only been deemed causal if it had been (1) *sufficient*, on its own, to bring about the defunct's death and (2) *sufficient at the time* of infliction. The case of *Thomas Crombie* certainly supports that supposition, suggesting that a wound that was not deadly on its own, but which led to death *later on* (perhaps because, to give a hypothetical example, it became infected), would not have been deemed causal of the defunct's death.

The trial of the panel in *Johnne Young* (1630) for the slaughter of Archibald Reid contained a lengthy court discussion about the nature of the wound and the defunct's subsequent death.⁷⁰ The panel alleged that the criminal letters were not relevant, 'to pass to the knowledge of the assize', as the strike was libelled to have

⁶⁵ SJC (Vol 1) 64 and JC2/6, 393.

⁶⁶ SJC (Vol 1) 18 and JC2/6, 324.

⁶⁷ *ibid.*

⁶⁸ *ibid.*

⁶⁹ For a description of the 'but for' test in modern criminal law see, eg, *McDonald v HM Advocate* 2007 SCCR 10, 11.

⁷⁰ SJC (Vol 2) 313 and JC2/6, 667.

been inflicted on the defunct's 'schakill', which can probably be interpreted as referring either to the defunct's wrist or ankle.⁷¹ In either case, then, the blow had clearly not been directed at a vital part of the body.⁷² After the alleged strike, the defunct had been able to continue his ordinary work as a smith, 'communing with his cronies, binding, staking and his other ordinary efforts up and down the country', and he, 'never lay bedfast all that time', until shortly before his death, which was caused by a fever.⁷³ The pursuers answered that the libel was relevant and argued that, '[although] the schakill bane [was not] a vital part' of the body, 'life consists [of] all (...) parts of the body'.⁷⁴ The panel countered that when any man survives for forty days after a wound it cannot be presumed that he died thereof.⁷⁵ The defunct had lived for four or five months after the strike,⁷⁶ and so easily satisfied this 'forty day' test. The panel was subsequently acquitted.

The cases of *Young* (1630) and *Crombie* (1630) make for a worthwhile comparison. In each, any prospect of conviction was thwarted by the lack of a deadly wound. The fact that both cases seemed to fail on that basis becomes interesting when one remembers that the defunct in *Crombie* survived for considerably less time (one day) than the defunct in *Young* (several months). In *Crombie*, the inference that the wound was not deadly was probably bolstered by the presence of medical evidence, which seemed to weigh compellingly in favour of a non-deadly wound, hence acquittal. By contrast, in *Young*, where the defunct survived for a much longer period, the evidence prayed in aid was of a much less specialised variety. For instance, the panel argued that the defunct had been able to continue his daily activities, notwithstanding his wound. Otherwise, however, the cases are very similar. Both essentially concerned causation and the inference that the wound inflicted had not been sufficient, on its own, to bring about the defunct's death.

Several cases subsequent to the two discussed above started to include a 'deadly strike' in the libel and to detail the location of the wound and the effect that it had had on the defunct, as shown in Table 2 below.

⁷¹ According to the Dictionary of Scots Language (hereinafter the 'DSL') the schakill was 'a fetter for the ankle or wrist of a prisoner.' 'DSL' <www.dsl.ac.uk/entry/dost/schakill_n> accessed 22 July 2016.

⁷² SJC (Vol 2) 313 and JC2/6, 667.

⁷³ *ibid.*

⁷⁴ *ibid.*

⁷⁵ *ibid.*

⁷⁶ *ibid.*

Case	Nature of stroke	Verdict
<i>William Hangetsides (1632)</i>	'deadly strike'	past simpliciter ⁷⁷
<i>James Wright (1632)</i>	'deadly strike in the breast'	acquit
<i>Andro Grahame (1632)</i>	'deadly strike'	convicted
<i>Jon Waterstony and others (1632)</i>	'deadly strike in his breast, bellie and sydies breaking of his internal and noble parts'	past fra ⁷⁸
<i>James Balfour (1632)</i>	'deadly strikes hurting and wounds upon his breast, sides, bellie'	deserted
<i>George Linly (1632)</i>	'deadly and cruel strike in his body and upon his sidies and his head'	past simpliciter
<i>James Jonston and another (1634)</i>	'deadly wounds in his head, breast, bellie and sides'	continued (i.e. does not appear again in period under review).
<i>Mr Gavin Dunbar and another (1637)</i>	'cruel and deadly strikes in his head and under parts of his body'	continued
<i>Thomas Stott (1643)</i>	'cruel and deadly strike'	convicted

~TABLE 2 – SHOWING THE CASE LIBELS TO INCLUDE THE WORDS 'CRUEL' AND/OR 'DEADLY' IN THE CHARGING OF THE PANEL.

It seems that it was simply not enough to allege that the strike given had been cruel: the pursuer had to show that the wound had been deadly or, which amounts to the same thing, that the panel's strike had caused the defunct's death. If the pursuer failed to show this, or the panel provided evidence suggesting that the wound inflicted had not been deadly, there were several possible outcomes: the assize would acquit the panel; the case would be deserted entirely by the Justices; or the pursuer would agree to 'past fra'.⁷⁹

A detailed exploration of the evidence needed to show the deadliness of the wound was again illustrated in the case of *Janet Corsair (1641)*. The panel was

⁷⁷ 'Past simpliciter' was one of the verdicts returned by the Justiciary Court in homicide cases from 1625-1650. According to the DSL, 'simpliciter' meant 'simply, without qualification or condition being placed upon the event described; without further restriction' ('DSL' <www.dsl.ac.uk/entry/dost/simpliciter> accessed 4 November 2016). Both the verdict of 'past fra' (see n 78 below) and that of 'past simpliciter' indicated that the pursuits had been dismissed. However, the verdict 'past simpliciter' seemed to differ from 'past fra' in indicating that the court, as against the pursuers, had decided to cease pursuit of the case.

⁷⁸ 'Past fra' was one of the verdicts returned by the Justiciary Court in homicide cases from 1625-1650. It seems, from the reading of the records, that a verdict of 'past fra' occurred when the pursuer agreed to forgo pursuit. The DSL does not define this verdict, nor do the records.

⁷⁹ See *ibid.*

charged with the cruel slaughter of James Johnstoun, an eight-year-old boy, by gripping him on his two ears and casting him with great force to the ground whereupon she gave him strikes with her hands, fists, knees and feet.⁸⁰ The panel alleged that, though libelled as 'deadly', the strikes allegedly inflicted upon the child could not have been deadly as the child survived for six months after the alleged incident. Several witnesses were presented, who verified that two or three months after the incident the child was walking about in as good health as, 'if not better' health than, he had been in prior to the incident; these circumstances were deemed incompatible with the infliction of a mortal wound.⁸¹

The panel also alleged that, a month after the incident, a great sickness of pox had raged for months among the children of Leith and Edinburgh, but that this child had subsequently managed to recover therefrom.⁸² Implicit in this statement was the argument that the child would not have survived the pox had the wound been deadly.⁸³ The panel also presented to the court a surgeon called Patrik Johnstoun, who had visited and attended to the child before his death. He gave evidence that the child had died of another cause, namely the fluxes.⁸⁴ The court continued to examine the issues of relevancy of defence witnesses to be called and ultimately acquitted the panel of the slaughter. The panel was able to prove to the court that the original wound was not deadly; hence a libel of the crime of slaughter was unsustainable. Clearly, the appearance of an expert witness, namely the aforementioned surgeon, would have facilitated that conclusion. Like *Crombie and Young*, the case of *Corsair* indicates the importance of connecting the panel's wrongful conduct with the defunct's subsequent death.

D. Conclusions

The above discussion shows that the nature of the wound was very important to a charge of slaughter. Firstly, it becomes apparent, from the prosecution of slaughter in all its sundry forms, that the pursuers had to prove that the wound had been 'deadly'. If, however, the panel could provide evidence to the contrary, the case would be deserted or the panel acquitted by the assize. The need for a 'deadly' wound was very much linked to causation. The pursuer had to prove that the defunct's death had *resulted* from the wound inflicted upon him or her. The notion of a 'cruel' strike was very different: it was a circumstance from which malice could be inferred, hence it was specifically related to the mental element of 'slaughter with *precogitat* malice'. In determining whether 'malice' had been present, some of the cases relied upon the wording of the criminal letters and the details therein of a cruel and deadly strike, while other cases relied on the testimony of witnesses to prove that there had been some form of hostile discourse between the panel and the defunct (as explored above). Finally, the term 'diverse strikes' was sometimes used

⁸⁰ SJC (Vol 2), 414 and JC2/8, 125.

⁸¹ *ibid.*

⁸² *ibid.*

⁸³ *ibid.*

⁸⁴ *ibid.*

to describe injuries to multiple parts of the body. ‘Diverse strikes’ do not, however, seem to have been a requirement for a successful conviction of slaughter.

It is important to consider what relevance, if any, the above three characteristics of a wound (‘deadly wound’, ‘cruel strikes’ and ‘diverse strikes’) had to the classification of ‘slaughter’ and the three different forms thereof. Clearly, the notion of a ‘cruel strike’ had great classificatory significance in distinguishing ‘slaughter with *precogitat* malice’ from the other two types of slaughter. From the presence of so-called ‘cruel’ strikes could be inferred ‘malice’, and from ‘malice’ could be inferred ‘slaughter with *precogitat* malice’. Certainly, the notion of ‘diverse’ strikes was similar to ‘cruel’ strikes in its clarification of the nature of the wound, but the former does not seem to have played a classificatory role; it did not relate to a distinctive mental element. Finally, though the notion of ‘deadly’ wound was very significant, given that ‘deadliness’ was necessary for a conviction of ‘slaughter’, it was arguably not of classificatory significance. It related, essentially, to causation, and so, being a requirement that applied to all forms of ‘slaughter’, was not a particularly useful point of classification. It follows that the nature of the wound was relevant to the classification of ‘slaughter’ in some, but by no means all, respects.

3. *The Crime of Murder*

A. Overview of Findings

This article suggests that, to sustain a charge of murder in the Justiciary Court, as distinct from a charge of slaughter, the Justices and assize needed to be convinced that the panel’s act had been intentional and deliberate. Here the notions of ‘intent’ and ‘deliberation’ referred specifically to *homicidal* forethought (or ‘aforethought’). In other words, to be guilty of murder, the panel must not merely have harboured hostile sentiments towards the defunct prior to the act of killing (‘*precogitat* malice’), but must have planned the killing itself. Such a specific, technical understanding of ‘intent’ is to be distinguished from ‘intent’ in the loose sense of the word. For instance, whereas someone who kills another in a spontaneous rage can, to the modern layperson, be understood to have acted with ‘intent’, such a killer would not have been deemed to have acted with ‘intent’ for the purposes of a ‘murder’ charge in seventeenth century Scotland.

In ‘murder’ cases, the intent or deliberate purpose preceding the homicidal act was not libelled by express mention of aforethought in the criminal letters; instead, this was done indirectly, by libelling in connection to murder certain circumstances from which deliberate conduct could be inferred. These circumstances included treason, killing in a relationship of trust, ‘unnatural’ killing (e.g. where a mother killed her child), or killing by stealth (e.g. at night or by poison).

Treasonable murder started to be libelled and classified as murder ‘under trust’ in the latter part of the period under review. Here the word ‘treasonable’ did not refer to betrayal of one’s country, but betrayal of a person; it concerned the scenario where the panel and defunct had been in a relationship of trust and the panel had betrayed that trust by killing him or her. A homicide where the defunct

had been under the trusted credit, assurance and power of the accused would have amounted to an aggravated form of murder known as 'treasonable' murder. The difference between 'treasonable murder' and 'murder under trust' was inconsequential; these were essentially different terms for the same crime.

The following study is arranged to show the classification of the aforementioned types of murder, each of which corresponded to the particular circumstance suggestive of 'deliberate' killing.

B. Murder under Trust and Treasonable Murder

The concepts of 'treasonable murder' and 'murder under trust' were indistinguishable. Consequently, the mere fact that the latter term became more commonly used will not, in itself, have effected a drastic change to this area of the law. In *Andro Rowane* (1627), the panel was charged with the 'treasonable murder' of Euphame Douglas, his spouse. It was put to the court that contained in an Act of Parliament was the rule that, 'if a party is slain under trust, credit, assurance or power of the slayer he shall be put to death for the crime'.⁸⁵ As the defunct had been the lawful spouse of the panel, he was considered to have, 'most cruelly and unnaturally under her trust and his power murdered her'.⁸⁶ This can be compared with the case of *Jon Jamesone* (1628), wherein the panel was charged with the 'cruel murder and slaughter' of his spouse by giving to her a cruel and deadly strike.⁸⁷ Jamesone was put to the knowledge of the assize, which returned a verdict of 'culpable' and convicted him of the 'horrible murder of his spouse'.⁸⁸ Thus, despite the similar facts with which *Rowane* and *Jamesone* were concerned, the approach to the legal question was slightly different: notably, the panel in the latter case was not charged with treasonable murder. Still, it is significant to note that the assize convicted him of the 'horrible murder' of his spouse. The conviction of murder, as against slaughter, was likely a reflection of the fact that the defunct had been the panel's spouse and, consequently, that the two had been in a trustful relationship.

These two cases also show the divergence of the statutory law and the law of homicide in practice. The aforementioned Act of Parliament states that:

(...) where the party slain is under the trusted credit, assurance and power of the slayer, all such murder and slaughter to be committed in time coming after the date hereof, the same being lawfully tried and the person dilated found guilty by an assize thereof, shall be treason, and the persons found culpable shall forfeit life, lands and goods.⁸⁹

This seems to indicate that 'treason', rather than 'murder', should have been libelled in the cases just discussed. It is likely that the first case of Rowan was libelled as 'treasonable' as a result of the reading of this statute. It could also be the case that 'murder under trust' indicated a circumstance that aggravated the libel of murder.

⁸⁵ APS III 45, c 34, RPS 1587/7/44.

⁸⁶ SJC (Vol 1), 73 and JC2/6, 437.

⁸⁷ JC2/6, 513.

⁸⁸ *ibid.*

⁸⁹ APS III, 45, c 34, RPS 1587/7/44.

This was certainly the case in *Walter Urquhart and others* (1642), in which the panels were charged and found culpable and guilty of the crime of, ‘treasonable, barbarous and cruel murder under trust and friendship’ of the defunct.⁹⁰ This case links the notion of ‘treasonable murder’ with that of ‘murder under trust’.

C. Murder by ‘Unnatural’ Circumstances

Several murder cases in the period under review involved a mother killing her infant child. Most importantly, however, the term ‘child murder’ or ‘infanticide’ was not used in the Justiciary Court records. Instead, this crime was termed the ‘unnatural murder and destruction’ of an infant. It is not clear why this violation of the maternal bond was placed into the special category of murder by ‘unnatural circumstances’, and not simply deemed another form of ‘treasonable murder’ or ‘murder under trust’.

Three of the ‘unnatural circumstances’ cases discussed below involved a mother killing her infant child to conceal adultery. This is not the only instance of the word ‘unnatural’ being used in connection with the classification of murder: as discussed above, *Andro Rowane* (1627) was alleged to have, ‘most cruelly and unnaturally [murdered his wife] under her trust and his power’.⁹¹ What is meant by the word ‘unnatural’ is unclear, but given its use in the five cases below and the case of *Rowane*, it seems to have indicated the violation of a relationship of the utmost sanctity, of which the two paradigms were spousal and parental.

Case	Year	Citation	Charge
<i>Jonet M’Craith</i>	1626	SJC (Vol 1) 47	‘horrible murthour’
<i>Christiane Hamilton</i>	1634	JC2/7, 305	‘unnatural murder and destruction’
<i>Margaret M’Linle</i>	1634	JC2/7, 317	‘cruel and unnatural murder and destruction’
<i>Margaret Cunningham</i>	1642	SJC (Vol 2) 532	‘cruel and unnatural murder’
<i>Jonet Campbell</i>	1644	JC2/8, 383	‘cruel murder and slaughter’
<i>Jonet Jonstone</i>	1644	JC2/8, 388	‘cruel and unnatural murder’
<i>Margaret Chalmers</i>	1648	SJC (Vol 3) 768	‘cruel murder, death and destruction’

~TABLE 3 – SHOWING CASES INVOLVING THE KILLING OF AN INFANT CHILD BY ITS MOTHER.

⁹⁰ SJC (Vol 2) 550.

⁹¹ SJC (Vol 1) 73 and JC2/6, 437.

The cases of *Hamilton*, *M'Linle*, *Cunningham* and *Jonstone* mentioned in Table 3 all used the phrase 'unnatural murder' in relation to the killing of an infant child. This terminology was used in three of the seven cases of homicide committed to conceal adultery. The case of *Campbell* deviates from that pattern. Despite the use therein of the phrase 'cruel murder and slaughter' to describe the killing of an infant, the facts of this case did not involve a killing intended to conceal adultery. The case of *Chalmers* also differs from the others in its use of the phrase 'cruel murder, death and destruction' to describe a mother's killing of her infant child in an attempt to conceal adultery. From *Campbell* and *Chalmers* it would seem that the term 'unnatural' was replaced by 'cruel' in the mid-1640s.

These six cases are notable for the absence of defence counsel. Although these were not the only cases in the period under review where no advocate appeared to defend the panel, this appears to have been a relatively rare occurrence, accounting for roughly 20 per cent of the homicide cases considered in this article.⁹² It has been suggested by Wasser that it was common for a person accused of a crime before the Justiciary Court to have an advocate present.⁹³ This is plausible in the light of the Parliamentary Act of 1424, which required judges to appoint counsel to 'poor men free of charge', although Wasser has suggested that this was likely intended for civil suits.⁹⁴ The various auld laws, practicks, treatises and Acts of Parliament, passed before the 1424 Act, concerning legal representation were either ambiguous in their scope or mentioned civil cases only when referring to aid being given to a defender.⁹⁵ By the beginning of the seventeenth century, it was probably common for a person accused of a crime before the Justiciary Court to be represented by an advocate.⁹⁶ In 1587, an Act of Parliament stipulated that no advocate or prolocutor should be stopped, in Parliament, from appearing, defending, or reasoning on behalf of a person accused of treason or otherwise.⁹⁷ Thus, even in the case of a very serious criminal charge (treason), the accused could avail himself or herself of an advocate. Against this backdrop, the lack of legal representation in cases involving infanticide is particularly marked.

In each of the six cases mentioned above, the woman on trial confessed judicially to the killing of her child, and was thereafter capitally punished. It is unclear whether the lack of defence counsel contributed to the women's convictions; it is, conversely, possible that it was their confessions that accounted for the absence of advocates in the first place. It is also unclear whether the confessions were obtained through the use of torture, as the records do not indicate as much.⁹⁸ One of these cases, however, states that a woman confessed in the avowed hope that God

⁹² In other words, between 1625 and 1650, 164 cases were decided. Of these, a mere thirty-six did not feature an advocate (including the six aforementioned cases on infanticide).

⁹³ See generally M Wasser, 'Defence Counsel in Early Modern Scotland: A Study Based on the High Court of Justiciary' (2000) 26 *JLH* 183.

⁹⁴ *ibid.*

⁹⁵ Wasser, 'Defence Counsel' (n 93) 184.

⁹⁶ *ibid* 183.

⁹⁷ APS III, 443, c 16, RPS 1587/7/26.

⁹⁸ B Levack, 'Judicial Torture in Scotland during the Age of Mackenzie' in H L MacQueen (ed), *Miscellany IV* (Stair Society (LexisNexis Butterworths) 2002) 185.

would show mercy.⁹⁹ Thus, some confessions were possibly motivated not simply by the fear of torture, but by heartfelt religious sentiment.

D. Murder by Stealth

The commission of homicide by stealth, or in the darkness of night, could successfully lead to the prosecution of a charge of murder. This scenario arose in the case of *Johnne Myller and another* (1636), wherein the panels were charged with ‘murder under night’ by giving six or seven cruel and deadly strikes to the sleeping defunct on his head.¹⁰⁰ This aspect of stealth was also evident in the case of *Robert Walker and others* (1642), in which the panel killed the defunct ‘under silence and cloud of night’.¹⁰¹ According to the libel, the defunct was sent from her house at night by her husband, Robert Walker, to a neighbour’s house for a pint of ale. She was then shamefully and cruelly strangled to death by that neighbour, David Grahame. He did so under the order of Robert Walker and David Grahame’s wife, Margaret Grahame. It was alleged that Robert had committed adultery with Margaret; had been called before the Session of the Kirk of Montrose for the adultery previously; and had planned the killing of his wife to prevent her from revealing the adultery.¹⁰² All three conspirators were found guilty of murder ‘art and part’. This case indicates a clear, deliberate purpose; indeed, ‘forethought’ was specifically mentioned in the libel.

E. Conclusion

For a successful conviction of murder to be brought in the Justiciary Court, the Justices and assize had to be convinced that there had been some ‘intent’ or deliberate purpose underlying the killing. This was done by including in the libel a reference to such circumstances as treason, unnatural killing and killing by stealth (e.g. at night or by poison). Treasonable murder was murder aggravated by the betrayal of a trustful relationship between the panel and defunct. It was aggravated in the sense that the homicide would be libelled as murder but could also lead to a charge of treason under the statutory law. The term was used in the Justiciary Court records at the beginning of the period under review but seems later to have been superseded by the term ‘murder under trust’.

Most cases do not explicitly indicate when and how ‘forethought felony’ played a part in the murder. In this, the practice of the prosecution of murder in 1625-1650 might have differed from the statutory law discussed above.¹⁰³ However, the circumstances by which murder was termed and libelled seem to indicate an intent or deliberate purpose behind the homicide. Invoking such a deliberate purpose seems to have been sufficient to import a ‘forethought felony’ in connection with the murder – the key element of a successful prosecution. It should be stressed

⁹⁹ *Margaret Chalmers* (1648), SJC (Vol 3) 768 and JC2/8, 579.

¹⁰⁰ SJC (Vol 1) 274 and JC2/7 662. The verdict for this case appears at JC2/6, 700.

¹⁰¹ SJC (Vol 2), 536 and JC2/8, 240.

¹⁰² *ibid.*

¹⁰³ RPS, 1372/3/6 (n 7).

that the nature of the forethought associated with murder was specifically homicidal in character; it signified that the panel had not only borne malice towards the defunct, in advance of the killing (*'precogitat malice'*) but had planned to kill him or her.

To some extent, the seventeenth century notion of 'murder' was similar to the older, more restrictive understanding of the term as *secret* killing specifically.¹⁰⁴ This is particularly apparent from the aforementioned importance of murder by stealth. Also, it might be hazarded that some of the other categories of murderous killing, for instance 'unnatural killing' of an infant child, were also, incidentally, committed in secret. However, the existence of *multiple* circumstances from which murderous 'intent' could be inferred is suggestive of a broader definition of 'murder' than one associated solely with secret killing.

4. *The Libel of 'Murder and Slaughter'?*

A. Overview of Findings

Several records refer to charges of both 'cruel slaughter and murder'.¹⁰⁵ This double charge has not received scholarly exploration and has been disregarded in the index and introduction to the printed Justiciary Court records.

Usually, the cases in which the panel was charged with both 'murder and slaughter' concern homicide in circumstances which, for reasons explored immediately below, were potentially indicative of an intentional, premeditated killing. Two key examples of these circumstances were killings committed at night and killings committed in the context of hospitality. It thus becomes slightly unclear why the double charge of 'cruel slaughter and murder' was libelled in these cases rather than just 'murder'. Perhaps this was done so that the pursuer had a higher chance of securing a conviction. Apprehensive as to whether he or she would be able to sustain a charge of murder, the pursuer may, quite plausibly, have libelled 'slaughter' in the alternative.

This impression seems particularly credible given the importance of the criminal letters and the fact that the cases would be deserted if a pursuer only libelled one charge but failed to prove it. As discussed above, just over 47 per cent of the 'murder and slaughter' cases were libelled as the *cruel* 'murder and slaughter' of the defunct.

¹⁰⁴ See text to nn 11-16.

¹⁰⁵ This double charge appeared in the following cases: *Hary Gordoun* (1627), JC2/6, 400; *Walter Jamesone* (1628), SJC (Vol 1) 4 and JC2/6, 477; *Jon Jamesone* (1628), JC2/6, 513; *David Quahte* (1629), JC2/6, 523; *Williame Lennox and another* (1635), JC2/7, 522; *James Grant* (1635), JC2/7, 629; *David Robert and others* (1637), JC2/8, 8; *James Spalding* (1637), JC2/8, 38; *James Donaldsone* (1638), JC2/8, 40; *James Graham* (1640), SJC (Vol 2) 399 and JC2/8, 107; *William Anderson and others* (1641), SJC (Vol 2) 409 and JC2/8, 122; *William Fraser* (1641), SJC (Vol 2) 443 and JC2/8, 151 & 155; *John Morrison and others* (1642), SJC (Vol 3) 524 and JC2/8, 213 & 225; *Laurence Mercer and others* (1643), SJC (Vol 3) 564; *John Bell* (1643), SJC (Vol 3) 582; *Johnne Bell* (1644), JC2/8, 374; *Johnne Campbell* (1644), JC2/8, 384; *Patrick Meldrum* (1646), SJC (Vol 3) 726; *James Urquhart* (1646), SJC (Vol 3) 728; *Margaret Lamber* (1646), SJC (Vol 3) 731; *Thomas Calander* (1647), JC2/8, 544; and *Alexander Milne* (1649), JC2/8, 667.

The libelling of specific circumstances in connection with the killing in the cases explored below (e.g. the fact that it was committed at night, in the context of hospitality, or in violation of a trustful relationship), was probably intended to infer the ‘intention’ or deliberation necessary for a conviction of murder.

B. Homicide Committed by Stealth

The previous section on murder suggests that a charge of murder would be successfully prosecuted if the killing appeared, on the evidence, to have been committed by stealth. In the case of *William Lennox* (1635), the panel was charged with the ‘cruel murder and slaughter’ of the defunct, which was committed ‘under night’.¹⁰⁶ The fact that the homicide was committed ‘under night’, hence covertly, seemed to indicate that murder, as against mere slaughter, could be inferred from the circumstances. The outcome of the case is unknown as it was ‘continued’ and did not appear again in the period under review. It seems likely that this was done merely as a safeguard to increase the chances of conviction. This also happened in the case of *James Donaldson* (1638), where the assize dropped the charge of murder (from the double charge of ‘murder and slaughter’) and returned a verdict of ‘slaughter’. The circumstances of that case are explored further below.

C. Homicide Committed in Such Contexts as Hospitality

In *William Anderson and others* (1641), the panels were accused of being guilty, ‘art and part of the cruel slaughter and murder’ of William Davidstone.¹⁰⁷ The pursuers alleged that the panel and defunct were dining together when, at some point, they ‘fell out in evil words’.¹⁰⁸ The case was eventually deserted for the lack of an assize. However, the fact that the panel and defunct were friends and were dining together seemed to indicate circumstances from which homicidal ‘intent’ could be inferred, which, had there been an assize in place to decide on the matter, would have been all too propitious for a charge of murder. This seems particularly credible when consideration is paid to the expectations, particularly those of honour and good behaviour, that hospitality tended to engender in the early-modern period.

Hospitality can be understood as, ‘preponderantly a private form of behaviour, exercised as a matter of personal preference within a limited circle of friendship and connection’.¹⁰⁹ One of the earliest uses of the term ‘law of hospitality’ was in Philip Sidney’s, *The Countess of Pembroke’s Arcadia*, which promoted good behaviour towards strangers.¹¹⁰ William Heale, who wrote against wife-beating in 1608, stated that, ‘none who entered into [the house of another], should for the time

¹⁰⁶ JC2/7, 522.

¹⁰⁷ SJC (Vol 2) 409 and JC2/8, 122.

¹⁰⁸ *ibid.*

¹⁰⁹ F Heal, *Hospitality in Early Modern England* (OUP 1990) 1.

¹¹⁰ P Sidney, *The Countess of Pembroke’s Arcadia* (1590, M Evans ed, London 1977) 385, as cited in *ibid* 5.

of his aboad there, suffer any kind of injury upon any occasion'.¹¹¹ Stair also mentions the, 'laws of hospitality' as, 'the mutual trust between the host and the guest, whom he has willingly received in his house, whereby neither of them can act anything prejudicial to the life or liberty of the other, while in that relation'.¹¹² At the beginning of the early-modern period, the elite became preoccupied with hospitality, in particular the perceived need to show honour through the household and to provide guests with good entertainment.¹¹³

Hospitality would thus infer a friendly, trusting relationship that imported certain societal and behavioural norms within the period under review. The above case does, however, describe a 'sudden and passionate exchange' between the defunct and the panel. Given the classification of slaughter and murder above, this description could mean that there was doubt that the circumstances of the case would sustain a successful charge of murder or even slaughter 'with *precogitat malice*', given the difference between '*precogitat malice*' and *chaudemella* (discussed above). As such, it is possible that the pursuers charged 'slaughter and murder' in the criminal letters in an attempt to straddle the uncertain classificatory boundary between those two crimes.

This case of *Andersone and others* can be compared with the case of *John Dick and others* (1649), in which it was alleged that the panel committed the cruel murder of the defunct by, 'poisoning him, under trust and friendship' within Dick's household.¹¹⁴ With great repentance and remorse, the panel confessed judicially to the killing in the manner specified in the criminal letters. The friendship between the panel and the defunct implied a relationship of trust, the violation of which by poisoning was all the more heinous for having occurred within the panel's household. This aspect of hospitality and friendship seems to have facilitated the inference that the killing was the product of forethought and premeditation, hence the unequivocal characterisation thereof as 'murder' rather than 'murder and slaughter'.

D. Homicide Committed in the Context of a Trustful Relationship

As mentioned above, the dual charge of 'murder and slaughter' seems to have featured in the Justiciary Court records of 1625-1650 because it increased the pursuer's chances of obtaining a conviction, an impression that is further reinforced by the case of *James Donaldson* (1638). Here the panel was charged with both 'cruel murder and slaughter', committed by giving a 'terrible and deadly strike' to the defunct, and also with adultery,¹¹⁵ although the latter charge can, for present purposes, be considered irrelevant. The assize returned a verdict of 'culpable' and

¹¹¹ W Heale, *An Apologie for Women, or an Opposition to Mr Dr G[ager] His Assertion. That it was Lawfull for Husbands to Beate their Wives* (Oxford 1609) 23, as cited in F Heal, *Hospitality in Early Modern England* (n 109) 5.

¹¹² Sir James Dalrymple, Viscount Stair, *The Institutions of the Law of Scotland* (1st edn, Andrew Anderson 1681) 1,10; Stair, *The Institutions of the Law of Scotland* (2nd edn, Andrew Anderson 1693) 1,1,11.

¹¹³ F Heal, *Hospitality in Early Modern England* (n 109) 11.

¹¹⁴ SJC (Vol 3) 811 and JC2/8, 656.

¹¹⁵ JC2/8, 40.

convicted James Donaldson of ‘slaughter and adultery’,¹¹⁶ thereby dropping the charge of murder and only returning a verdict of ‘guilty’ on the ‘slaughter’ charge. There is no indication as to why the assize returned that verdict, as only the criminal letters were relied upon and no extra evidence was produced. However, based on the classification of murder proposed hitherto by this article, it could be argued that the assize returned a verdict of ‘slaughter’ alone because the pursuers had failed to aver a relationship from which intentional, premeditated (hence murderous) killing could be inferred.

Conversely, ‘deliberate’ or ‘intentional’ killing was potentially present in the case of *Thomas Calander* (1647), wherein the panel was charged, and found culpable of, the ‘murder and slaughter’ of his son-in-law.¹¹⁷ This would suggest that the circumstances of the homicide, namely the trustful relationship between the defunct and panel, were *potentially* enough for a successful prosecution of ‘murder’. This certainly differs from the case of *Donaldson* and is more in keeping with the above cases wherein the only charge was ‘murder’. This calls into question why the pursuers included both charges of ‘slaughter’ and ‘murder’. Perhaps this was done for fear that the tie between father-in-law and son-in-law was not as close as, say, that between father and son or husband and wife, hence too weak a basis on which to ground a charge of murder.

E. Conclusions

The records do not make it completely clear why some cases were libelled as ‘murder and slaughter’. However, the likely explanation is that this double charge would have given the pursuer a better chance of securing a conviction in more borderline cases, where there was some uncertainty as to whether a murder conviction would be successful.

5. Conclusions

From the present survey of the Justiciary Court records, a number of conclusions can be drawn regarding the early seventeenth century Scottish notion of unlawful killing.

The cases considered indicate a division of the law of homicide into ‘slaughter’ and ‘murder’. For a successful prosecution of murder or ‘forethought felony’, in the Justiciary Court, the pursuer had to convince the court that the killing had been committed in circumstances from which a ‘certain and deliberate purpose’ could be inferred.¹¹⁸ This was made clear by a statute enacted in 1372.¹¹⁹ The above cases of murder illustrate that there tended to be an absence of explicit written confirmation of when and how ‘forethought felony’ had played a part in the homicide. However, the specific circumstances libelled in connection to the murder,

¹¹⁶ *ibid.*

¹¹⁷ JC2/8, 544.

¹¹⁸ RPS, 1372/3/6 (n 7).

¹¹⁹ *ibid.*

for instance treason, 'unnatural circumstances' and covert killing, seem to have been included to signal the presence of 'forethought felony': that is, a deliberate and premeditated intention to kill the defunct. The criminal letters relied upon by the pursuers, the nature of the terminology included therein, and the circumstances in the libel were a reflection of what was required to obtain a successful conviction of murder.

A tentative suggestion can be made regarding the historical context of the 'murder' cases considered in this article and what role they might have played, if any, in the wider development of the Scots law of homicide. As mentioned in the introduction, the pre-seventeenth century conception of 'murder' was, for many years, defined narrowly in terms of secret killing.¹²⁰ By the eighteenth century, the 'murder' concept had been liberalised into the notion of premeditated killing per se. Considering this article's focus on the seventeenth century, it is only fitting to ask: what happened in the interim? What inferences might be drawn regarding the development of homicide in the period 1625-1650? From the cases considered, it is clear that murder did not, by the first half of the seventeenth century, consist exclusively of 'secret killing'. 'Secret killing' was certainly one of the circumstances that would give rise to an inference of 'forethought' but it was by no means the *only* circumstance. On the other hand, the seventeenth century concept of 'murder' was arguably not one of premeditated killing per se, given pursuers' tendency to plead *specific* circumstances (for instance treason and killing by stealth) from which premeditated intent might be inferred. This could be taken to suggest that 'murder' had not yet completely freed itself from the strictures of the old, more rigid understanding of that term, and that the seventeenth century conception thereof was possibly not as wide as it would become in the eighteenth century. What conclusion might be drawn from this development? Quite possibly, the seventeenth century could be seen to represent an intermediate stage in a broader trend: namely, the liberalisation of the 'murder' concept. Equally, however, this may simply be taken to indicate that the notion of 'murder' was in flux during the seventeenth century and that the lawyers of the time had not yet agreed on a definition. Regardless, more scholarly exploration of the prosecution of crime in the seventeenth century could shed light on this particular issue.

Besides 'murder', the other key form of 'homicide' considered in this article was that of 'slaughter'. The term 'slaughter' could, in turn, be sub-divided into 'slaughter *chaudemella*', 'slaughter with *precogitat* malice' and 'accidental slaughter'. The prosecution of slaughter entailed the use of evidence regarding the nature of the wound. Pursuers ultimately had to allege and prove that the wound was deadly and inflicted maliciously. This was either done directly through the production of witness testimony (in court) or by referring to the wording of the criminal letters, which would detail whether the strike was given cruelly; was deadly; and included diverse strikes. The judicial preoccupation with the nature of the wound concerned causation; the pursuers had to show that the defunct's death had resulted from the wound inflicted by the panel.

¹²⁰ See text to nn 11-16.

In some sense, the precise meaning of the term 'slaughter' is difficult to pinpoint. 'Slaughter with *precogitat malice*' was not the only form of slaughter; there was, in addition, 'slaughter without *precogitat malice*', which, in turn, sub-divided into 'slaughter *chaudemella*', relating to culpable killing in the heat of the moment, and 'accidental slaughter'. Since it was used as a label for different crimes, it is hard to give the term a consistent, concrete meaning. As a term applicable to all three of those crimes, the term 'slaughter' was perhaps simply a convenient 'catch-all' referring to all culpable killing short of 'murder'.

One of the more fruitful comparisons that can, based on the above survey, be drawn is that between 'murder' and a specific *type* of 'slaughter', namely 'slaughter with *precogitat malice*'. In a broad sense, the two crimes were very similar. Both 'slaughter with *precogitat malice*' and 'murder' entailed the killing of another coupled with some element of malicious forethought on the part of the killer. The terms 'forethought felony' and '*precogitat malice*' both related to some sort of premeditation, which poses the questions: Were some forms of 'slaughter' simply 'murder' by a different name? Were the differences between 'forethought felony' (associated with murder) and '*precogitat malice*' (associated with slaughter) merely linguistic, rather than conceptual; more apparent than real? In fact, it was in this fundamental similarity between murder and 'slaughter with *precogitat malice*', namely the shared notion of killing preceded by some element of forethought, that the two crimes also seemed to differ in a fundamental way. With murder, malicious forethought consisted of a premeditated desire to kill the defunct. Contrariwise, the notion of malicious forethought associated with 'slaughter' (*precogitat malice*) seemed to consist of an antecedent course of dealings, which, while suggestive of prior ill-will and animosity towards the defunct on the part of the killer, did not indicate that the killing had been planned. At least to that extent, 'slaughter' did not simply refer to all unlawful killing less serious than 'murder'; it had a positive, coherent definition in its own right.

GUIDELINES FOR CONTRIBUTORS

The Editorial Board is keen to receive submissions from students, past and present, of the University of Aberdeen.

The purpose of the ASLR is to showcase the work of the students of Aberdeen, highlighting the many areas of law which are taught and researched at this university. As such, we welcome submissions on any area of law.

Several types of articles will be considered:

- Case Notes (500 to 1,500 words on a recent judicial development)
- Book Reviews (500 to 1,500 words on a recent publication)
- Essays and Short Articles (1,500 to 4,000 words)
- Full Research Papers (4,000 to 10,000 words)

All the limits specified above include footnotes.

Submissions are welcome from students at all stages, from first year undergraduate to postgraduate PhD level. Students may submit a piece of work which has been written as part of their degree, or may write something specifically for the ASLR.

Articles must provide a critical analysis of a particular area of law, publication or judicial decision. They should be original, relevant and aim to make an interesting contribution to the academic debate. Every piece of work is reviewed anonymously by the Editorial Board, before being sent for peer-review. Submissions may be accepted outright, accepted subject to modification or rejected. Constructive comments will be sent to the author.

Published articles will be compiled in an annual issue which will be available on HeinOnline and in UK law libraries.

All submissions must conform to the ASLR House Style and must be accompanied by a title page and a signed declaration of originality; further information on this can be found on the ASLR website: www.abdn.ac.uk/law/student-activities/aberdeen-student-law-review-95.

Finally, all submissions (including the title page, declaration of originality and manuscript) must be sent to aslr@abdn.ac.uk under the subject 'Submission for ASLR' by 15 September of every year.

This publication has been generously sponsored by Stronachs LLP.

All those involved with the Aberdeen Student Law Review would like to thank Stronachs for their support.



Stronachs

For further information please contact

Stronachs LLP
34 Albyn Place
Aberdeen
AB10 1FW

info@stronachs.com
01224 845845

www.stronachs.com