Annual Banking Law Update 2019

ABLU 2019 (Part ONE).indb 1 2019/10/09 9:49 AM

Annual Banking Law Update 2019

Recent Legal Developments of Special Interest to Banks

Editors: **Sarel du Toit**
Charl Hugo

Annual Banking Law Update 2019

First published 2019

Juta and Company (Pty) Ltd
First Floor
Sunclare Building
21 Dreyer Street
Claremont
7708

www.juta.co.za

© 2019 Juta and Company (Pty) Ltd

ISBN: 978 1 48513 552 4

All rights reserved. No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage or retrieval system, without prior permission in writing from the publisher. Subject to any applicable licensing terms and conditions in the case of electronically supplied publications, a person may engage in fair dealing with a copy of this publication for his or her personal or private use, or his or her research or private study. See section 12(1)*(a)* of the Copyright Act 98 of 1978.

The author and the publisher believe on the strength of due diligence exercised that this work does not contain any material that is the subject of copyright held by another person. In the alternative, they believe that any protected pre-existing material that may be comprised in it has been used with appropriate authority or has been used in circumstances that make such use permissible under the law.

Although every care is taken to ensure the accuracy of this publication, its supplements, updates and replacement material, the authors, editors, publishers and printers do not accept responsibility for any act, omission, loss, damage or the consequences thereof occasioned by a reliance by any person upon the contents hereof.

Production Co-ordinator: Valencia Wyngaard-Arenz
Typesetter: Elinye Ithuba dtp solutions
Cover Design: Drag and Drop
Printed by: Digital Action

Typeset in 10pt on 12pt Times New Roman

Preface

The *Annual Banking Law Update* of 2019 covers a range of topics within the wider field of banking law, authored by leading academics, hailing from both South Africa and abroad. All contributions should be of interest to academics conducting further research, but, as importantly, also to all banking lawyers outside the realm of academia, and thus fulfilling the aim of the Centre for Banking Law to make a contribution to knowledge in the field for the benefit of both academics and society at large.

In the first chapter the allocation of risk in the case of payment scams is examined from an international and comparative perspective. The second chapter considers the legal and policy implications of the introduction of deposit insurance in Namibia and South Africa. In a departure from chapters with a mostly legal focus, Chapter 3 provides an economic analysis of the South African Reserve Bank's financial stability policies. The always topical and often-litigated sphere of the National Credit Act is never far from a banking law book and so Chapter 4 deals with, inter alia, the undefined concept of "on the road fees". Twin Peaks, also close to any current banking law discussion, provides the backdrop to Chapter 5, where the focus is on the impact of the Consumer Protection Act on financial products and services. The Conduct of Financial Institutions Bill forms the subject matter of Chapter 6, in what must be one of the first detailed academic analyses of the Bill. The book concludes with a slight shift in emphasis to taxation, and a chapter on what the author calls the myth of financial emigration.

For the second year running, all chapters have individually been subjected to a rigorous double-blind peer-review process, conducted under the supervision of the publisher, making use of leading experts in their respective fields. Each contribution was accepted, inter alia, as being an original contribution to knowledge in the field and contributing something new or interesting to the existing scholarship on the issue – after suggestions of the reviewers were incorporated where required.

The *Annual Banking Law Update* is also an ongoing collaboration between Juta and the Centre for Banking Law, now in its fourth year, and one which I hope will endure far into the future. I wish to express my appreciation to Stephen Allcock, Leila Samodien and Valencia Wyngaard, for their invaluable assistance in publishing the book – often under strict deadlines mainly due to the tardiness of this particular co-editor. As Charl Hugo has often expressed in previous prefaces, the timelines necessitate a robust editing approach – and I bear responsibility for the mistakes which will inevitably remain.

I nevertheless trust that the book will be a worthy contribution to the still somewhat sparse collection of academic writing in this country on banking law.

Sarel du Toit
Co-editor
Centre for Banking Law
University of Johannesburg

3 October 2019

Contents

Preface v

SANDRA BOOYSEN — Tackling payment scams: A comparative review 1

DUNIA ZONGWE — Deposit insurance in Namibia and South Africa: Pricing its necessity and design 21

HYLTON HOLLANDER, DAWIE VAN LILL — A review of the South African Reserve Bank's financial stability policies 49

CORLIA VAN HEERDEN, STEFAN RENKE — Cost of credit in terms of the National Credit Act: "On the road fees", administrative fees and/or handling fees 79

MONICA LAURA VESSIO — Twin Peaks and the impact of the Consumer Protection Act on financial products and services 113

DALEEN MILLARD — Fair play? The Conduct of Financial Institutions Bill and the new face of the financial services industry 129

LG TREDOUX — The taxation of South African expatriate employees: Dispelling the myth of financial emigration 147

Table of Cases 171

Legislation 175

Conventions and Treaties 181

Tackling payment scams: A comparative review

SANDRA BOOYSEN*

1 Introduction

Rogues seeking to get their hands on money in other people's bank accounts have two options: to deceive either the bank or the customer into making a payment to them. The former will be referred to as an unauthorised payment scam and the latter as an authorised payment scam; collectively they will be referred to as payment scams. Payment scams involve dishonesty and must be distinguished from, for example, payments for defective goods or services. In the latter case, the recipient of the defective performance must generally seek recourse against the supplier; there may be a breach but there is no dishonesty.[1] The majority of payment scam case law in the 20th century involved cheques because cheques were the primary payment mechanism in that era.[2] In the cheque context, an unauthorised payment scam typically involves forgery of the customer's, *ie* the drawer's, signature on the cheque or altering the amount and/or the payee of the cheque; an authorised payment scam involves tricking the customer into issuing a cheque when they have no reason to do so.

Since the scammer will either disappear or be impecunious when caught, the loss from a payment scam invariably falls on one of two relatively innocent parties, the bank or the customer. At common law the loss generally falls on the person that was deceived. In other words, banks bear the risk of unauthorised payments, and customers bear the risk of authorised, but unintended, payments. This outcome follows from the fundamental principle that a bank must observe its customers' instructions.[3] As a general principle, if the bank has no instructions or departs from

* Associate Professor, Faculty of Law, National University of Singapore. My thanks to Sarel du Toit and the Centre for Banking Law, University of Johannesburg, South Africa for inviting me to present my research at the Annual Banking Law Update 2019. My thanks also to the Centre for Banking and Finance Law at the National University of Singapore for their research support, and to Peter Ellinger, Helena Whalen-Bridge and Emma Leong for their helpful comments, as well as participants at the Society of Legal Scholars Annual Conference 2019 where I presented some of the arguments made here.

1 See also the definition of "APP Scam" in the Contingent Reimbursement Model Code for Authorised Push Payment Scams ("CRM Code"), 28 May 2019, under "Definitions and Scope" (https://www.lendingstandardsboard.org.uk/wp-content/uploads/2019/05/CRM-code.pdf). There will no doubt be borderline situations in which the distinction is difficult to draw.

2 See, for example, Booysen "Cheques: To be or not to be?" 2018 *Journal of Business Law* 283.

3 See, for example, *Hall v Fuller* (1826) 5 B & C 750 at 757; 108 ER 279 at 282. In the words of Bayley J: "The banker, as the depository of the customer's money, is bound to pay from time to time such sums as the latter may order. If, unfortunately, he pays money belonging to the customer

ABLU 2019 (Part ONE).indb 1 2019/10/09 9:49 AM

them, *ie* if the payment is unauthorised, the bank cannot deduct the payment from the account and must reimburse the account if it has done so. Conversely, if the bank follows authentic instructions, *ie* if the payment is authorised, it is entitled to deduct the monies from the account.

The instruction to make a payment is the most common instruction given by customers to their banks. Hence, the allocation of risk governing unauthorised and authorised payments, and its development to meet the modern payment environment, is of great practical importance. Because of the dominance of cheques in the last century, the cardinal principles governing the allocation of risk for payment scams were developed in the context of cheques, but these principles have come under pressure because they are not well-suited to govern the payments landscape in the 21st century. The point can be illustrated with reference to an English case from the late 19th century. In *Smith v Mercer*, Heath J stated that "[t]he situation of bankers is most peculiar: they are bound to know the hand-writing of their customers",[4] while Dallas J thought this was a reasonable requirement since the banker would see the customer's writing "every day".[5] The idea that a bank, no longer an individual or group of individuals but a large corporate with many employees, will know the customer's handwriting, and the reasoning that the bank sees the customer's writing regularly, ceased to reflect reality many decades ago. Hence, in *National Westminster Bank Ltd v Barclays Bank International Ltd*,[6] Kerr J rejected the maxim that a bank is under a duty to know its customer's signature. Nevertheless, the risk of paying without the customer's authorisation is on the bank.

As a result of changes in the banking industry the allocation of risk for both types of payment scam has migrated from the common-law position. This chapter primarily focusses on authorised payment scams and considers measures that are being taken to combat them. The chapter starts, however, with a discussion of unauthorised payment scams where the default risk lies on the bank. This risk is now shared by customers in many jurisdictions, the extent of such sharing varying from jurisdiction to jurisdiction. Risk sharing for unauthorised payments started more than a hundred years ago and was driven, at least initially, by banks through their contractual terms. Customers, on the other hand, have considerably less scope to spread the risk of authorised scams to banks and the advent of such risk sharing is, probably for that reason, more recent and can be attributable, at least in part, to the activism of consumer bodies. The discussion looks at how such risk sharing is happening in three jurisdictions – Australia, Singapore and the United Kingdom (UK), with a particular emphasis on the UK, which has since 2016 taken a number of noteworthy steps towards a new allocation of risk for authorised payment scams. These steps have been in response to a growing concern about authorised payment

upon an order which is not genuine, he must suffer, and to justify the payment, he must shew that the order is genuine, not in signature only, but in every respect."

4 (1815) 6 Taunt 76; 128 Eng Rep 961 at 965.

5 *Smith v Mercer* (n 4) 963.

6 [1975] 1 QB 654 at 666.

scams raised by the consumer organisation "Which?",[7] and highlights the effective role that consumer bodies can play to advance the consumer cause. This chapter highlights the trend of risk sharing that is evident in both types of payment scam, and argues that this development is needed for the sustainability of the bank-customer relationship. Because of the historical dominance of banks in the payment services industry, the chapter will focus on banks, but it should be remembered that today there are other providers of such services to whom the discussion may be similarly applicable.

Aside from the above-mentioned jurisdictions, reference will be made in varying degrees to a number of other jurisdictions – namely Canada, Hong Kong, Malaysia, New Zealand and the United States (US). These jurisdictions all inherited the English common law to a greater or lesser extent for historical reasons. The US, while not a commonwealth jurisdiction, also has a common-law link with England for similar reasons. That link is evident in US banking law, although the relationship of bank and customer in the US is now largely governed by the Uniform Commercial Code (UCC),[8] along with other statutes, including the Truth in Lending Act of 1968 and the Electronic Funds Transfer Act of 1978. Unsurprisingly, the various jurisdictions broadly agree on the basic legal principles applicable to the unauthorised and authorised payment scenarios under discussion, although their response to greater risk sharing will vary.

According to a report published by UK Finance, the UK's banking and finance industry body, unauthorised card, cheque and remote banking payments for 2018 accounted for the larger segment of payment scams at £844.8 million, while authorised scams totalled £354.3 million.[9] The first year that figures for authorised payment scams were collected by UK Finance was in 2017 in response to a realisation in the payments industry that such scams are on the rise and need to be tackled.[10] Authorised payment scams are "one of the fastest-growing" frauds,[11] and have been described in the *Financial Times* newspaper as "modern day bank robbery".[12] One of the reasons for the escalation in authorised payment scams is apparently the reasonably effective anti-fraud measures that have been introduced to curb unauthorised payment scams, and the increase in mobile and online banking.[13]

7 Which? super-complaint "Consumer safeguards in the market for push payments" 23 September 2016 (https://www.psr.org.uk/sites/default/files/media/PDF/which-super-complaint-sep-2016.pdf).

8 The Uniform Commercial Code is implemented in the United States by state legislation and hence may vary in its detail from state to state.

9 UK Finance Report "Fraud the facts 2019: The definitive overview of payment industry fraud" 7 (https://www.ukfinance.org.uk/policy-and-guidance/reports-publications/fraud-facts-2019).

10 UK Finance "2018 half year fraud update: Unauthorised payment card, remote banking and cheque fraud and authorised push payment scams" 19 (https://www.ukfinance.org.uk/system/files/2018-half-year-fraud-update-FINAL.pdf).

11 Which? News "Fraud complaints hit record high as banks' new anti-scam measures delayed" 15 May 2019 (https://www.which.co.uk/news/2019/05/fraud-complaints-hit-record-high-as-banks-new-anti-scam-measures-delayed/).

12 Barrett "Push payment scams — or modern day bank 'robbery'" *Financial Times* 12 October 2018.

13 See Payment Systems Regulator "Which? authorised push payments super-complaint: PSR response" December 2016 par 3.45 (https://www.psr.org.uk/sites/default/files/media/PDF/PSR-Which-super-complaint-response-December-2016.pdf).

The extensive adoption of technology in the payments sector in the 21st century has also afforded additional ways for scammers to identify and contact scam victims.

Scammers manipulate their victims, typically remotely, through a technique known as social engineering, which involves deceptive but often highly convincing emails, websites, text messages and phone calls.[14] A scammer's modus operandi has many possible permutations, such as posing as bank staff or law enforcement officers and threatening adverse consequences unless certain payments are made, scaremongering about monies being at risk and needing to be moved to safe accounts, intercepting communications and changing the details of an intended payee's account, and offering goods or services which do not exist. A warning of an actual scam could be found on the website of the Australian Securities & Investments Commission (ASIC) in July 2019, concerning fake invoices purporting to emanate from ASIC relating to the renewal of company names.[15]

2 Unauthorised payment scams

The unauthorised payment scam,[16] as noted above, involves the rogue deceiving the bank into thinking that the bank's customer has authorised a payment from her or his account. According to UK statistics, most payment scams today involve payment cards,[17] of which a substantial percentage will be unauthorised transactions. Before the advent of payment cards, cheques were the primary mechanism. Hence, a proper understanding of the risk allocation for unauthorised payments requires brief reference to the legislation governing cheques. This legislation in all the commonwealth jurisdictions considered here is similar since it was derived from the UK's Bills of Exchange Act 1882 (BEA).[18] Also in the US, the governing UCC article 3 can be traced to the BEA.[19]

As noted earlier, unauthorised cheque payment scams were, and still can be, perpetrated in one of two ways: by getting access to the customer's cheque book

[14] See Hurley in the UK Financial Ombudsman Service *Ombudsman News* August 2018, Issue 145 (https://www.financial-ombudsman.org.uk/publications/ombudsman-news/145/145-ombudsman-focus-fraud-and-scams.html).

[15] See https://asic.gov.au/online-services/service-availability/scams-targeting-asic-customers/.

[16] This section of the paper draws, in part, on a previous publication: see Booysen (2019) 135 *LQR* 437.

[17] See UK Finance Report (n 9) 10, 12 *et seq*. As a result, stronger authentication measures are being introduced for card transactions in the UK: see Recital 6, Commission Delegated Regulation (EU) 2018/389 of 27 November 2017 supplementing Directive (EU) 2015/2366 of the European Parliament and of the Council. In Singapore, see Association of Banks Singapore "FAQs: Enhancing payment card security – new measures to be phased in from 2nd quarter 2010 to 1st quarter 2011" (https://www.abs.org.sg/consumer-banking/consumers/payment-card-security).

[18] See, for example, McKeehan "The negotiable instruments law" (1902) 50 *U Pa L Rev* 561 at 563. Also Holden *The History of Negotiable Instruments in English Law* (1993 Reprint) 201–202; Coleman *The Cheques and Payment Orders Act 1986* (1987) 3; Geva "The modernization of the Bills of Exchange Act: a proposal" (2011) 50 *Can Bus LJ* 26 at 27.

[19] Through the Uniform Negotiable Instruments Act 1896; see Lee and Zinnecker *Payment Systems, Banking and Documentary Transactions* (2003) 5. See also Geva (n 18) 27 n 8.

and forging the customer's, *ie* the drawer's, signature,[20] or by altering the amount and/or the payee of a properly signed cheque. The former will be called a forged cheque and the latter an altered cheque, notwithstanding that the alteration also constitutes forgery. The distinction is made because the BEA treats the two somewhat differently, although the practical effect may be similar. A forged signature is of no effect, hence if the drawer's signature is forged there is basically no signature and the drawee bank has no authority to pay such a cheque. A bank which does pay on such a forged signature is unable to debit its customer's account unless the customer is estopped from denying her or his signature.[21] Such an estoppel arises, in particular, if the customer becomes aware of the forgery but fails to alert the bank. This duty to alert the bank is known as the *Greenwood* duty and originates from a case of the same name.[22] What is significant about the *Greenwood* duty is that the customer has no duty to check bank statements and other notifications to detect forgery, but if she or he discovers it, there is a duty of disclosure.

The BEA has a different provision when the cheque has been materially altered.[23] It provides that a material alteration voids the cheque,[24] which means that the cheque gives no enforceable rights and becomes a "worthless piece of paper".[25] The BEA indicates that altering the sum payable is material but it makes no mention of altering the payee. Case law indicates, though, that changing the payee of a cheque that is not freely transferable[26] does constitute a material alteration.[27] Where the bank has paid on a materially altered cheque, the consequences are similar to forgery – it is unable to debit its customer's account. Here, too, there are common-law qualifications. If the customer has facilitated the alteration, for example, by signing cheques in blank or leaving gaps around the payee or the amount, thus facilitating alteration by the rogue, the customer is in breach of her or his *Macmillan* duty, again derived from a case of the same name.[28] A customer that becomes aware of an altered cheque would also, pursuant to the *Greenwood* duty, be obliged to notify the bank promptly.

[20] It is possible for indorsements to be forged but due to the decline in the use of the cheque as a negotiable instrument, there is likely to be only one signature on a cheque, that of the person making the payment – the drawer.

[21] Bills of Exchange Act 1882 s 24.

[22] *Greenwood v Martins Bank* [1933] AC 51.

[23] See Gleeson *Chalmers and Guest on Bills of Exchange, Cheques and Promissory Notes* (2017) par 3-044, 8-077.

[24] Bills of Exchange Act 1882 s 64(1).

[25] See *Smith v Lloyd's TSB Group plc* [2001] QB 541 at 556–557, 558. There is a qualification in favour of the holder in due course to enforce the cheque according to its original provisions where the alteration is not apparent; Bills of Exchange Act 1882 s 64(1).

[26] A freely transferable cheque is known as a "bearer" cheque under the Bills of Exchange Act 1882: see ss 2, 8.

[27] See, for example, *Smith v Lloyd's TSB Group plc* [2001] QB 541 at 550; *Goldman v Cox* (1924) 40 TLR 744. See also Gleeson (n 23) par 8-081.

[28] *London Joint Stock Bank Ltd v Macmillan & Arthur* [1918] AC 777.

The *Macmillan* and *Greenwood* duties have been widely recognised and adopted in the commonwealth jurisdictions referred to here.[29] These two duties are the common law's fairly modest contribution to spreading the risk of unauthorised transactions so that it is partly borne by the customer.[30] However, the reduction in cheque use around the world and the concomitant migration to other payment methods has had an effect on the utility of the *Macmillan* and *Greenwood* duties. The *Macmillan* duty was formulated specifically in the context of cheques and has considerably reduced relevance beyond a paper instrument since there is little scope for an electronic payment instruction from a customer to be altered. The *Greenwood* duty continues to have utility in the modern payment environment, particularly if it is not limited to forgery of the drawer's signature and embraces fraud in the broader sense. There is case support in New Zealand for such an extension of the *Greenwood* duty to encompass a duty to report detected unauthorised conduct or interference in the account.[31] Presumably courts in other commonwealth jurisdictions would agree that such an extension is uncontroversial and consistent with the rationale of the duty. Even assuming an extension of the duty beyond forgery to fraud more generally, *Greenwood* is still limited in that it only applies when the customer has knowledge of the unauthorised activity – it does not impose a duty on the customer to be vigilant in relation to her or his account.

On the whole, the courts did not develop the *Macmillan* and *Greenwood* duties to spread more of the risk for unauthorised payment scams from banks onto customers – a position that has been criticised as not ultimately working in favour of the customer.[32] There were a few attempts by courts to do so, particularly in Canada and Singapore,[33] but they did not find traction.[34] By contrast, the law in the US developed differently and from an early stage, the US courts recognised that the customer had to show greater vigilance towards her or his bank account,[35] a position confirmed, for example, by the UCC.[36] Despite the reluctance of the

[29] Australia: *National Australia Bank Ltd v Hokit Pty Ltd* [1996] 39 NSWLR 377; Canada: *Canadian Pacific Hotels Ltd v Bank of Montreal* (1988) 40 DLR (4th) 385; Hong Kong: *Tai Hing Cotton Mill Ltd v Liu Chong Hing Bank* [1986] 1 AC 80; Malaysia: *United Asian Bank Bhd v Tai Soon Heng Construction Sdn Bhd* (*United Asian*) [1993] 1 MLJ 182; New Zealand: *National Bank of New Zealand Ltd v Walpole and Patterson Ltd* [1975] 2 NZLR 7.

[30] There have been initiatives by the courts in some jurisdictions to embrace broader common-law duties on the part of the customer, but on the whole, they have not taken hold: see Booysen (n 16).

[31] See *Bank of New Zealand v Auckland Information Bureau (Inc)* [1996] 1 NZLR 420 at 423.

[32] See Booysen (n 16).

[33] *Arrow Transfer Co Ltd v Royal Bank of Canada* (1972) 27 DLR (3d) 81; *Canadian Pacific Hotels Ltd v Bank of Montreal* [1987] 1 SCR 711; *Khoo Tian Hock v Oversea-Chinese Banking Corporation Ltd* [2000] 4 SLR 673.

[34] See, for example, *Canadian Pacific Hotels v Bank of Montreal* (1987) 40 DLR (4th) 385 at 432; *Pertamina Energy Trading Limited v Credit Suisse* [2006] 4 SLR(R) 273 [54].

[35] See, for example, *Frank v Chemical National Bank of New York* (1881) 84 NY 209; *Leather Manufacturers' Bank v Morgan* (1886) 117 US 96; *Potts & Co v Lafayette National Bank* 269 NY 181; *Critten v Chemical Nat'l Bank* (1902) 171 NY 219; *Thomson v New York Trust Co* 293 NY 58; *Maryland Casualty Co v Central Trust Co* 297 NY 294.

[36] See, for example, Uniform Commercial Code § 4-406. See also Facciolo "Unauthorized payment transactions and who should bear the losses" (2008) 83 *Chicago-Kent Law Review* 605 at 607.

ABLU 2019 (Part ONE).indb 6 2019/10/09 9:49 AM

courts to intervene, in all the jurisdictions considered here, there has been a trend towards greater spreading of the risk for unauthorised payments away from the bank and onto the customer in other ways. The means vary from jurisdiction to jurisdiction and take the form of one or more of the following: express contractual terms, legislation and/or codes of conduct.

A prime example of contractual risk spreading is the verification and conclusive evidence clause,[37] the use of which is more than a century old.[38] This express duty expands the *Greenwood* duty and, subject to its drafting, generally requires customers to actively examine statements, transaction notes and other communications from the bank and report any transaction the customer disagrees with. Failure to do so within a specified period of time is stated to render the communication conclusive and binding on the customer. Such clauses are in use in a number of jurisdictions, including Canada, Hong Kong, Malaysia, New Zealand and Singapore. The clauses have been the subject of legal challenge on a number of occasions, particularly regarding their interpretation and pursuant to unfair terms legislation.[39] In Singapore, where the clauses are prevalent, they have been upheld as reasonable in the business context;[40] they have been regarded as reasonable in obiter statements in the consumer context;[41] but they have been rejected where the fraudster was the bank's own employee.[42] The checking of bank notifications by the customer is also supported in Singapore by the financial regulator, the Monetary Authority of Singapore (the MAS).[43] Canada is another jurisdiction where the verification and conclusive evidence clause is widely used and has been enforced on numerous occasions.[44]

Another approach to sharing the risk for unauthorised payments is by statutory provision. Such an approach can be found in Malaysia,[45] the UK and the USA. The provision in Malaysia is quite limited and requires the customer to exercise care to avoid unauthorised cheque payments.[46] It does not extend to other payment mechanisms and may, accordingly, be of decreasing relevance with the trend away

[37] For a fuller discussion of the verification and conclusive evidence clause, see Booysen (n 16).

[38] See the Canadian case of *Columbia Gramophone Co v Union Bank of Canada* (1916) 34 DLR 743.

[39] For example, in Singapore there is legislation that bans certain exemption clauses and requires others to be reasonable: see Unfair Contract Terms Act (Chapter 396, 1994 Revised Edition). This statute is based on the UK's Unfair Contract Terms Act 1977, as is the Unfair Contract Terms Ordinance that operates in Hong Kong.

[40] *Pertamina Energy Trading Limited v Credit Suisse* [2006] 4 SLR(R) 273 [55–71].

[41] *Tjoa Elis v United Overseas Bank* [2003] 1 SLR(R) 747 [92–96].

[42] *Jiang Ou v EFG Bank AG* [2011] 4 SLR 246.

[43] See MAS "E-Payments User Protection Guidelines" s 3.3 (available on the website of the Monetary Authority of Singapore: www.mas.gov.sg).

[44] See Perrett "Account verification clauses: Should bank customers be forced to mind their own business?" (1998–1999) 14 *BFLR* 245. For a more recent summary, see Booysen (n 16).

[45] As noted earlier, the verification and conclusive evidence clause is used by banks in Malaysia, but it seems not to have been as prominent in determining disputes as in jurisdictions such as Canada and Singapore: see Booysen (n 16).

[46] Bills of Exchange Act 1949 s 73A. See also *Malaysia Plastics Sdn Bhd v United Overseas Bank (M) Bhd* [2012] 9 MLJ 336 (High Court, Kuala Lumpur) [53–54], discussed in greater detail in Booysen (n 16).

from cheques, a trend that is being encouraged by the Malaysian government.[47] The statutory approach in Malaysia is combined with the contractual approach and the verification and conclusive evidence clause can also be found in the terms governing bank accounts in Malaysia.[48] Another example of the statutory approach is the UK's Payment Services Regulations 2017 (PSR 2017), a regulation of the European Union that seeks to harmonise the legal framework for payments in EU countries.[49] The PSR 2017 disallow a challenge to a transaction by a customer after 13 months have lapsed.[50] The US has legislation to a similar effect, for example, the UCC disallows the raising of unauthorised transactions a year after the transaction information is available to the customer.[51] As noted earlier, also at common law, the US expects bank customers to take more responsibility to ensure the integrity of their bank accounts.

Many jurisdictions also have banking codes which may set out what is required from a customer in the conduct of the bank account. Subscription to such codes is commonly voluntary and they typically represent soft law, meaning they may not be directly enforceable but the position will vary from jurisdiction to jurisdiction. The Hong Kong Code of Banking Practice supports the practice of transaction verification by the customer,[52] but notes that contract terms should not be used to excuse the bank where its staff are implicated in the unauthorised activity.[53] Australia is an example of a jurisdiction where codes of conduct are a prominent way in which greater diligence from customers is required. The ePayments Code, for example, says the customer is liable for loss arising from fraud where the customer unreasonably delays reporting situations that can result in unauthorised transactions, such as a security breach relating to account access codes or devices.[54] Subscription to the Australian Code is voluntary but widespread, and compliance with the Code is achieved contractually by incorporating its provisions into the contract between the subscribing bank and the customer.[55]

47 See "Strategy to Accelerate Migration to e-Payments in Malaysia" 20 September 2016 (http://pubdocs.worldbank.org/en/935551479484734382/GPW2016-afr-asia-Bokhari-Strategy-to-accelerate-migration-Malaysia.pdf).

48 Discussed in Booysen (n 16) 446.

49 See the Explanatory Memorandum to the Payment Services Regulations 2017 (http://www.legislation.gov.uk/uksi/2017/752/pdfs/uksiem_20170752_en.pdf).

50 Payment Services Regulations 2017 reg 74(1).

51 Uniform Commercial Code § 4-406(f). See also *Associated Home & RV Sales, Inc v Bank of Belen* 2013-NMCA-018 [9].

52 Code of Banking Practice (February 2015) cl 18.4 (http://www.hkma.gov.hk/media/eng/doc/code_eng.pdf). The Code is issued by the Hong Kong Association of Banks and the Hong Kong Association of Restricted Licence Banks and Deposit-taking Companies and endorsed by the Hong Kong Monetary Authority.

53 The verification clause in Hong Kong, as an exemption clause, is subject to controls similar to those found in the UK, such as restrictive interpretation and the Control of Exemption Clauses Ordinance Chapter 71 Hong Kong, which is similar to the UK's Unfair Contract Terms Act 1977; see Fisher and Greenwood *Contract Law in Hong Kong* (expanded 3 ed 2017) par 8.6.4.

54 Australian Securities and Investments Commission (ASIC) "ePayments Code", effective 29 March 2016, cl 11–12. See also Booysen (n 16).

55 ASIC "ePayments Code" (n 54) cl 2.2, see also p 2.

In summary, the risk of unauthorised payments from a bank account is on the bank under the common law, which requires banks to be vigilant to ensure that they have authority to act when making payments out of customers' accounts. That risk has, however, been shared through various means such that bank customers also play an important role in foiling the activities of fraudsters. It is interesting to consider briefly what prompted the risk sharing that is evident in the context of unauthorised transactions. Aside from the US, courts were generally reluctant to extend risk sharing beyond the *Macmillan* and *Greenwood* situations. In the jurisdictions where risk sharing was introduced by banks by contractual means, it was motivated by self-interest. Yet, that self-interest was arguably justified since in other jurisdictions we have seen that parliaments or other bodies, such as regulators, have stepped in to endorse more extensive risk sharing provisions than the courts were prepared to do. At least part of the reason for such intervention is almost certainly the need for greater customer involvement in the new payments environment, where safeguarding by customers of account access codes and other devices, and diligence regarding bank communications, is essential if rogue activity is to be combatted.

3 Authorised payment scams

The authorised payment scam occurs when the customer is tricked into making a payment to someone they would not otherwise pay. The customer's authorisation to the bank is therefore genuine; there is no forgery of the customer's signature or alteration of the payee or the amount, but the customer has been defrauded into giving the authorisation, or possibly been subjected to duress. Such scams are often called authorised push payment scams (APP scams) because they tend to arise in the "push" payment scenario. Push payments are payments in which the paying customer communicates the payment instruction to her or his bank. Most one-off funds transfer payments are push payments. Push payments can be contrasted with payments in which the payment instruction is given or communicated to the payee who, directly or indirectly, hands it to the paying bank. The latter payments are known as "pull" payments – cheques, credit and debit cards and direct debits are the main examples.[56] UK figures indicate that £354.3 million was lost in 2018 from APP scams, of which £228.4 million involved consumers.[57]

Much of the debate about authorised payment scams in the UK has been in the push payment context, but it is important to note that pull payments can also produce authorised scams, as illustrated by the two well-known English "rubber" cases: *Selangor United Rubber Estates Ltd v Cradock (No3)*[58] and *Karak Rubber Co Ltd v Burden (No2)*.[59] The facts of the two cases are similar: the rubber companies were induced by elaborate fraudulent schemes to fund the purchase by rogues of their own shares, the payments being made by cheques. Another example of an

[56] See, for example, Cranston et al *Principles of Banking Law* (2017) 338–339.
[57] UK Finance Report (n 9) 41.
[58] [1968] 1 WLR 1555.
[59] [1972] 1 WLR 602.

authorised pull payment scam comes from *Marfani & Co Ltd v Midland Bank Ltd*,[60] in which a dishonest employee induced his employer to sign a cheque which the employee subsequently deposited into an account he had opened under a false name.

The reason for the focus by "Which?" on APP scams in 2016 was that they considered that the existing regime in the UK governing authorised pull payments and unauthorised payments was more balanced and fairer to consumers, perhaps because banks have incentives to detect such scams.[61] In contrast, consumers bear full liability for authorised push payments and consequently banks do not have the incentives to manage the risk more.[62] This view was based on the prevailing UK regime which was, and still is, largely governed by the PSR.[63] Under this regime, for example, the bank's primary liability for unauthorised payments serves as a protection for customers, and notably, the verification and conclusive evidence clause never became prevalent in the UK. There are other mechanisms in place that protect customers from authorised pull payments by card and direct debit.[64] Cheques are not covered by the PSR. While their declining use reduces their significance, it is noteworthy that statistics produced by UK Finance show an increase in the UK in 2018 of the number and value of cheque frauds compared with 2017.[65]

Authorised payment scams are primarily at the risk of the customer because of the authenticity of the customer's authorisation but, as for authorised payment scams, there is a common-law qualification that spreads the risk – this time to the bank. This spreading of the risk is achieved by the duty on a bank to its customers to exercise care in the conduct of the customer's bank account. The paradigm example of what that duty requires is that the bank must not execute a payment instruction if it has reason to believe that its customer has been tricked or pressurised into giving it. The leading UK case is the Court of Appeal decision in *Lipkin Gorman v Karpnale & Co*,[66] in which a solicitor defrauded his firm by withdrawing monies from its client account. The rogue solicitor was a signatory on the account and hence the payments were authorised, but the question was whether in all the circumstances the bank should have known that the law firm was being defrauded and hence bear the loss. On the facts, the court found in the bank's favour that it had not breached its duty of care. It stressed that banks are not expected to view every

60 [1968] 1 WLR 956.

61 Which? super-complaint (n 7) 8–9.

62 Which? super-complaint (n 7) 8.

63 Initially the Payment Services Regulations 2009, now the Payment Services Regulations 2017. The Payment Systems Regulator is a subsidiary of the financial regulator, the Financial Conduct Authority, and oversees the UK's payments industry: see https://www.psr.org.uk/.

64 For card payments, the contractually based chargeback system enables cardholders to have a card transaction reversed if, for example, they do not receive the goods or services they contracted for, or if a charge was duplicated or made for the wrong amount. The Consumer Credit Act 1974 s 75 also entitles consumer credit card users to reimbursement/compensation in similar circumstances for transactions between £100 and £30 000 in value. For direct debits, a direct debit guarantee protects customers from errors in direct debit payments.

65 UK Finance Report (n 9) 30–31.

66 [1989] 1 WLR 1340. The case went on appeal to the House of Lords, but the duty-of-care point was not raised.

transaction with suspicion but, at the same time, recognised that if facts come to the bank's attention which suggest wrongdoing against their customer, they cannot ignore it.[67] This duty of care is recognised and applied widely in the jurisdictions discussed here.[68]

Authorised payment scams take many forms and can be classified in different ways. Prevalent examples are the sale of goods or services that never materialise,[69] and "malicious redirection", which includes intercepting/redirecting a payment that is due to a creditor by, for example, supplying bank account details that belong to the scammer.[70] The latter scam has targeted, inter alia, conveyancing lawyers and their clients – no doubt because large sums are involved. A specific example of the interception scam can be found in the English case of *Tidal Energy Ltd v Bank of Scotland plc*,[71] ("*Tidal Energy*"). Tidal Energy instructed Bank of Scotland (BoS) to pay €217 000 to a supplier. It transpired, however, that Tidal Energy had been tricked into giving the bank the wrong account number. The payment was made into the rogue's account and in due course the unintended recipient withdrew most of the funds. Tidal Energy sued BoS because, notwithstanding the incorrect account number, they had correctly stipulated the name of the intended payee, *ie* there was a discrepancy between the account number and the named payee. Tidal Energy argued that the bank had not executed the payment instruction as the intended payee had not received the funds. Since the receiving account was held with another bank, Barclays Bank, Tidal Energy were seeking to hold BoS liable for Barclays' failure to detect the discrepancy.[72] The majority of the Court of Appeal ruled in favour of the bank because BoS was entitled to follow banking practice for the high value payment system that was used, and this practice ignores the payee's name and considers only the account number when making a payment. The facts of *Tidal Energy* illustrate the limited scope of the bank's duty of care to protect the customer from such scams. Because the receiving account was held with Barclays, there was nothing to alert BoS that its customer had been deceived.

4 Tackling authorised payment scams

Jurisdictions are starting to introduce measures that address authorised payment scams, particularly from mobile and internet banking platforms. Measures that

[67] *Lipkin Gorman v Karpnale* (n 66) 1356–1357, 1376. See also the English Court of Appeal's dicta in *Barclays Bank Plc v Quince Care Limited* [1992] 4 All ER 363 at 376: "[A] banker must refrain from executing an order if and for as long as the banker is 'put on inquiry'."

[68] See, for example, in Australia: *Varker v Commercial Banking Co of Sydney Ltd* [1972] 2 NSWLR 967 at 976–978; Hong Kong: *Dex Asia Ltd v DBS Bank (Hong Kong) Ltd* [2007] HKCA 516 [23–24]; Singapore: *Hsu Ann Mei Amy v OCBC* [2011] 2 SLR 178.

[69] See UK Finance Report (n 9) 43–44.

[70] See the discussion by forensic fraud investigator, Richard Emery, in the UK Financial Ombudsman Service *Ombudsman News* August 2018, Issue 145 (https://www.financial-ombudsman.org.uk/publications/ombudsman-news/145/145-ombudsman-focus-fraud-and-scams.html).

[71] [2014] 2 Lloyd's Rep 549. See also Booysen "Payment scams: *Tidal Energy v Bank of Scotland* and recent developments" (2018) 33(7) *Journal of International Banking and Financial Law* 405.

[72] Discussed further in Booysen (n 71).

will be considered here include Australia's ePayments Code[73] and Singapore's E-Payments User Protection Guidelines (the E-Payments Guidelines).[74] The bigger focus, however, will be on more general and extensive measures taken in the UK to combat APP scams per se. As noted above, the UK measures are largely in response to activism from the consumer organisation "Which?".[75] The UK's PSR responded to the concerns with a report in November 2017,[76] and numerous measures such as data collection and "confirmation of payee" (CoP) were identified as part of a strategy to tackle APP scams. A particularly significant development since the 2017 report is a voluntary code of practice, the Contingent Reimbursement Model Code for Authorised Push Payment Scams (the CRM Code), that took effect on 28 May 2019. The major retail banks in the UK have all signed up to the CRM Code.[77] The UK's Lending Standards Board will oversee its operation. The CRM Code applies only to UK payments involving consumers,[78] small businesses[79] and small charities.[80]

The CRM Code's provisions can be divided into three groups: those that reduce APP scams, those that improve the response once scams have materialised, and those that reduce the impact of a scam on the victim. The CRM Code identifies conduct standards that signatories, referred to as "firms", will follow with a view to preventing, handling and responding to APP scams. These standards vary according to whether they are the "sending" or the "receiving" firm in the transaction in question. A failure to meet the standards can result in a liability to reimburse the consumer who has been a victim of a scam. Sending and receiving firms are expected to take "reasonable steps" to protect their customers from APP scams, including to detect, prevent and respond to APP scams.[81] In particular, sending firms must collect and analyse data, train staff so that higher risk payments can be identified, warn customers of the risks and give practical advice on reducing the risk. If sending firms have "sufficient concerns", they must delay the payment pending an investigation, and relay the concern to the receiving firm in a specified timeframe.[82] The CRM Code does not say what amounts to a sufficient concern, but firms are encouraged to employ a risk-based approach. Receiving firms must

[73] ASIC "ePayments Code" (n 54).

[74] MAS "E-Payments User Protection Guidelines" (n 43).

[75] Which? super-complaint (n 7); Fulton "Account raid – who bears the cost?" *Fraud Intelligence* 20 October 2016; Kouchikali "All wired up – redress for dishonest account transfers" *Fraud Intelligence* 7 February 2017.

[76] See Payment Systems Regulator "Authorised push payment scams" Report and Consultation CP 17/2 November 2017.

[77] A list of signatories is available on the website of the Lending Standards Board which oversees the implementation of the CRM Code at https://www.lendingstandardsboard.org.uk/contingent-reimbursement-model-code/#firms-that-have-signed-up-to-the-code.

[78] *Ie* individuals not acting in the course of a business or trade: see CRM Code "Definitions and Scope" (n 1).

[79] A business with less than ten employees and an annual turnover and/or balance sheet total not exceeding €2 million: see CRM Code "Micro-enterprise" in "Definitions and Scope" (n 1).

[80] Charities with an annual income of less than £1 million: see CRM Code "Definitions and Scope" (n 1).

[81] CRM Code (n 1) SF(1), SF(2).

[82] CRM Code (n 1) SF1(5).

take reasonable steps to identify accounts that may be used for scam purposes, and conduct customer due diligence before opening accounts, both measures being consistent with the anti-money laundering duties already widely applicable in the financial services industry worldwide, and reminiscent of what is required from a bank when collecting payment of a cheque if it is to avoid liability for the tort of conversion.[83] The receiving firm must also respond to concerns to which they have been alerted by the sending firm.[84] Particular measures that have been identified as useful in the battle against authorised payment scams, both in the CRM Code and elsewhere, are discussed in more detail in the sub-sections that follow.

4.1 *Customer education*

Customers are in the frontline of authorised payment scams and improving customer awareness and detection through education is vital to reducing the success of payment scams. The CRM Code places much emphasis on customer education as an important preventative measure,[85] and requires firms to educate customers about authorised scams and their methodologies, and to provide care following a scam which could include further education and advice.[86] Australia and Singapore also recognise the importance of promoting customer awareness through government agency websites,[87] talks and other publicity.[88] In addition, bank websites in these jurisdictions contain security alerts warning customers and offering tips on how to avoid being a scam victim.[89] Australia's ePayments Code requires banks to give customers clear warning, particularly at the point of making a payment, of the dangers of mistaken internet payments and the importance of using the correct identifying information.[90]

4.2 *Data collection and sharing*

The need for more information regarding APP scams was recognised by the PSR in their 2017 Report.[91] UK Finance started collecting such data in 2017. The CRM

[83] Conversion is a strict liability tort but legislation, for example the UK's Cheques Act 1957 s 4 and Singapore's Bills of Exchange Act (Chapter 23, 2004 Revised Edition) s 86, give the collecting bank a defence if it was not negligent in collecting the payment.

[84] UK Finance is developing Best Practice Standards which provide guidance for firms responding to reports of scams: see CRM Code (n 1) SF2(4).

[85] CRM Code (n 1) GF1, SF1(2). See also Payment Systems Regulator CP 17/2 (n 76) 17.

[86] CRM Code (n 1) GF(3).

[87] See, for example, ASIC's MoneySmart website at https://www.moneysmart.gov.au/scams/banking-and-credit-card-scams; also the Australian Competition and Consumer Commission's website, Scamwatch at https://www.scamwatch.gov.au/.

[88] See, for example, an educational talk organised by Singapore's national financial education programme, Moneysense: https://www.nlb.gov.sg/golibrary2/e/beware-of-scams-26970055.

[89] See, for example, DBS Bank (Singapore)'s website on "Security alerts & news" at https://www.dbs.com.sg/personal/deposits/security-and-you/default.page; also Bankwest (Australia) website titled "Security centre" at https://www.bankwest.com.au/security-centre.

[90] ASIC "ePayments Code" (n 54) s 25.

[91] Payment Systems Regulator CP 17/2 (n 76) 24.

Code boosts this initiative by requiring subscribing firms to collect data on scams and share it with trade bodies, thereby increasing the pool of information.[92] By building up a data record, the payments industry will have a greater appreciation of the size and methods of authorised payment scams. Greater insight into such scams will enable better prevention through the detection of patterns, followed by targeted measures such as warning customers of particular scams. In this way, data collection and customer education are complementary. In Australia, data on payment scams is available from Scamwatch, a website run by the Australian Competition and Consumer Commission. It facilitates reports of scams by the public. Scamwatch announced recently that it expects scam losses to reach record levels in 2019, with a predicted figure of AUD532 million – exceeding half a billion for the first time.[93]

4.3 *Confirmation of payee*

CoP is a preventative measure endorsed by the CRM Code that is pending implementation. CoP involves checking whether the payee's account corresponds with the designated payee in the customer's instruction. This information is generally available only to the receiving bank. As a mechanism, CoP is very effective at reducing the type of scam seen in *Tidal Energy*. It can also detect and prevent some error payments where no scam is involved.[94] CoP is not, however, able to detect scams where the payer is openly dealing with a rogue who plans to abscond without delivering her or his side of the bargain, *ie* the wolf-in-sheep's clothing in a fake goods/services scam or a romance scam. A typical permutation of the latter involves monies being sent to a scammer who was met online in order that they can buy an air ticket to meet the victim who has been manipulated into believing that the scammer is romantically interested in them.

CoP is already used for e-payments in Australia, Singapore and the UK. In Australia, for mobile phone and internet banking payments, customers can create their own payment identity, known as a PayID, which can be given to anyone needing to make a payment to them. The payment identity can be random and may not have an obvious connection to the payee, but the payer receives a notification of who the recipient is before giving final confirmation for the payment.[95] The PayID system thus enables payments to be made without the need for disclosure of the recipient's account details,[96] while the CoP helps prevent unintended payments. In Singapore, a similar system has been established for a funds transfer system,

[92] CRM Code (n 1) GF(2).

[93] Scamwatch "Record losses expected as scammers target Australians" 12 August 2019 (https://www.scamwatch.gov.au/news/record-losses-expected-as-scammers-target-australians).

[94] Apparently only a third of non-scam, error payments are recovered: see Payment Systems Regulator "Confirmation of payee: Response to the first consultation and draft specific direction for further consultation" May 2019 CP 19/4 par 2.120.

[95] For further information on PayID, see the Australian Securities & Investments Commission's MoneySmart website at https://www.moneysmart.gov.au/.

[96] On the other hand, the system has been criticised as enabling identity fraudsters to put names to random numbers: see https://www.theregister.co.uk/2018/02/19/payid_accidental_reverse_telephone_number_lookup/.

PayNow, which enables mobile and online payments using a mobile phone or identity number, or for corporates, a unique entity number.[97] In the UK, such a scheme is operative for its mobile payment system known as Paym. These existing systems require a database with the requisite data to facilitate its operation.

The CoP system proposed by the CRM Code in the UK is a self-standing system that is not embedded in a particular payment system, and it itself performs the task of checking the payee name against the account name when a new beneficiary is created.[98] It then communicates one of three outcomes to the payer: match, close match and no match. The payer is discouraged from proceeding in the case of the latter two but can override and proceed. CoP will apply prospectively to new payees created after CoP becomes operative as well as to any existing payees to whom no payment has ever actually made.[99] Conservative estimates are that CoP will produce a benefit of £145–£150 million per year.[100] The UK's Payment Systems Regulator has been involved in its implementation and has engaged in a series of consultations on its scope and operation.[101]

CoP will be implemented for payments made via two of the UK's payments systems, *ie* the Clearing House Automated Payment System (CHAPS) and Faster Payments Service (FPS), which are the two systems that have had the most APP scams by value.[102] The initial proposal was that all payment services providers (PSPs) processing payments via CHAPS and FPS should participate in CoP, but difficulties, especially for smaller PSPs, have resulted in a more limited proposal that applies only to the six largest UK banking groups, known as the "directed" PSPs. The directed PSPs are involved in 90% of CHAPS and FPS payments so the coverage will still be extensive. The intention is for CoP to be obligatory for directed PSPs, irrespective of their subscription to the Code, and it will apply to business accounts and hence operate more widely than the Code.[103] The details of its operation are still being worked out, with implementation being staggered between the end of 2019 and March 2020.[104]

4.4 *Recovery of mistaken payments*

Unintended payments to scammers are in principle recoverable on the basis of unjust enrichment.[105] The claimant will ordinarily be the person, either the bank or

[97] See the PayNow fact sheet at https://abs.org.sg/docs/library/paynow_factsheet.pdf. The confirmation of payee facility is supported by the MAS "E-Payments User Protection Guidelines" (n 43).

[98] I am grateful to Mr Brian Cunnington at Pay.UK for answering my questions relating to confirmation of payee.

[99] See Pay.UK "Frequently asked questions" (https://www.wearepay.uk/10-common-confirmation-of-payee-questions-answered/).

[100] Payment Systems Regulator CP 19/4 (n 94) 1.9, 2.33–2.35.

[101] See Payment Systems Regulator CP 19/4 (n 94).

[102] For further discussion of which systems were chosen and why, see Payment Systems Regulator CP 19/4 (n 94) 2.54–2.56, 3.5.

[103] Payment Systems Regulator CP 19/4 (n 94) 2.72, 2.89, 2.119, 3.7.

[104] Payment Systems Regulator CP 19/4 (n 94) 3.1, 3.6–3.7.

[105] See *Kelly v Solari* (1841) 9 M&W 54; *Barclays Bank Ltd v W J Simms Son & Cook (Southern) Limited* [1980] 1 QB 677; *Kleinwort Benson Limited v Lincoln City Council* [1999] 2 AC 349.

customer, who otherwise will bear the loss. The legal principles entitling a claimant to the recovery of mistaken payments have received much academic attention.[106] Australia, England and Singapore take a similar approach which asks whether the defendant was unjustly enriched at the claimant's expense – a test that is likely to be satisfied in the context of a scam payment, although each case will ultimately depend on its facts. This section will focus not on the legal requirements for recovery of mistaken payments, but on the more practical measures that can be adopted to assist the customer that has the legal right of recovery following an authorised payment scam. These measures are important and can make a significant difference because even when the right to recovery is legally clear, the exercise of the right may be regrettably elusive.[107]

One practical problem is that legal action may not be viable due to the expense and time it will take to reach a conclusion, particularly when a relatively small amount is involved. In other words, a non-litigious solution is needed which generally requires the assistance of the banks involved. Another problem in the scam context is that at the earliest opportunity, the scammer is likely to withdraw the funds and abscond – hence prompt action by the relevant banks is critical to recovery. This point is illustrated by the facts of *Tidal Energy.* Tidal Energy alerted the sending bank, BoS, to their error. BoS contacted the receiving bank, Barclays. Barclays, however, would not intervene without a court order directing them to freeze the funds, and a few hours later the monies were withdrawn.[108] The receiving bank is in a difficult position when it receives an alert of monies paid by mistake as it holds account funds on behalf of its customer and must ordinarily meet the customer's demands for withdrawal.[109] It may not be able to assess the merits of a mistake claim and will risk breach of contract with its customer if it freezes the funds. As such, the development of practices by a regulator or industry body to guide or direct banks on how to react to scam payment alerts is needed. In *Tidal Energy*, the discrepancy between the account name and number was a strong clue that the payment was unintended, and it is unfortunate that Barclays did not feel able to assist and prevent the loss of a substantial sum of money.

Australia,[110] Singapore[111] and the UK [112] have developed measures to assist in the recovery of mistaken payments, particularly e-payments. Where the payer knows the mobile number to whom the money was transferred, it is generally recommended that they call the number to alert the recipient and request its return. In addition, customers are advised to alert the sending bank promptly of the error.

106 For a recent theoretical discussion, see Penner "We all make mistakes: A 'duty of virtue' theory of restitutionary liability for mistaken payments" (2018) 81(2) *MLR* 222.

107 See, for example, Leow "Enforcing unjust enrichment rights: The recovery of mistaken payments in practice" 2018 *SJLS* 22.

108 [2013] 2 Lloyd's Rep 605 [9–10].

109 See, for example, *London Joint Stock Bank Ltd v Macmillan & Arthur* [1918] AC 777 at 814, 824.

110 See ASIC "ePayments Code" (n 54) par 24–34.

111 See MAS "E-Payments User Protection Guidelines" (n 43) s 6.

112 See the process applicable to the Faster Payments system, which processes payments up to £250 000 in a short timeframe, available under "Consumers" at http://www.fasterpayments.org.uk/. See also CRM Code (n 1) SF2(5).

After notification to the sending bank, a common feature of the recovery measures in the above jurisdictions is that both sending and receiving banks are expected to take steps to assist in the recovery of the monies. If the claimant is not happy with the action taken, they can consider taking the matter to the financial dispute resolution mechanism available in their jurisdiction.[113] Expectations from the sending bank are that they will investigate the claim by obtaining particulars from the claimant and notifying the receiving bank. The receiving bank must communicate with the recipient and seek the return of the monies. In all three countries, it is an offence to appropriate monies received in error, which information should be relayed to the unintended recipient.

The ePayments Code in Australia has particularly notable provisions to assist with the return of the monies where the receiving bank is satisfied that the payment was made in error:[114] if the claimant alerts the sending bank within ten business days of the payment and the money is still in the recipient's account, the receiving bank must return the monies to the sending bank; if the alert is given after ten days but within seven months, the receiving bank must freeze the funds and give the recipient ten business days to prove their entitlement to the monies, failing which it must be returned; if notification is after seven months of the payment, the monies can be returned only if the recipient agrees. One of the problems noted earlier is how the receiving bank can satisfy itself that the monies were paid in error. This is where the investigation by the banks is important and it seems that a discrepancy between the account name and number can be taken as prima facie evidence of an erroneous payment[115] – contrast the reaction of Barclays Bank in *Tidal Energy*.

4.5 *Reimbursement*

Assuming that there is no recovery of the scam payment, one of the most significant provisions of the CRM Code is for reimbursement of APP scam victims by the firms involved. The default position is that the customer will receive reimbursement unless she or he is guilty of one of a number of defaults which would have helped to prevent the scam had they been followed.[116] The defaults which may block reimbursement include a customer's failure to respond appropriately to clear warnings on the payment platform when creating a new payee, altering a payee, making the payment, or after receiving notification of a discrepancy after CoP. Customers who are particularly vulnerable because of their personal circumstances, the nature of the scam, or its effect on them, should be reimbursed regardless of such defaults. However, in making a reimbursement decision, firms can take into

[113] See, for example, the Financial Ombudsman Service Australia "Fact sheet: Mistaken internet payments" 2 (https://www.fos.org.au/custom/files/docs/fact-sheet-mistaken-internet-payments.pdf).

[114] ASIC "ePayments Code" (n 54) par 28–30.

[115] See Financial Ombudsman Service Australia "Fact sheet" (n 113) 1.

[116] CRM Code (n 1) 8, R1–R2.

account the customer's conduct in the reporting of the scam, such as dishonesty or obstructing the investigation.[117]

A decision on reimbursement must be made without delay and generally within 15 business days of the scam being reported, and the customer must be informed of her or his right to take an adverse decision to the Financial Ombudsman Service. If the decision is to reimburse, payment must be made without delay.[118] The CRM Code provides guidance on how any reimbursement should be shared between firms, the idea being that reimbursement is borne by the firm that has breached the CRM Code, and if both, shared.[119] Reimbursement where there has been no beach of the Code by either of the firms involved or by the customer, a so-called no-blame situation, is being funded until December 2019 through contributions from PSPs. A more enduring funding plan is currently being devised to take over from January 2020.[120] There are different ways of financing and structuring a reimbursement fund, including a mandatory modest levy on customers, or on both customers and the payments industry, or an optional insurance scheme for customers.

It will be interesting to see how the reimbursement scheme unfolds in practice. The reimbursement measures, particularly the default position which favours reimbursement, go beyond legal notions of fault-based liability and arguably tilt the balance too far in favour of the gullible. They give payers greater protection when making payments than they have in other areas, such as when buying products or services they do not need, or which offer poor value. One might question why, for example, someone who falls for a romance scam should be entitled to protection from their own folly at the expense of those who fund the scheme. The measures are even more generous to vulnerable customers, who are widely defined. There are evident dangers of customers becoming complacent and being less careful about parting with their money. The availability of reimbursement may also give rise to some difficult questions about whether a loss is attributable to an APP scam or whether it's simply a bad bargain for which there is no protection under the CRM Code.

5 Conclusion

The measures discussed above are a significant qualification to the common-law risk allocation for authorised payment scams. As noted earlier, the bank's duty of care not to make payment when they have reasonable grounds to believe that the customer does not intend to benefit the payee is the only qualification at common law to the customer's default liability for an authorised payment. This duty of care, at least as it has been applied to date, is unlikely to be triggered in many situations in which authorised scam payments arise, and yet there are steps banks can take

[117] CRM Code (n 1) 8, R2.

[118] CRM Code (n 1) 9, R3–R4. For more detailed provisions on dispute resolution, see 11–13, DR 1–DR3.

[119] CRM Code (n 1) 10–11, ALL1–ALL4.

[120] Payment Systems Regulator "PSR welcomes major boost in protection from authorised push payment scams, to begin in May 2019" 28 February 2019 (https://www.psr.org.uk/psr-publications/news-announcements/psr-welcomes-major-boost-in-protection-from-APP-scams).

to reduce scam payments. The UK's CRM Code, with its detailed provisions on customer education, detection, prevention, recovery and reimbursement, is more proactive and detailed than the common law. The standards stipulated by the Code are also now likely to start informing the content of the common-law duty of care, particularly for Code signatories but over time also non-signatories.

An important question which has not yet been raised is why banks should be responsible for addressing and bearing some of the burden of the growing authorised payment scam problem. After all, as Tomlinson LJ noted in *Tidal Energy*, in an authorised payment scam it is the customer that falls for the rogue's ploy, not the bank.[121] There are a number of reasons why banks should share more of the burden of tackling authorised payment scams:

- Banks are better positioned to take measures that can make a significant difference in combatting authorised payment scams.[122] For example, CoP and recovery of mistaken payments are valuable measures that require action from the banks involved in the payment; data collection leading to intelligence that can be effectively harnessed to combat future scams is also within the banks' domain but beyond the reach of the customer. As the earlier discussion shows, customers have been co-opted into the fight against unauthorised payments by requiring greater vigilance and care from customers in relation to their accounts. One of the justifications for more responsibility on the customer in the unauthorised payment scam scenario is that the customer is uniquely placed to contribute to the fight against unauthorised transactions. This view reflects the concept of mutuality or reciprocity in the bank-customer relationship.[123] Mutuality is similarly needed in the fight against authorised scams, and banks should play their part and do what they reasonably can to help reduce the incidence and effect of such scams, which are primarily at the customer's risk. No doubt measures such as customer education, data collection and CoP involve an increase in costs, but it is reasonable to expect banks to incur such costs to provide a safer system that will reduce losses to society as whole. Practically, such costs are, in any event, likely to be passed on to customers through bank charges.
- A second reason why banks should share more of the burden of tackling authorised payment scams was aired in the *Financial Times* article mentioned earlier and springs from the connection between authorised payment scams and the mechanisation of bank systems. Unavoidably, the changes in the way in which banks engage in banking business have led to reduced human contact and the reduction of traditional banking settings, thus increasing the risk of deception by remote means.[124] A hundred years ago, a customer would have conducted almost all of her or his banking business by going

[121] [2014] 2 Lloyd's Rep 549 [40].
[122] An argument made in the Which? super-complaint (n 7) 6–7.
[123] See, for example, Booysen "Banks and exclusion of losses for forged cheques – is it reasonable?" (2010) 252 *Singapore Academy of Law Journal* 22 at 28.
[124] Barrett (n 12).

into an established and familiar branch building and interacting with familiar persons, which is considerably more difficult to mimic than the remote contact we tend to have with our banks today. The common-law rules governing risk allocation for payment scams were developed in an environment that is much changed, and they do not respond adequately to this new banking environment. The extensive use of technology to provide account services to customers has allowed for cost savings through the reduction of staff and closure of bank branches. It is not unreasonable, therefore, that banks contribute to the negative effect of these changes on their customers and address the vulnerabilities that have arisen.

- A third reason is that it is vital for banks to retain, and one might say restore, the trust that is so essential to the financial system, and which has been eroded since the global financial crisis.[125] The UK's Financial Ombudsman Service has reported that complaints about the handling of fraud and scams by banks in 2018–2019 saw a 43% increase from the previous year,[126] a development that could negatively impact trust. Since payment services are so widely and extensively used by customers, it is essential that they should feel confident about, and secure in using, the payment services on offer. For this reason, banks should go to greater preventative lengths to protect customers from scammers than was previously required.

As the above discussion shows, the evolution of risk allocation for unauthorised and authorised payment scams reveals a similar trend. Under the common law, the bank is the primary risk bearer for unauthorised payments, subject to limited common-law duties imposed on the customer to reduce the risk to the bank. The common law has been modified in all the jurisdictions mentioned here, and customers now bear more risk for unauthorised transactions than they did under the common law. Where transactions are authorised, the customer is the primary risk bearer under the common law, subject to the duty imposed on the bank to exercise care in executing payment instructions. This study shows that the common-law allocation of risk for authorised payment scams is similarly undergoing modification such that banks are now expected to do more than the common law has traditionally required from them to address authorised payment scams. It has been argued here that this recalibration of the risk for scam payments, unauthorised and authorised, is a necessity in the modern banking environment.

[125] The issue of trust remains critical to both consumers and financial regulators since the global financial crisis. See, for example, White "British public don't trust banks 10 years after crisis, survey finds" *Reuters, UK* 16 August 2018 (https://uk.reuters.com/article/uk-britain-banks/british-public-dont-trust-banks-10-years-after-crisis-survey-finds-idUKKBN1L11EL); speech by Ravi Menon, Managing Director, Monetary Authority of Singapore "Strengthening trust in finance" Opening Address at Symposium on Asian Banking and Finance, 3 June 2019.

[126] See Which? News (n 11).

Deposit insurance in Namibia and South Africa: Pricing its necessity and design

DUNIA ZONGWE*

1 Introduction

The collapse of SME Bank two years ago in Namibia is reportedly "the worst economic disaster since independence".[1] It also revealed that no explicit deposit-insurance scheme (DIS or DI scheme) exists in the country.[2] Depositors can only claim N$25 000.00.[3] (Namibia has pegged its currency, the Namibian Dollar, one-to-one to the South African Rand.)[4] Beyond that amount, the depositors enjoy no preference vis-à-vis the bank's non-deposit creditors.

If Namibia had had a deposit insurance law at the time the SME Bank went bankrupt, the Bank's 18 000-odd depositors would not have lost a fortune. And the fact that the depositors consisted of small and medium entrepreneurs aggravates the loss suffered by corporate Namibia and the economy in general.

Deposit insurance plays a key role in fending off severe banking crises. Although they came out of the narrow focus they used to hide in before the 2008 global

* Associate Professor and Head of the Mercantile Law Unit, Department of Legal Studies, Walter Sisulu University. I appreciate the time that Charl Hugo and two anonymous reviewers spent reading earlier drafts of this chapter and offering useful comments and suggestions. This chapter reflects events and the law as in September 2018.

1 *Bank of Namibia v Small & Medium Enterprises Bank Ltd (3)* 2018 1 NR 193 (HC) (hereinafter referred to as "*SME Bank* case (final order)"). Judge Hannelie Prinsloo did not portray SME Bank's downfall accurately. Actually, by 2018, when she handed down the judgment in which she portrayed SME Bank's collapse as "the worst economic disaster since independence", experts were already pointing out that the recession that started in 2016 marked the worst economic crisis since independence. Notably, Namibian Finance Minister, Calle Schlettwein, admitted as early as in December 2016 in his mid-year budget review speech that "[t]he Namibian economy has never before been in such a precarious situation". See also Nakashole "Schlettwein speaks on Moody's rating" *The Namibian* 12 December 2016.

2 Zongwe and Katjaimo "SME Bank closure exposes cracks in the banking system" *The Namibian* 1 December 2017 at 11; International Monetary Fund *Namibia: 2017 Article IV Consultation* (2018) 17 (remarking that the recent liquidation process of the SME Bank has shown that the crisis management and resolution framework of Namibia needs "substantial improvement").

3 On 15 June 2017, the Governor of the Bank of Namibia determined that if a bank is wound up (the liquidators of) that bank must pay deposit liabilities up to an amount of N$25 000 to each depositor. See Bank of Namibia: Determination under the Banking Institutions Act, 1998 (Act 2 of 1998) as amended: Priority of Claims in the Event of Winding-up of a Banking Institution or Controlling Company (*GG* 158) s 6(2)*(c)*.

4 Namibia belongs to the Multilateral Monetary Area, an arrangement in terms of which Namibia, Lesotho and Swaziland operate a fix peg against the South African Rand, without restricting capital flows in the Area. See Bank of Namibia *Challenges of Monetary Policy for Namibia within the Common Monetary Area (CMA) Agreement* (2000) 4.

ABLU 2019 (Part ONE).indb 21 2019/10/09 9:49 AM

recession and although many countries embraced them, deposit-insurance schemes are still poorly understood. As a result, these schemes are “ill-conceived”[5] and “inadequately designed”.[6]

At the outset, this chapter hammers home the point that Namibia needs a DIS, and hence the real question is how Namibian policy-makers should conceive and design it. While DI schemes can shield depositors from losses, not all of them can achieve the same results. What, however, affects all such schemes – and all insurance schemes for that matter – boils down to moral hazard.

Indeed, a notable body of research shows that, lying at the heart of insurance design, is the question of pricing bank risks[7] – the question as to whether a deposit insurer should adopt fixed-rate or risk-adjusted premiums, or whether it should combine the two systems. This chapter favours risk-based premiums as the policy choice that strikes the right balance between preventing bank runs and discouraging inordinate risk-taking. More broadly, this choice promotes financial stability (by forestalling bank runs) and economic efficiency (by imposing market discipline).

To chart the way on how best to tackle moral hazard through deposit insurance, this chapter draws on the experience in South Africa, Namibia’s much larger neighbour. Like South Africa,[8] Namibia does not yet have a DIS, and South Africa’s stance may have influenced Namibia’s delay in taking up a DIS.

The remainder of this chapter is divided into six sections. The first part dissects the idea of deposit insurance. The next section examines the philosophical foundations of deposit insurance. In the process, it emphasises the vital significance of these schemes for the banking sector and the national economy as a whole. The third part surveys and assesses various DIS designs across the world, searching for the design that best suits Namibia’s unique circumstances. The fourth section focuses on the (lack of) deposit insurance in Namibia, using the downfall of SME Bank as a case in point. It also seeks a formula for a Namibian DIS by going through the draft Designated Institutions Resolution Bill, which introduces a DIS in South Africa. The fifth section evaluates risk-pricing and, in the concluding paragraphs, the chapter proposes one particular deposit-insurance model for Namibia.

Even though a DIS plays a huge part in a country’s financial safety net, to date no one has published any in-depth study on deposit insurance in Namibia. The fact that the banking industry operates without a DIS has rendered Namibia vulnerable to systemic bank failures. The government is working on a deposit insurance bill,

[5] *Eg* Blad “Searching for the ideological foundations of Federal Deposit Insurance” 1990 *Annual Review of Banking Law* 533 at 534.

[6] Demirgüç-Kunt, Kane and Laeven “Deposit insurance database” IMF Working Paper 14/118 (2014) 15 (hereinafter “IMF Database”).

[7] *Eg* Suphap “Toward effective risk-adjusted bank deposit insurance: A transnational strategy” 2004 *Columbia Journal of Transnational Law* 830; Blad (n 5) 533–557; Macey, Miller and Carnell *Banking Law and Regulation* (2001) 258; Shaffer “Deposit insurance pricing: The hidden burden of premium rate volatility” 1997 *Cato Journal* 81; Keeley “Deposit insurance, risk, and market power in banking” 1990 *The American Economic Review* 1183; Horvitz “The case against risk-adjusted deposit insurance premiums” 1983 *Housing Finance Review* 253.

[8] Republic of South Africa: Department of National Treasury *Strengthening South Africa’s Resolution Framework for Financial Institutions* (2015) 33 (hereinafter “National Treasury”).

but given the absence of any careful study or report on deposit insurance, the general public and key stakeholders do not know the policy options behind and the ideological bases of these schemes. In South Africa, by contrast, the National Treasury and the Central Bank have published three policy documents for public comment.[9] These documents have shaped the bill currently being drafted.[10]

This chapter thus aims to correct the situation and commence the debate on deposit insurance in Namibia. It focuses on national statutory deposit-insurance schemes for commercial banks.

2 Basic deposit insurance

2.1 *Insurance, risk and insurable risks*

To conceptualise deposit insurance, one needs to understand insurance in the first place. By means of insurance, one person (the insured or policyholder) protects himself, herself or itself against the occurrence of certain risks by passing them on to another person (the insurer) in exchange for the payment of a fee (the premium).[11] "Risk" refers to the probability that a peril or danger (for example, losing property rights or some other valuable interests permanently) may occur and cause harm to the person exposed to it.[12] Insurable risks include death, grievous injury, unemployment or loss of income, fire, floods, or loss of bank deposit.[13]

Insurance benefits policyholders in two ways: it spreads losses across institutions or policyholders (cross-sectional smoothing) and over time (intertemporal smoothing).[14] In the banking sector, this smoothing presumably works to the better advantage of small depositors.[15] Nonetheless, maintaining a DIS is costly. And, in fact, depositors share the burden of that insurance scheme since banks pass a portion of the premiums they pay to the deposit insurer on to the depositors themselves.[16]

Compared to other insurance types, deposit insurance has emerged fairly recently. It seems to have originated in the United States of America (US).[17] Economists,

[9] Republic of South Africa: Department of National Treasury *A Safer Financial Sector to Serve South Africa Better* (2011) (hereinafter "National Treasury *Safer Financial Sector*"); National Treasury (n 8); South African Reserve Bank *Designing a Deposit Insurance Scheme for South Africa: A Discussion Paper* (2017) 2 (hereinafter "SARB").

[10] See International Monetary Fund *South Africa: 2017 Article IV Consultation* (2017) 72.

[11] See Davis *Gordon and Getz on the South African Law of Insurance* (1993) 79; Sutherland and Van der Bijl "The law of insurance" in Scott and Cornelius (eds) *The Law of Commerce in South Africa* (2014) 303–304; Birds *Modern Insurance Law* (1993) 10–11; Van Niekerk *The Development of the Principles of Insurance Law in the Netherlands From 1500–1800* (1998) 11–12; and McCall "Insurance" in *The New Palgrave Dictionary of Economics* (2018) 6611.

[12] See Sutherland and Van der Bijl (n 11) 303; Zongwe "Conjuring systemic risk through financial regulation by SADC central banks" 2011 *SADC Law Journal* 99 at 103.

[13] Sutherland and Van der Bijl (n 11) 303.

[14] Shaffer (n 7) 82.

[15] *Ibid.*

[16] *Ibid*.

[17] Calomiris and White "The origins of federal deposit insurance" in Goldin and Libecap (eds) *The Regulated Economy: A Historical Approach to Political Economy* (1994) 145–188.

bankers and law-makers have debated deposit insurance since 1829.[18] However, in those earlier days, the federal state had not yet rolled out any deposit insurance at a national scale. Instead, individual businesses or sub-national governments would issue such policies or protections.[19]

2.2 *Bank runs*

Through the ages, states have used national DI schemes, whether explicit or implicit, to avert the risk of systemic bank runs. An imitative run, "a run on the bank" or simply a "bank run" usually happens when the failure of one bank prompts the depositors of other banks, especially those who are not insured, to fear that their banks will fail as well and to withdraw their deposits.[20] This mass hysteria and mass withdrawal of deposits bring on a liquidity crisis, which in turn causes banks to fail.

Bank runs may be triggered by events other than the failure of one or more banks. In the case of the Namibian SME Bank, investors rushed to withdraw their money following news that the central bank took over SME Bank.

The threat of runs leads bankers to under-produce liquidity.[21] Bank runs nullify the banking sector's special contribution to economic activity.[22]

2.3 *Implicit and explicit deposit insurance*

Deposit insurance is "implicit" when it emerges from unwritten promises and other non-binding obligations.[23] Implicit deposit-insurance schemes always exist, irrespective of whether or not a country has an explicit scheme and irrespective of how much coverage the explicit scheme extends.[24] Following its implicit insurance, the South African government has compensated in the past the depositors of failed commercial banks on a case-by-case basis, which meant that ultimately taxpayers had to bear the costs of bank failure.[25] After the South African finance minister placed VBS Mutual Bank under curatorship in March 2018, the South African Reserve Bank (SARB) secured guarantees later in July from the National Treasury of up to R100 000 for each depositor to compensate the bank's depositors.[26] Here again, the South African government intervened without any explicit DIS, thereby draining taxpayers' money.

18 Hust "Federal deposit insurance and some of its constitutional aspects" 1939 *George Washington Law Review* 595 at 596.

19 See Hust (n 18) 596.

20 Scott "The reduction of systemic risk in the United States financial system" 2010 *Harvard Journal of Law and Public Policy* 671 at 673; Zongwe (n 12) 107.

21 Carns "Should the $100,000 deposit insurance limit be changed?" 1989 *FDIC Banking Review* 13.

22 Carns (n 21) 13.

23 Suphap (n 7) 833.

24 See also IMF Database (n 6) 4.

25 SARB (n 9) 2.

26 For background information on the VBS Mutual Bank, see Motau *The Great Bank Heist: Investigator's Report to the Prudential Authority* (2018).

In contradistinction, deposit insurance becomes "explicit" when it is provided for in a legally binding document such as a statute, regulations, or a written contract.[27] In practice, countries start with an implicit DIS before graduating to an explicit DIS.[28]

The number of countries that have opted for an explicit DIS has increased significantly. Before the 2008 global financial recession, only about 84 countries subscribed to a DIS; after the recession, more than 110 countries have embraced some or other form of deposit insurance.[29] Some jurisdictions have more than one DIS. In these countries, several schemes, whether private or public, exist alongside a national statutory deposit-insurance scheme.

DI schemes are especially common among high-income countries.[30] Almost all countries in Europe have adopted a DIS.[31] By contrast, with about a 30% adoption rate, most low-income countries have no explicit DI schemes.[32] In Africa, particularly, roughly four-fifths of the countries had no DIS by 2013.[33]

Explicit DI schemes spare the state and taxpayers from having to intervene in major banking crises to save banks through huge bail-outs. However, the mere fact that a government in a certain country issues temporary blanket guarantees does not necessarily entail that the country concerned has an explicit DIS.[34]

2.4 *The "paybox" function of deposit insurance*

In essence, the role of an explicit DIS is to pay out the value of a customer's deposit up to the statutory limit if the customer's bank fails (*ie* becomes insolvent), also known as the "paybox" role of a DIS.[35] The SARB explains the function of a DIS as follows:[36]

> "A DIS provides a mechanism to ensure a pre-planned, orderly and efficient provision of protection rather than an unprepared scrambling for funds, haphazard policy decisions made under pressure and/or disorderly and non-transparent compensation arrangements."

However, the bulk of DI schemes perform other roles, over and above their core paybox function.[37] Thus, some deposit insurers, in addition to pay-outs, wind up

[27] See Suphap (n 7) 833.
[28] *Ibid.*
[29] IMF Database (n 6) 11.
[30] *Ibid.*
[31] *Ibid.*
[32] IMF Database (n 6) 12.
[33] IMF Database (n 6) 32.
[34] See IMF Database (n 6) 4.
[35] See IMF Database (n 6) 6.
[36] SARB (n 9) 2.
[37] IMF Database (n 6) 12 (indicating that about 57% of DI schemes in *ex ante* schemes play other roles in addition to their main paybox role, including the task to reduce losses to the deposit-insurance fund).

failed banks, supervise or license banks, or act as macro-prudential[38] regulators.[39] And, in some countries, banking laws may empower insurers to minimise losses to the taxpayer by allowing them to employ a series of measures, such as replacing negligent bank managers or setting up bridge banks.[40]

3 The philosophical foundations of deposit insurance

3.1 *The necessity of deposit insurance*

The importance of a DIS cannot be overstated. The future health of the entire financial services industry hinges on the integrity of the deposit insurance system.[41] The World Bank advises governments to adopt DI schemes as part of its adjustment programs while the International Monetary Fund (IMF) offers technical and policy-based advice on how they can design DI schemes.[42] Moreover, elected local officials tend to favour DI schemes to avoid the political backlash that accompanies major banking disasters.[43]

In particular, the IMF recommended in its latest country reports that Namibia and South Africa enact a DIS. It proposes that Namibia develop a full crisis management and resolution framework featuring notably a DIS that meets international norms.[44] Similarly, the IMF advises South Africa to introduce an explicit, *ex ante* (*ie* pre-funded), privately funded DIS, with a back-up credit line from the National Treasury.[45]

Establishing a DIS becomes necessary when governments' ability and willingness to pay for the cost of bank failures diminish.[46] This was especially the case in Namibia after the economy slipped into recession in the closing months of 2016.[47] The Namibian government apparently lost its ability to rescue SME Bank. In these circumstances, uncertainty exists about which depositors to compensate, how much to pay out, and where the funding will come from.[48] It is these uncertainties that a DIS aims to remedy by setting up a pre-planned, predictable and efficient framework for liquidating failed banks and compensating their insured depositors.

38 Regulations or policies are "macro-prudential" when they encompass all the rules that a country has adopted to protect its banks from systemic risk or systemic failure.
39 IMF Database (n 6) 6, 12.
40 *Ibid*.
41 Williamson "Regulatory theory and deposit insurance reform" 1994 *Cleveland State Law Review* 105 at 106–107; Granatstein "Deposit insurance reform in Canada" 1986 *Manitoba Law Journal* 45.
42 Miller "Deposit insurance for economies in transition" 1997 *Yearbook of International Financial and Economic Law* 103 at 104.
43 Miller (n 42) 104.
44 International Monetary Fund (n 2) 17, 68 and 69.
45 International Monetary Fund (n 10) 15–16, 21 and 72.
46 See SARB (n 9) 2.
47 See "Namibia goes into 'technical' recession" *The Namibian* 16 December 2016 at 15.
48 See SARB (n 9) 2.

ABLU 2019 (Part ONE).indb 26 2019/10/09 9:49 AM

Policy goals of deposit insurance

In spite of the widespread adoption of DI schemes after the 2008 global financial meltdown, most people do not fully understand them. This knowledge gap has one likely cause: the first nationwide explicit DIS set up by a piece of legislation was passed in a hurry to stem the Great Depression in the 1930s,[49] arguably the worst economic crisis this world has ever suffered. In that hurried process, the US Congress never got to resolve the ideological foundations of the DIS, nor did it disentangle that law's internal contradictions.[50]

When US law-makers enacted the first ever nationwide explicit DIS law in 1933 via the Banking Act (the "Glass-Steagall Act")[51] they targeted three goals: to protect small depositors, to restore confidence in the financial system,[52] and to save the country from systemic banking failure.[53] How the DIS would attain these goals, however, remained unclear.[54] To make matters worse, some of these goals appeared to contradict one another[55] – a point which this chapter revisits below.

To begin with, a DIS protects small, less financially sophisticated depositors – the class of depositors who are often unable to assess the risks of different, competing depository institutions.[56] By the way, critics have blamed some DI schemes precisely for their insistence on protecting ordinary depositors. They level criticism at this tight focus on small investors because they view it as undermining the other goals of these schemes.[57] These other goals include shielding the banking sector and the economy from systemic risks.

Murton submits that the state insures deposits chiefly because it wishes to promote financial stability by stemming bank runs.[58] Notably, during the 2008 global financial recession, bank deposits that were not insured experienced massive withdrawals.[59] In other words, the lack of deposit insurance led to widespread bank runs on uninsured deposits.

One big (non-official) reason for passing the Banking Act of 1933 lies in the perceived immorality of the US government during the Great Depression. The

[49] Blad (n 5) 534.

[50] Blad (n 5) 533.

[51] Banking Act of 1933 (United States) ch 89, 48 Stat 162 (codified as amended throughout 12 USC chs 2,3 and 6 (1988)). See also Driscoll "Deposit insurance in theory and practice" 1988 *Cato Journal* 661 at 663*ff*.

[52] Balderston "Statement on proposed changes in federal deposit insurance" 1963 *Federal Reserve Bulletin* 626 (affirming that one of the major purposes of the US federal insurance for bank deposits is to maintain public confidence).

[53] Blad (n 5) 537 (recounting how the US Congress was quite open in stating the goals for the federal deposit insurance system: to protect small depositors, to restore public confidence in the banking system, and to save the country from a systemic banking failure).

[54] Blad (n 5) 537.

[55] *Ibid*.

[56] Suphap (n 7) 834. See Economides, Hubbard and Palia "Federal deposit insurance: Economic efficiency or politics" 1999 *Regulation* 15 (arguing that, in introducing federal deposit insurance, US law-makers in fact aimed at protecting small banks).

[57] Blad (n 5) 535.

[58] Murton "Bank intermediation, bank runs, and deposit insurance" 1989 *FDIC Banking Review* 1.

[59] IMF Database (n 6) 15.

government ought to have provided a guarantee to depositors because it urged them, through public pleas, to cease hoarding their money and to deposit that money in a banking system that it knew to be unsafe.[60] In that moral obligation lay a *fundamental contradiction* in the goals of the DIS: to guarantee the deposits of all depositors, on the one hand, and to protect small depositors *only*, on the other.[61]

With regard to the faith-retaining policy goal of deposit insurance, the SARB states that "maintaining public confidence in the banking sector is at the heart of financial sector regulation".[62] A credible DIS maintains depositor confidence even in broken and dangerously fragile banks.[63] Banks play a key part in pricing and risk-taking in other economic sectors. For this reason, a well-functioning economy cannot dispense with the appropriate regulation of the banking sector.[64] In this sense, a DIS constitutes the crux of a country's financial safety-net system.

Do banks really need DI schemes?

The absence of a DIS or of sufficient funding when a bank goes bankrupt delays the resolution of the failed bank and increases costs.[65] Similarly, the absence of a DIS in Namibia explains why liquidating SME Bank consumes so much time and money.[66]

Of course, banks can always self-insure. But self-insurance does not protect depositors if the bank itself fails, nor can it diversify losses across banks.[67]

Though states should not avoid them, DI schemes have more than one substitute.[68] One alternative to a DIS is capital requirements. Together with deposit insurance, requiring banks to maintain good capital-to-asset ratios could avert bank failures.[69] But even alone, capital-to-asset ratios can ward off bank runs.

Another option is management. Regulation does not replace management.[70] On closer scrutiny, a DIS would not have stopped SME Bank from going bust: the bank failed because it mismanaged its assets,[71] not because Namibia lacks a DIS. Yet an explicit DIS would have contained the fallout from the SME Bank's bankruptcy. No matter how optimally policy-makers design a DIS, the deciding factor of the

60 Blad (n 5) 540.
61 See Blad (n 5) 541.
62 SARB (n 9) 2; see also Zongwe (n 12) 99–100.
63 Demirgüç-Kunt, Kane and Laeven "Determinants of deposit-insurance adoption and design" 2005 World Bank 21 (http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.694.8400&rep=rep1&type=pdf (accessed 9 August 2019)).
64 See Demirgüç-Kunt, Kane and Laeven (n 63) 23.
65 See Suphap (n 7) 836.
66 Zongwe and Katjaimo (n 2) 11.
67 Shaffer (n 7) 82.
68 See Demirgüç-Kunt, Kane and Laeven (n 63) 24.
69 Keeley (n 7) 1184.
70 SARB (n 9) 2.
71 See *Bank of Namibia v Small & Medium Enterprises Ltd (2)* 2018 (1) NR 183 at 186C–D (hereinafter referred to as "*SME Bank* case (provisional order)") (observing how, over and above the loss of 196 million Namibian dollars, SME Bank lost vast sums in its lending and its other activities).

safety and soundness of banks is management, coupled with the market forces that exercise discipline on that management.[72]

3.2 *Moral hazard*

Above all, most DI schemes are under-priced.[73] This implies that moral hazard is rife in the world's various DI schemes. Moral hazard, a term borrowed from 19th century insurance practice,[74] arises when behaviour alters as mispricing occurs in a buyer-seller contract.[75] This means that, as a result of the mispricing, individuals use too much of some resource or engage in too little care.[76]

The essence of the deposit-insurance problem is therefore one of accurately pricing risk. When risk is optimally priced, conventional wisdom holds that insurance frees itself from moral hazard and that it precludes that risk from materialising. Moral hazard looms the largest in situations where insurers effectively grant unlimited coverage and where they do not adjust premiums for risk.[77]

To fully grasp how moral hazard works, imagine a world before or without deposit insurance. In such a world, it is the market that disciplines banking institutions. If a bank takes too much risk and loses assets as a result, its clients will rush to withdraw their money from the bank, pushing the bank into bankruptcy.[78] The fear of a bank run keeps banks on their toes. However, in such a perfectly efficient world – Chang observes – banks will become less, not more, stable.[79]

Moral hazard emerges because deposit-insured banks tend to care less about the risks they take than deposit-uninsured banks. Hence the irony of moral hazard: *the heavy loss that the insurance scheme tries to avoid is the very harm that the scheme is more likely to bring on*. In the aftermath of the savings and loans crisis in the 1980s, many observers in the US started to question the desirability of deposit insurance. They believed that the deposit insurance's blanket protections and policies actually brought about the crisis.[80] Likewise, as contended later in this chapter, the dubious investments and the eventual fall of the SME Bank were induced by the moral hazard created by the Namibian government's generous financial backing of that bank.

[72] See SARB (n 9) 2.
[73] IMF Database (n 6) 19.
[74] Miller (n 42) 109.
[75] Ippolito *Economics for Lawyers* (2005) 350.
[76] Miller (n 42) 109 (stating that moral hazard is simply that, with deposit insurance in place, bankers will tend to undertake risky investments and activities that they would otherwise avoid, increasing the exposure of the deposit insurance fund to pay-outs in the event of failure); Williamson (n 41) 105 (concluding that deposit insurance encourages banks to engage in inefficient and risky behaviour, *ie* to assume risk even when disutility from doing so is greater than the risk premium earned, except for the under-priced deposit insurance subsidy available from the government); Ippolito (n 75) 350.
[77] Miller (n 42) 110.
[78] See also Suphap (n 7) 840.
[79] Chang *23 Things They Don't Tell You About Capitalism* (2011) 231–241.
[80] Keeley (n 7) 1183; Blad (n 5) 533.

Deposit insurance distorts incentives,[81] such that deposits would sooner move away from conservative banks to riskier banks.[82]

In the context of motor vehicle insurance, for example, an insured driver may not worry about the financial repercussions of reckless driving (he is insured, after all!) while the uninsured driver may act more carefully. This reasoning also applies in bank deposit insurance. However, moral hazard as it manifests in ordinary insurance markedly differs from the way it operates in deposit insurance. In ordinary insurance, the insured party is also the party who controls the risky activity.[83] In deposit insurance, on the other hand, the insured party (*ie* the depositor) does not control the risky business.[84] Based on this difference, one may assume that moral hazard in the banking setting poses less of a problem than in other settings, but this assumption would be wrong because the existence of deposit insurance may tempt banks to take inordinate risks.[85]

3.3 *The financial safety net*

A DIS evolves within the wider framework of a country's financial safety net. The fairness and efficiency of a country's safety-net design can be measured "by the extent to which design features promise to preserve the system's financial integrity without either subsidising or penalising bank risk-taking".[86] A sound financial safety system also preserves the core function of banking, such as linking savers and borrowers (*ie* financial intermediation), maintaining the country's payment and settlement systems, and supporting monetary policy.[87]

Though they may take on different forms, sound financial safety nets contain the following four components: (1) banks' access to a lender of last resort;[88] (2) final, riskless settlement of payment system transactions; (3) prudential supervision of banks; and (4) deposit insurance.[89] The way deposit insurance interacts with other components of the safety net can "mend or break" the financial system.[90]

The absence of a DIS in Namibia and South Africa means that there is a yawning hole in these countries' financial safety net. A sound DIS and financial safety net

[81] Williamson (n 41) 106; Kane "A six-point program for deposit-insurance reform" 1983 *Housing Finance Review* 269 at 271; Carns (n 21) 13.

[82] Carns (n 21) 13.

[83] Miller (n 42) 109.

[84] *Ibid*.

[85] Miller (n 42) 109–110.

[86] *Ibid*.

[87] Ketcha *Deposit Insurance System Design and Considerations* Bank for International Settlements Policy Paper (1999) 221.

[88] See Bank of Namibia Act 15 of 1997, s 32 (providing that the Namibian central bank may, where it considers it necessary to maintain a sound financial system, act as a lender of last resort). South Africa has similar provisions: see South African Reserve Bank Act 90 of 1989, s 10(1)*(f)*. See also De Jager "The South African Reserve Bank: A central bank in the firing line" 2018 *Annual Banking Law Update* 115 at 120–122.

[89] Ketcha (n 87) 222.

[90] Yokoi-Arai "The relationship between a single financial regulator and the deposit insurance system: Analyzing Japan" 2005 *The International Lawyer* 63.

must balance competing interests. Depositors have an interest to put their money in a safe place that earns interest (*ie* profit); bankers have an interest in a stable stream of deposits to loan money with interest (*ie* profit); and society has an interest in stable banks that make it easier for money and credit to flow within the economy.[91] A sound DIS should preserve the balance of interests within the banking system, but not at the cost of absolute stability.[92]

The global recession that started in 2008, the most severe systemic financial crisis since the Great Depression,[93] offers an excellent example of how a financial safety net can cushion the worst effects of major economic or banking crises. During and shortly after that financial tsunami, national governments spread their financial safety nets. To maintain the confidence of the public in the financial system, they applied one or more of the following six measures:[94]

1. introducing deposit insurance;
2. increasing the statutory limit of the national deposit insurance;
3. abolishing co-insurance;[95]
4. introducing government guarantees on bank deposits, bank assets or non-deposit liabilities;
5. extending liquidity support through the central bank; and
6. nationalising targeted banks.

National governments in many countries expanded financial safety nets.[96] This expansion raises serious questions on whether governments can wholly finance the promises of present DI schemes in future periods of stress and on how to juggle between the objective of preventing bank runs and the objective of avoiding moral hazard and threats to the financial system from incentives for aggressive risk-taking.[97]

Expectations that governments will extend the financial safety net, as they did during the 2008 crisis and beyond, reduce the effectiveness of deposit insurance.[98] Research found that generous financial safety nets increase bank risks and the vulnerability of the financial system.[99] This is a consequence of moral hazard.

[91] Economides, Hubbard and Palia (n 56) 15; Blad (n 5) 535.
[92] Blad (n 5) 556.
[93] Chu "Deposit insurance and banking stability" 2011 *Cato Journal* 99.
[94] See also IMF Database (n 6) 14–16; Chu (n 93) 99; Miller (n 42) 104.
[95] Co-insurance refers to the arrangement whereby depositors are insured for only a pre-determined portion of their funds and for a rate lower than 100% of the insured deposits, *eg* 70%. Before the crisis, few banks used co-insurance to combat moral hazard. See IMF Database (n 6) 10.
[96] IMF Database (n 6) 3.
[97] *Ibid*.
[98] IMF Database (n 6) 17.
[99] Anginer, Demirgüç-Kunt and Zhu "How does deposit insurance affect bank risk? Evidence from the recent crisis" 2014 *Journal of Banking and Finance* 312.

4 Comparing the designs of deposit-insurance schemes

4.1 *The parameters of this comparative exercise*

Poor designs

The crucial element with deposit insurance is not the insurance as such, but the way policy-makers design this insurance. Accordingly, this section looks at various deposit-insurance models.

A World Bank paper shows that wealthy countries tend to adopt an explicit DIS and that they also manage the design features of the scheme better.[100] By contrast, DI schemes installed in crisis circumstances or under pressure to emulate other countries are often poorly designed.[101] Pressure and efforts to emulate the regulatory frameworks and DI schemes of developed countries lead developing countries to choose design features and DI schemes that inadequately control risk-shifting.[102]

If anything, the 2008 global financial recession proves that many governments had poorly designed their DI schemes. For this reason, they could not avoid moral hazard. Indeed, these DI schemes could not stop risk from building up in the banking system because they failed to impose market discipline and to compensate for the risks transferred to them.[103]

The statistical basis of the comparisons

The design comparisons undertaken in this section are mainly informed by an IMF study. Carried out in July 2014, this study provides a database of DI schemes around the world (hereinafter "IMF Database").[104] It covers DI schemes in 188 countries plus Liechtenstein. And it finds tentatively that DI schemes fulfil their core mission of preventing open runs on the banks by insured depositors.[105] This is a significant finding since one of the foremost aims of a DIS consists in keeping away bank runs. Also significant, this finding belies ample empirical evidence that DI schemes could not maintain financial stability.[106] At the very least, evidence suggests that DI schemes do not suffice in monitoring risk-taking and imposing market discipline in the midst of a systemic banking crisis.[107]

Overview

A host of criteria could be used to compare and contrast DI schemes around the globe. These include institutional structure, administration of that structure,

[100] Demirgüç-Kunt, Kane and Laeven (n 63) 23–24.
[101] Demirgüç-Kunt, Kane and Laeven (n 63) 24.
[102] Demirgüç-Kunt, Kane and Laeven (n 63).
[103] IMF Database (n 6) 15.
[104] IMF Database (n 6).
[105] IMF Database (n 6) 3.
[106] Chu (n 93) 109 (showing that deposit insurance is no panacea for banking instability as some countries with DI schemes still experienced banking crises regardless of the extent of the deposit-insurance coverage).
[107] IMF Database (n 6) 18.

compulsory or voluntary membership of the deposit insurance fund, funding, scope of the insurance, coverage, premiums, pay-outs, and statutory limits.[108] This section is arranged according to the following comparison criteria: organisational structure, funding, scope, coverage, premiums, pay-outs, statutory limits and systemically important banks.

4.2 *Organisational structure*

Organisation types

A variety of institutional and administrative structures exist with respect to DI schemes. In the midst of this institutional variety, four general models emerge. A DI scheme can be organised (1) as a separate, independent legal entity; (2) as an entity falling under a country's banking supervisory structure; (3) as an entity falling under the authority of the national central bank; or (4) as an entity under a national government ministry, especially the finance ministry.[109] Because experience suggests that, in crisis circumstances, political pressures lead to decisions that do not further the long-term best interest of a sound and efficient banking system, an independent authority stands in the best position to withstand such pressures.[110] And, as a matter of fact, the majority of DI schemes are organised as independent entities, though they may be housed within the banks' regulator, the central bank, a government ministry or the ministry of finance.[111]

Stand-alone funds

An independent or stand-alone deposit-insurance fund increases the likelihood that money will be available when needed. However, a fund that would form part of government institutions may create obstacles when the fund needs money. Even as a stand-alone structure, the fund will have money at its disposal only if the premiums charged are high enough and assumptions made about possible losses are realistic.[112] It is the financial capacity of the insurer that lends credibility to a deposit insurance guarantee.[113]

Administration of deposit insurance funds

A public institution, a private entity, or a public-private partnership can administer funds, however organised. The better half of DI schemes is administered by a

[108] Talley and Mas "Deposit insurance in developing countries" World Bank Working Paper WPS 548 (1990); IMF Database (n 6).
[109] See IMF Database (n 6) 5.
[110] Ketcha (n 87) 226.
[111] IMF Database (n 6) 12.
[112] Ketcha (n 87) 31.
[113] Ketcha (n 87) 225.

public institution,[114] such as the FDIC in the US, the Canada Deposit Insurance Corporation (CDIC), or the Nigeria Deposit Insurance Corporation (NDIC).

4.3 *Funding the deposit-insurance scheme*

A good DIS must commit credibly to pay out depositors if their banks fail.[115] The DIS can establish its credibility by collecting funds before a systemic crisis arises (*ie ex ante*) or after the crisis (*ie ex post*). Although some have blamed it for draining the liquidity of the banking sector,[116] *ex ante* funding still represents about 80% of DI schemes.[117]

Ex ante *versus ex post funding*

Ex ante funding often takes the form of premiums paid regularly by participating banks. On the other hand, *ex post* funding generally consists in a deposit-insurance entity collecting funds from surviving participating banks only after a bank fails and the money available to cover the depositors of the failed banks falls short.[118]

Ex post funding costs less than *ex ante* funding, and – unlike its *ex ante* counterpart – *ex post* funding may encourage banks to monitor other banks (*ie* peer monitoring).[119] Despite these benefits, *ex post* funding arrangements practically do not exist in low-income and lower-middle-income nations.[120] But these arrangements have their own weaknesses. For one, with *ex post* funding, the resolution of failed banks proceeds more slowly.[121] For another, failed banks will not have to contribute to the fund, which raises questions of fairness since the contributing banks will have to pay in circumstances where they have not taken excessive risks while the failed banks escape the losses they have caused.[122]

That said, *ex ante* and *ex post* schemes are not mutually exclusive. For instance, in several countries, regulators build a deposit-insurance fund *ex ante*, but have the power to impose charges *ex post* when a bank collapses.[123]

Funding sources

The funding of a DIS can come from various sources. It can come from government, the private sector (when participating banks regularly contribute to the funding of the DIS), or from both the government and the private sector. Funding flows primarily from the private sector (*ie* participating commercial banks).[124]

[114] IMF Database (n 6) 12–13.
[115] IMF Database (n 6) 7.
[116] Suphap (n 7) 836.
[117] IMF Database (n 6) 12.
[118] IMF Database (n 6) 7.
[119] Suphap (n 7) 836.
[120] IMF Database (n 6) 12.
[121] Suphap (n 7) 836.
[122] See Suphap (n 7) 836.
[123] Suphap (n 7) 837.
[124] IMF Database (n 6) 13.

Each funding source presents policy-makers with downsides and advantages. A government-funded DIS promises greater financial assistance than a private fund. But, with a government-funded DIS, the danger is that, in the event of a massive bank failure, the government may baulk at using taxpayers' money to salvage the failed banks.[125]

A private fund may encourage peer monitoring among banks,[126] and, unlike the publicly funded DIS, it pays out money to depositors of failed banks with little regard to political expediency. In cases of insolvency, the deposit insurer does not take a cent from the national purse. However, if a systemic banking crisis develops, the private fund may run out of money to pay back insured depositors.[127] In the late 1980s, the savings and loans crisis in the US depleted the federal deposit-insurance fund.[128] In such conditions, if the government does not intervene, depositors will bear the losses.

4.4 *Scope of the insurance*

Another element of the framework for comparing DI schemes is the scope of activities that fall under the insurance coverage. For instance, a DIS should not cover the non-banking activities carried out by banks. Some jurisdictions have limited their DI schemes to activities that qualify as traditional banking and closely related functions by requiring banking and non-banking organisations to carry out banking and non-banking activities respectively.[129] They restrict activities because extending the insurance coverage to non-banking operations would expose the deposit insurer to greater risks that the insured institutions would fail.[130]

4.5 *Coverage*

DI schemes comprise the coverage of the insurance. Different deposit insurers cover different types of deposits. Some DI schemes cover all types of deposits, including foreign-currency deposits; several exclude inter-bank deposits; and some cover only household accounts.[131] Still, other DI schemes cover the deposits held by branches or subsidiaries of foreign banks located within their jurisdiction.[132] These different DI schemes reflect the difference between the emphasis placed on stabilising the financial system and that placed on protecting the small, less sophisticated savers.[133]

125 IMF Database (n 6) 7–8.
126 IMF Database (n 6) 8.
127 Miller (n 42) 111.
128 Blad (n 5) 533.
129 Ketcha (n 87) 227.
130 *Ibid*.
131 Ketcha (n 87) 229; IMF Database (n 6) 7.
132 See IMF Database (n 6) 6.
133 Ketcha (n 87) 229.

4.6 *Premiums*

Most DI schemes charge premiums on the banks they insure. An essential feature of a good deposit-insurance design relates to premium pricing. The pricing must not induce inordinate risk-taking. If a deposit insurer under-prices its premiums, banks will have an incentive to take greater risks.

The issue of premium pricing remains vital as evidence from the 2008 crisis reveals that, though effective in warding off large-scale bank runs, DI schemes were bad at correctly pricing risk.[134]

Flat-rate system

DI schemes could implement various systems with regard to premiums. One such alternative is the fixed-rate or flat-rate system. By the turn of the 21st century, an estimated 50 countries chose the flat-rate pricing scheme,[135] plausibly the most common of deposit-insurance premiums.[136] The US, the first country to introduce an explicit DIS, applied the flat-rate system for nearly six decades.[137] Low failure rates characterise most of this history.[138] Under this system, subscription must be compulsory in order to avoid a situation whereby the DIS attracts riskier entities,[139] as opposed to safer ones. However, the flat-rate system is more likely to exacerbate than to lessen the problem of moral hazard. Its rigidity means that more often than not it misprices bank risks.

Risk-adjusted system

To avoid the situation of "adverse selection", policy-makers can opt for a risk-based premium pricing scheme. Under a risk-adjusted system, banking regulators classify depository institutions into risk categories and impose higher premiums on institutions falling under higher risk-categories.[140] Thus, the risk-adjusted system resembles conventional types of insurance.[141]

Risk-adjusted premiums offer several benefits, namely ensuring a level playing field among banks of different risk profiles; rectifying the inequitable subsidising of riskier banks by safer banks; bridging the information asymmetries between banks and their depositors; reducing regulatory and enforcement costs; and, more importantly, curbing moral hazard.[142] Indeed, some experts praise the risk-adjusted pricing system for discouraging the moral hazard of excessive risk-taking.[143]

134 See IMF Database (n 6) 19.
135 Suphap (n 7) 837.
136 Suphap (n 7) 831.
137 Keeley (n 7) 1183; Suphap (n 7) 837; Ketcha (n 87) 232.
138 Keeley (n 7) 1183.
139 Ketcha (n 87) 232.
140 Suphap (n 7) 830–831.
141 Suphap (n 7) 837.
142 Suphap (n 7) 830.
143 Suphap (n 7) 831.

Nonetheless, risk-adjusted pricing confronts big hurdles. First, designing and implementing a risk-adjusted premium system is "extremely complex",[144] which is probably the reason why the flat-rate model still obtains in different parts of the globe. By contrast, governments can design and apply the flat-rate model more easily. Secondly, insofar as the premiums do not fully reflect the risks the DIS faces, the risk-based system also calls for mandatory membership[145] – a system adopted by at least 11 states, including the US, Canada and Argentina.[146]

Nevertheless, like *ex ante* and *ex post* funding, the flat-rate and risk-adjusted systems are not mutually exclusive, and some countries combine elements of both systems.

4.7 *Pay-outs*

Under the standard model, the DIS reimburses each depositor per bank. Certain models insure each depositor account or each depositor.[147] These models sharply differ from the standard (*ie* the each-depositor-per-bank) model. By insuring each depositor account, the insurer pays out much more, while, by insuring each depositor, it pays out less than it does with the standard model.

Almost invariably, insurers cover deposits up to a certain statutory limit. Unsurprisingly, these pay-out limits vary widely across countries. Typically, they are denominated in local currency. International best practice recommends that the deposit insurer pay out within seven working days.[148]

4.8 *Statutory limit*

Before the 2008 financial storm, deposit insurers in seven countries (Finland, Iceland, Japan, Kuwait, Mexico, Norway and Turkey) covered deposits in full.[149] Nonetheless, the IMF generally recommends that countries multiply by one or two the GDP per capita to calculate the appropriate amount of deposit-insurance coverage.[150]

Furthermore, some bank observers have suggested decreasing the scope of statutory-limit coverage to reduce moral hazard and restore market discipline.[151] In particular, Chu demonstrated that low coverage beats both high and full coverage in maintaining banking stability.[152]

Statutory limits encourage depositors to diversify their risks across banks as DI schemes insure their deposits on an each-depositor-per-bank basis. DI schemes also tell a cautionary tale, warning depositors how much money they can invest in any single banking institution.

[144] Suphap (n 7) 832; see also Blad (n 5) 557.
[145] Ketcha (n 87) 232.
[146] Ketcha (n 87) 233.
[147] IMF Database (n 6) 8–9.
[148] SARB (n 9) 6.
[149] Chu (n 93) 108.
[150] Ketcha (n 87) 229.
[151] Chu (n 93) 113; Carns (n 21) 11; Blad (n 5) 533.
[152] Chu (n 93) 113.

4.9 *Too-big-to-fail banks*

The failure of some large banks may exceed the ability of a DIS to maintain financial and banking stability. This is often referred to as the "too-big-to-fail" problem and these banks are called "systemically important financial institutions" (SIFI). When a SIFI fails, it may threaten the nation's financial stability and prompt governments to intervene.[153] Such intervention undermines market discipline and puts small banks and their customers at a disadvantage.[154]

Policy-makers can resolve this problem in at least two ways. First, they can create a class of creditors with clear motivations to monitor a bank's risk-taking.[155] Secondly, they can make large banks internalise the costs of extending special protections to them.[156] In practical terms, this solution implies that large banks must pay extra costs to cover those incurred in bailing them out.[157] This chapter endorses this solution.

5 The position in Namibia and South Africa

5.1 *Namibia*

Banking and deposit insurance

The banking sector in Namibia has been dominated by four[158] commercial banks, three of which are headquartered in South Africa: FNB Namibia, Standard Bank and Nedbank.[159] Well capitalised, these banks have embraced conservative banking policies.[160]

The banking sector in Namibia and South Africa may have grown stronger in fact because of the absence of DI schemes or generous financial safety nets. Interestingly, even though policy-makers usually advance moral hazard as one of the main reasons for implementing an explicit DIS, moral hazard was, in the South African experience, the reason for *not adopting* an explicit DIS.[161]

Namibia lacks an explicit DIS; it only has an implicit scheme. A bill on deposit insurance is being developed, but government has not yet tabled it in Parliament.

SME Bank

The collapse of the SME Bank in 2017 illustrates the current state of deposit insurance in Namibia. The bank sprang from a joint venture between the Namibian government (through the Ministry of Industrialisation, Trade and SME

153 Ketcha (n 87) 230.
154 *Ibid*.
155 *Ibid*.
156 Ketcha (n 87) 231.
157 *Ibid*.
158 First National Bank (FNB), Standard Bank, Nedbank and Bank Windhoek.
159 Sherbourne *Guide to the Namibian Economy 2017* (2017) 391.
160 Sherbourne (n 159) 391, 401.
161 SARB (n 9) 9.

Development) and two Zimbabwean banks (the Metropolitan Bank of Zimbabwe and World Eagle Properties). With a 65% stake, the Namibian government held the majority of shares. Though the central bank, Bank of Namibia (BON), initially hesitated to issue a licence to SME Bank because controversial figures would participate in running it, the central bank eventually bowed to political pressure and allowed SME Bank to operate.[162]

The debacle

The SME Bank invested approximately 196 million Namibian dollars into Mamepe Capital (Mamepe), an investment company in South Africa. Mamepe then contracted with VBS Mutual Bank for the bank to act as Mamepe's banker in the investment deal binding Mamepe to SME Bank.[163] In addition, SME Bank sustained heavy losses from its lending and other activities.[164] In January 2017 BON informed SME Bank that it no longer met its local assets requirements.[165]

On 1 March 2017 BON assumed control of SME Bank in terms of the Banking Institutions Act.[166] On 31 May 2017, in terms of the Banking Institutions Act,[167] BON requested by letter the shareholders of SME Bank to inject by 13 June 2017 an amount of N$359 million to recapitalise their ailing bank.[168]

On 21 June 2017 the Ministry of Industrialisation, Trade and SME Development (Trade Ministry) replied to BON that "no resources are available to ensure timely recapitalization of SME Bank" as requested.[169] Seeing that the two Zimbabwean shareholders could also not inject funds into SME Bank, BON moved on to apply for the compulsory winding up of the bank.

When the High Court of Namibia granted the provisional order on 10 July 2017, the total liquidity available to SME Bank stood at N$3 895 994.25[170] while its liabilities exceeded N$468 000 000.25.[171] The cash left to SME Bank was so little that it would "dissipate in an instant".[172] Therefore, the court held (and rightly so) that the bank had become factually insolvent (as its liabilities exceeded its assets) and commercially insolvent (as it could no longer pay its debts as they fell due).[173]

162 Sherbourne (n 159) 398–399. He mentions the names of the Zimbabwean businessman Enock Kamushinda and Andrew Ndishishi, the then permanent secretary at the Namibian health ministry.
163 Van Rensburg "How VBS Scheme Broke Namibian Bank" *City Press* 3 December 2018 (https://city-press.news24.com/Business/how-vbs-scheme-broke-namibian-bank-20181203 (accessed 9 August 2019)).
164 *SME Bank* case (provisional order) (n 71) 186C–D.
165 *SME Bank* case (provisional order) (n 71) 185H–186C.
166 Banking Institutions Act 2 of 1998, s 56.
167 Banking Institutions Act 2 of 1998, s 28(4).
168 *SME Bank* case (provisional order) (n 71) 185H–186C.
169 *SME Bank* case (provisional order) (n 71) 186G–H.
170 *SME Bank* case (final order) (n 1) 196C.
171 Referring specifically to the investments by the National Energy Fund (*ie* N$368 442 770.04) and the Government Institutions Pension Fund (*ie* N$100 million).
172 *SME Bank* case (final order) (n 1) 205H.
173 *SME Bank* case (final order) (n 1) 195E–I.

Government-induced moral hazard caused SME Bank to fail

One aspect of the SME Bank's downfall has eluded observers: the bank's failure was caused by moral hazard. The moral hazard was, in turn, induced by the colossal amounts of money injected by the government into the bank. As Prinsloo J noted, the government and parastatals have invested hundreds of millions of Namibian dollars into the bank. The government alone poured approximately N$900 million of taxpayers' money into the bank.[174] This financial mattress encouraged the bank to make dubious investments with Mamepe, confident that it could count on the government if the investments failed. Tellingly, even during the bank's liquidation proceedings, the two Zimbabwean shareholders still expected the government to recapitalise the bank.[175] Clearly, the government's financial largesse brought about moral hazard, which precipitated the spectacular failure of SME Bank.

What the SME Bank debacle says about the banking sector in Namibia

The fact that commercial banks in the country conduct their business prudently has probably kept Namibia from any major banking crisis. The SME Bank's demise only constituted an isolated, non-systemic bank run, which may explain why BON did not extend lender-of-last-resort (LOLR) assistance to SME Bank.[176] The bankruptcy was caused, not by the bank run, but by the dubious investment with Mamepe. The bank run was touched off by BON taking over SME Bank following that failed investment.

The SME Bank closure shows that banking crises in Namibia may come from an economic downturn rather than an internal crisis in the banking industry. It also shows that developing-country governments frequently shrink from intervening during these downturns that hugely strain their finances.

5.2 *South Africa*

Why Namibia could learn from South Africa

South Africa pertains to a DIS in Namibia for more than one good reason. First, the lucrative banking sector in Namibia is dominated by four banks, three of which come from South Africa (*ie* FNB, Standard Bank and Nedbank),[177] implying that the banking sectors of the two countries intersect closely. Secondly, the South African Rand circulates in Namibia as legal tender such that the monetary policy set up by South Africa's central bank anchors Namibia's monetary policy. This matters because DI schemes cannot work effectively in the context of a worthless currency or a currency hit by hyperinflation.

[174] *SME Bank* case (final order) (n 1) 205E–F.

[175] *SME Bank* case (final order) (n 1) 207H–208C.

[176] However, because central banks typically provide LOLR emergency assistance under strict confidentiality (as they dread bank runs) the parties would not disclose in court papers that BON had ever previously supported SME Bank through LOLR assistance. See also De Jager (n 88) 118 and 120.

[177] Sherbourne (n 159) 391.

South African law and policies also speak to a study of DIS in Namibia because Namibia has largely inherited the South African legal system,[178] including banking laws, and the Namibian economy heavily leans on South Africa's economy. Thus, if South Africa suffers a systemic banking crisis, it will likely contaminate Namibia's banking sector.

The policy environment

The South African government is writing a Designated Institutions Resolution (DIR) Bill, which introduces a DIS into South Africa. In terms of the DIR Bill, the South African Reserve Bank (SARB) serves as the deposit insurer.[179]

The South African government expects most of the design features of the DIS to emanate from regulation or other forms of secondary legislation rather than in the DIR itself.[180] Similarly, the Financial Services Regulation Act (FSRA) does not constitute a law aimed at a DIS, but nonetheless contains some provisions relevant to the DIS.

Three policy documents inform the DIR Bill, namely *A Safer Financial Sector to Serve South Africa Better* (2011) (by the National Treasury),[181] *Strengthening South Africa's Resolution Framework for Financial Institutions* (2015) (by the National Treasury),[182] and *Designing a Deposit Insurance Scheme for South Africa* (2017) (by the SARB).[183] Together, these documents, the regulatory framework and the DIS form the comprehensive architecture for reducing the social and economic costs of failing banks.[184]

Rundown of the salient features

A team consisting of local academics and experts from the SARB, National Treasury and the World Bank worked on the latest version of the proposed South African DIS.[185] Overall, while well motivated, the proposed DIS suffers from a glaring omission: currently, the proposals do not address the indispensable element of premiums or risk-pricing. The author of this chapter submits that this shortcoming dramatically weakens the proposed South African DIS. The salient features of South Africa's DIS are:

[178] Mandate for South West Africa, 1919, art 2 (mandating South Africa on behalf of the League of Nations (and later the United Nations) to apply its legal system in Namibia). See Zongwe *International Law in Namibia* (2019) 135, 416. Proclamation 20 of 1919 (applying the law of South Africa); and Administration of Justice Proclamation 21 of 1919, s 1(1) (introducing Roman-Dutch law "as existing and applied in the Province of the Cape of Good Hope" into South West Africa/Namibia). See Amoo *An Introduction to Namibian Law: Materials and Cases* (2008) 55–77.

[179] See National Treasury (n 8) v, vi, 10, 11 and 14–17.

[180] National Treasury (n 8) 32.

[181] National Treasury *Safer Financial Sector* (n 9).

[182] National Treasury (n 8).

[183] SARB (n 9).

[184] SARB (n 9) 1.

[185] SARB (n 9) 1.

1. The DIS is explicit and credible, in line with best practices outlined in international standards.[186] Furthermore, the design of the DIS does not intend to place an excessive cost on the banking system, distort competitiveness in the banking sector, or cause moral hazard to the extent that it threatens financial stability.[187]
2. The DIS aims to protect covered deposits in the event of a bank failure, thereby helping the state to protect customers and stabilise the financial system.[188]
3. The DIS has a paybox-plus mandate, provided that it does not cost the DIS less than that which it would have to pay out in the event of a bank liquidation.[189]
4. The DIS structures the deposit insurer as a subsidiary of the central bank – a separate legal entity with its own legislative and governance framework, but physically located in the SARB.[190]
5. Membership of the DIS is compulsory and automatic for all registered banks.[191]
6. Qualifying deposits comprise all deposits held by banks, except deposits by banks, deposits by the non-private financial sector, deposits by government,[192] and bearer deposit instruments (*eg* promissory notes).[193]
7. The SARB recommends 5% as the target size of covered deposits, to be maintained on a continuous basis.[194]
8. The DIS covers all qualifying deposits up to R100 000 per depositor per bank.[195] This chapter considers that amount as unnecessarily high given that most deposits in South Africa do not exceed R3 000.[196] This amount implies that the target size of the deposit-insurance fund (which is based on covered deposits) may underestimate the extent of its legal commitments (which are based on the statutory limit) after a bank fails. In the specific context of Namibia, deposits are concentrated in the hands of four commercial banks. If one of them fails, it will likely wipe out the entire fund because it would probably hold more than 5% of covered deposits.

[186] SARB (n 9) 3, 8*ff*.
[187] SARB (n 9) 3.
[188] SARB (n 9) 3, 8–16.
[189] SARB (n 9) 3, 44–47.
[190] SARB (n 9) 4, 23–25.
[191] SARB (n 9) 4, 26–27.
[192] Thus, the proposed South African DIS will not cover deposits by municipalities, such as those made illegally in the infamous VBS Mutual Bank case.
[193] SARB (n 9) 4, 28–30.
[194] SARB (n 9) 5, 34.
[195] SARB (n 9) 4, 31.
[196] See International Monetary Fund (n 10) 72.

9. With respect to deposit coverage, the DIS covers deposits by foreign nationals, and deposits in foreign currency.[197] However, it does not cover deposits at foreign branches and subsidiaries of South African banks abroad.[198] This means that the South African DIS will not cover the deposits held by most major banks (*eg* FNB, Standard Bank and Nedbank) in Namibia. All the same, the IMF recommends that Namibia cooperate with the South African central bank in planning the recovery and resolution of these large banks.[199]
10. South Africa selected a partially pre-funded approach for the DIS, supplemented by emergency liquidity in the event of shortfalls.[200]
11. When a bank goes bankrupt, the DIS must pay out within 20 working days after the closure of the banks for accounts where ownership can be easily ascertained.[201]

6 Evaluation of deposit insurance designs

6.1 *The governing principle: Market discipline*

Basically, deposit-insurance coverage strives to find the right balance between avoiding devastating bank runs and exposing banks to market discipline. In much the same way, the history of the US DIS is one defined by a constant tug-of-war between safety (*ie* financial stability) and efficiency.[202] And, under the rubric of safety, policy-makers must negotiate a trade-off between two potential sources of instability: bank runs and excessive risk-taking.[203] In other words, the coverage must be extensive enough to prevent crippling bank runs, but not so extensive as to eliminate market discipline on the bank's risk-taking.[204]

Market discipline refers to "market-determined incentives to control risk-taking".[205] The sources of market discipline comprise depositors, shareholders and other non-deposit creditors.[206] The challenge is that deposit insurance weakens market discipline. Specifically, it weakens the threat of bank runs and the demands of depositors for higher interest from riskier banks.[207]

Ideally, a DIS should aim to eliminate that portion of market-determined risk premium reflecting the threat of bank runs without changing the portion reflecting

[197] SARB (n 9) 4, 28–30.
[198] *Ibid*.
[199] International Monetary Fund (n 2) 17.
[200] SARB (n 9) 5, 32–41.
[201] SARB (n 9) 6, 46–47.
[202] Blad (n 5) 556.
[203] Carns (n 21) 12.
[204] Carns (n 21) 12 (writing that the policy question is whether the trade-off between bank runs and excessive risk-taking, represented by the statutory limit, is optimal); Ketcha (n 87) 229.
[205] Carns (n 21) 11.
[206] *Ibid*.
[207] Carns (n 21) 12.

other risks.[208] In short, a DIS must infuse market discipline while avoiding the social costs of bank runs.[209]

6.2 *Risk-pricing and premiums*

The risk-adjusted premium pricing enjoys a competitive edge over flat-rate pricing in at least four respects: blunting the moral hazard for excessive risk-taking, offering a fair methodology for pricing premiums, correcting banking-related information asymmetries, and decreasing regulatory and enforcement costs.[210] This section addresses these aspects one by one.

Taming moral hazard

It follows from the foregoing discussion that risk-adjusted premiums best tackle the moral hazard. Because of competition in the industry, banks would have an increasing incentive to take on risks if they did not have to pay a premium for it.[211] Imposing premiums restrain risk-taking.[212]

By contrast, a flat-rate system induces banks to engage in riskier behaviour because they do not have to incur extra expenses for these extra risks.[213] Deposit insurers and depositors shoulder the costs of extra risks.[214] In this manner, bank managers can reap the benefits of a successful investment but pay nothing if the investment fails.[215] This is the perfect formula for risky behaviour in the banking industry.

To stop banks from taking inordinate risks or engaging in other bad practices, a DIS must feature supervision. Deposit insurance and bank supervision go hand in hand.[216] Supervision intends to identify excessively risky bank behaviour. In the US, the FDIC inspects insured banks periodically and requires them to provide it with a “report of condition”, which basically functions as a bank’s balance sheet.[217] Most importantly, supervision would allow a deposit insurer to tailor-make and charge risk-adjusted premiums for each bank.

However, Horvitz criticised the risk-adjusted premiums on two grounds. First, he argues that the claims that the risk-adjusted system is efficient and fair neglects the fact that risk to a DIS depends more on the amount of loss suffered when a bank fails than on the calculated probability (*ie* risk) of failure.[218]

[208] Carns (n 21) 11.
[209] *Ibid*.
[210] Suphap (n 7) 839.
[211] Keeley (n 7) 1185 (stating that increased competition in the 1950s and 1960s, as a result of liberalisation in the banking sector, reduced bank charter values, capital-to-asset ratios and banks’ incentives to act prudently); Suphap (n 7) 841.
[212] Carns (n 21) 11.
[213] See Suphap (n 7) 841; Keeley (n 7) 1183.
[214] Suphap (n 7) 841.
[215] Blad (n 5) 535.
[216] Carns (n 21) 13 (stating that, when deposit-insurance coverage is extensive, supervision becomes essential to preventing an increasing overexposure to risk in the banking industry); Blad (n 5) 544.
[217] Blad (n 5) 543–544.
[218] Horvitz (n 7) 253.

Secondly, relating premiums to risk calls for a means of measuring risk, which proves very difficult while the existing methods of measuring risk do not yet work perfectly.[219] Williamson has suggested that a technology-driven securities market presents a viable and efficient option to price bank risks.[220] But this alternative would not completely solve the issue of accurate risk-pricing either, as the securities market is not entirely objective and may not necessarily prevent systemic risks.

Eliminating free riding

Risk-adjusted premiums are also efficient because they reduce the free-riding phenomenon. For this reason, most economic agents would find it fair as well. Higher-risk banks should pay higher-premiums. If a higher-risk bank pays a lower-risk premium, the lower-risk bank will effectively subsidise the higher-risk bank's excessive risk-taking. That would amount to free riding, which would be distributionally unfair, and would inescapably lead to wasteful behaviour.

Correcting information asymmetries

Risk-adjusted premiums resolve information asymmetries. Generally speaking, banks tend to know more about their risk profiles than depositors.[221] As a result, the signal that banks send to depositors may mislead them. While deposit insurance aims to protect small, less financially savvy depositors, flat-rate premiums would not provide depositors with the precise signals about the relative riskiness of the banks they want to do business with.

A risk-adjusted premium, on the other hand, precisely signals which bank is more or less risky. In addition, DIS-reform advocates push for deposit insurers to release more information about each bank's financial condition.[222]

However, risk-adjusted pricing relies on market discipline. The asymmetric information associated with bank assets and the combination of these assets with callable liabilities make market discipline – especially depositor discipline – "costly, complex, and subject to error".[223] This problem, in turn, complicates complete contract-writing and, most importantly, the accurate pricing of risk.[224] In any event, a DIS must avoid any form of depositor discipline that increases the probability of bank runs.[225]

Saving regulation and enforcement costs

Insofar as premiums are priced accurately, a DIS can substantially reduce the costs of regulating and enforcing DIS-related standards. Some experts have gone as far

[219] *Ibid*.
[220] Williamson (n 41) 130.
[221] Suphap (n 7) 843.
[222] Blad (n 5) 535.
[223] Carns (n 21) 11.
[224] *Ibid*.
[225] Carns (n 21) 13.

as saying that perfectly risk-adjusted schemes could potentially replace a large part of existing banking regulations.[226]

6.3 *Risk-pricing methodology*

Six major challenges stand in the way of optimal risk-pricing: risk-pricing difficulties, operational costs, confidentiality concerns, cross-border complexities, political opposition and potential conflicts with existing risk-based capital standards.[227]

Risk-pricing difficulties

The fluctuations that come with a risk-adjusted premium system cost the banking industry more than a stable premium system.[228] Banks pass on these costs to their customers, which result in banks tightening credit.[229]

To surmount these difficulties, a DIS should determine a premium level based on the long-term average loss rate. In other words, the "average premium income [must] equal the long-run average expenses of the deposit insurance fund".[230] This long-term average loss rate would avoid both the inefficiency of a flat-rate system and losses occasioned by the constant volatility of a pure risk-adjusted premium system. At the same time, it will not drain banks' liquidity, and it will maintain the solvency of the deposit-insurance fund over time.

A premium above that loss rate would make banks pay too much in an actuarial sense[231] and raise the social costs of linking savers with borrowers.[232] Conversely, a premium below that rate would deplete the fund over time.[233] By mispricing risks, this rate would waste scarce resources.[234]

Setting the statutory limit

Difficulties arise in weighing the costs and benefits of changes in statutory-limit coverage.[235] Whether a DIS increases or decreases its coverage depends on whether the risk of bank runs or excessive risk-taking increases or decreases.[236] If the coverage is decreased, perceptions of costs associated with isolated bank runs differ.[237] The costs of these bank runs and the risk of contagion (*ie* the risk of

[226] Macey et al (n 7) 258; Suphap (n 7) 843.
[227] Suphap (n 7) 844–850.
[228] Shaffer (n 7) 81–82.
[229] Shaffer (n 7) 82.
[230] *Ibid*.
[231] *Ibid*.
[232] Carns (n 21) 11.
[233] Shaffer (n 7) 82.
[234] Kane (n 81) 270–271.
[235] Carns (n 21) 12.
[236] Carns (n 21) 14.
[237] Carns (n 21) 12.

systemic failure) are hard to measure objectively.[238] The rich history of banking crises or failures[239] does not give much guidance on how to determine the viable-in-the-long-run level of statutory coverage.[240] However, Chu argues that to raise coverage to eradicate bank runs is not necessarily the optimal policy response because bank runs, though threatening financial stability in the short run, reinforce banking stability in the long run.[241]

7 Policy options for Namibia

Even a well-designed DIS does not guarantee that it will achieve all its goals. All the same, an optimally designed DIS can go a long way in dealing with bank runs and systemic crises. The stark reality, however, is that the majority of DI schemes under-price risks, thereby subsidising potentially ruinous risk-taking by banks.[242]

Talley and Mas recommend that DI schemes in developing countries be public, compulsory, partial, pre-funded and flexible with regard to procedures for resolving bank failures.[243] Miller recommends that economies in transition must pick explicit and limited DI schemes, and DI schemes that involve the private sector to the greatest extent feasible.[244]

In light of its long discussion on DIS, this chapter prescribes the following plan for Namibia. The DIS that Namibia should implement must obviously be explicit. It must balance the necessity to ensure banking safety with concerns for greater efficiency.

Moreover, it must embrace the three common goals of DI schemes: to prevent bank runs, to protect small investors, and to instil public confidence in the banking system. The Namibian scheme must retain the core paybox function of deposit insurance while leaving the winding up of failed banks to the central bank. In addition to the paybox mandate, the Namibian DIS must require that banks maintain adequate capital-to-asset ratios.

The Namibian DIS should organise the deposit insurer as an independent, stand-alone legal entity or fund. Namibian policy-makers may have to think twice before imitating the proposed South African DIS whereby the deposit insurer operates as a subsidiary of the central bank. Unlike the Namibian central bank, the South African Reserve Bank is still privately owned.[245] The central bank's private ownership means that the deposit insurer does not need formal independence from the government.

238 *Ibid*.
239 For a chronicle of the world's major banking crises, see Reinhart and Rogoff *This Time is Different: Eight Centuries of Financial Folly* (2009).
240 See Carns (n 21) 12.
241 Chu (n 93) 101.
242 See IMF Database (n 6) 19.
243 Talley and Mas (n 108) 70.
244 Miller (n 42) 120–126.
245 South African Reserve Bank Act 90 of 1989, s 11, 21, 22, 23 and 24 (basically opening the SARB's shareholding to the general public subject to a number of restrictions). See De Jager (n 88) 131–132 (explaining that the South African Parliament established the SARB as a central bank with private shareholders and with no mission to maximise profit, and that it did not introduce the private ownership in the SARB primarily with a corporate governance objective).

Although its funding should come from commercial banks, whether public or private, it must be administered by a public body. Plus, given the fairly small size of its banking industry, Namibia should make membership of the DIS compulsory. Banks must fund it *ex ante*, with provisions for *ex-post* funding when a crisis strikes.

The scope of the deposit-insurance coverage must not extend beyond traditional banking activities. It must cover mainly individual and household accounts. The fund will make pay-outs on an each-depositor-per-bank basis. If the DIS mainly aims to protect small, less financially sophisticated savers, the statutory limit should hover around N$65 000, roughly the annual per-capita income in Namibia.[246]

Last but most importantly, the premium pricing model should adjust for risk, calculated for each bank as a long-term average loss rate. At the same time, the DIS must identify too-big-to-fail banks and oblige them to pay higher premiums.

[246] The average annual personal income in Namibia is US$4 640, as determined by the gross national income (GNI) per capita: World Bank, Namibia: Data, available at https://data.worldbank.org/country/namibia (accessed 26 April 2018).

A review of the South African Reserve Bank's financial stability policies*

HYLTON HOLLANDER**
DAWIE VAN LILL***

1 Introduction

In the immediate aftermath of the 2007–09 global financial crisis the (then) Minister of Finance Pravin Gordhan announced, during the 27 October 2010 Medium Term Budget Policy Statement, an expanded mandate of the South African Reserve Bank.[1] This mandate sanctioned the SARB to maintain and enhance financial stability. Now, after nearly seven years, the imposition of the Financial Sector Regulation Act 9 of 2017 (FSR Act) on 21 August 2017 introduced sweeping financial sector reforms and aligned South African regulatory and supervision practices with global standards. In addition to providing an explicit mandate for the SARB, this Act establishes two juristic authorities under the so-called "Twin Peaks" model of financial regulation. The first peak is the Prudential Authority.[2] The PA is a juristic person within the administration of the SARB. It facilitates the sound management of all deposit-taking institutions (*eg* traditional and cooperative banks), non-bank financial institutions (*eg* insurers and microloan organisations), financial conglomerates, and key market infrastructures such as the national payments system for clearing and settling in the interbank funding market. The second peak, the Financial Sector Conduct Authority,[3] is responsible for market conduct, regulation and supervision orientated toward financial consumer protection. This paper critically appraises the SARB's mandate to maintain financial stability, coordinate with other regulatory and supervisory bodies (including relevant departments within the SARB) and implement regulatory instruments for macroprudential policy.[4] We summarise the risks and vulnerabilities to the resilience and functioning of the system and appraise

* We are especially grateful to Greg Farrell, Roy Havemann, Charl Hugo, Krige Siebrits and two anonymous referees for the Annual Banking Law Update for their detailed comments and suggestions. In addition, we would like to thank Onesmo MacKenzie, Lisa Martin, Jamie Millar and Roan Minnie for their research assistance in helping to prepare this chapter. Finally, we would like to acknowledge conference participants at the 2018 Annual Banking Law Update (Centre for Banking Law, University of Johannesburg) and the South African Reserve Bank participants at the Monetary Economics Seminar Series (Stellenbosch University) for their useful feedback in the early stages of this chapter.
** Lecturer, Stellenbosch University, Department of Economics.
*** Lecturer, Stellenbosch University, Department of Economics.
1 The common acronym is SARB.
2 The common acronym is PA.
3 The common acronym is FSCA.
4 The common acronym is MaPP.

ABLU 2019 (Part ONE).indb 49 2019/10/09 9:49 AM

the SARB's approach to mitigating unintended consequences, with respect to both the institutional design and implementation of MaPP.

South Africa has a well-established regulatory and supervisory system compliant with international regulatory best practice in banking, insurance and securities regulation. Financial institutions are generally well capitalised and equipped to weather liquidity stresses, as observed during recent systemic events such as the global financial crisis of 2007–09, the European sovereign debt crisis from 2010, and the emerging market "taper tantrum" episode in response to US Federal Reserve monetary policy normalisation. Fissures in the domestic financial system appear to be isolated to unsecured lending and micro-lending activities, with African Bank, as an example, placed under curatorship in August 2014 due to significant wholesale funding shortages.[5] Although clearly resilient, and notwithstanding the politico- and socio-economic climate since 2011,[6] the resilience and unabated provision of intermediation services of the financial sector faces a number of challenges going forward.

Current risks and vulnerabilities to financial stability, as identified by the SARB[7] in their financial stability report, include the precarious domestic fiscal position, low growth levels, the associated decline in the quality of assets on the balance sheets of banks, a sharp increase in global risk premia, and the potential impact of protectionist policies stemming from the United States and the resulting impact on trade agreements.[8] Risks emphasised in their report are more in favour of shocks emanating outside the borders of the country than endemic risk arising because of local behaviour. This is indicative of the nature of shocks experienced historically by most small open economies, where one broadly divides the origination of risk into domestic and international origins. We focus on three domestic originations of risk to which the SARB could directly counteract: market concentration and concentration risk, lending risk, and funding liquidity risk. Counteracting these risks pose trade-offs between financial stability and real economic development that may or may not be welfare improving.

It is also important to realise that all of these risks and vulnerabilities are aggregate market (*ie systematic*) risks that contribute to *systemic* risk. Moreover, these risks can even trigger a *systemic event.*[9] But the presence of systematic risks

[5] South Africa experienced a small banking crisis and liquidity shortages from 2000–2002 – the most significant insolvencies were Saambou, Board of Executors and UniFer. See Schoombee "South African banks and the unbanked: Progress and prospects" 2004 *South African Journal of Economics* 72 at 581–603. More recently, in March 2018, VBS Mutual Bank was placed under curatorship for similar lackluster lending standards.

[6] These include the police shooting at Marikana, the mining charter, populism, public finance constraints and growing economic inequality.

[7] SARB *Financial Stability Report: First Edition* (2018a).

[8] Other systemic risks, identified as less likely to occur, include global geopolitical events (*eg* Brexit), domestic political uncertainty, and land expropriation without compensation.

[9] According to s 1 of the FSR Act a systemic event "means an event or circumstance, including one that occurs or arises outside the Republic, that may reasonably be expected to have a substantial adverse effect on the financial system or on economic activity in the Republic, including an event or circumstance that leads to a loss of confidence that operators of, or participants in, payment

are not mutually exclusive, and therefore do not have a unique mapping from a specific policy instrument to a specific risk. MaPP therefore creates conflicts between its own instrument-dependent intermediate objectives and other policy objectives. As such, there is now wide acknowledgement for the need to coordinate macroprudential policy with microprudential regulation and supervision, monetary policy, fiscal policy, and structural policies.[10]

Our discussion raises two broad themes related to this policy implementation and coordination problem. On one hand, financial sector regulation and supervision should correct incentive compatibilities that lead to market failures. These perverse incentives may arise from, for example, banking sector objectives to maximise the return on shareholder equity, a lack of competition, or risk-shifting behaviour. To address such issues, most central banks have access to a wide range of targeted macroprudential policy instruments (*ie* "tools"). But the selection and implementation of tools is complex, and their direct and indirect transmission channels are not well understood. This uncertainty can generate policy coordination failures between MaPP instruments (regulatory arbitrage) and outright conflicting macroeconomic policies. In other words, there are unintended consequences that MaPP can have on the financial sector, the real economy, and other macroeconomic policies.[11] On the other hand, even if financial imbalances can be identified, it is nearly impossible to measure the costs and benefits of using alternative MaPP instruments *ex ante*.[12] As such, there are weak incentives to take potentially costly actions. This "inaction bias" is an important challenge for macroprudential policy and includes undesirable interactions with the political cycle, which is not discussed here.[13] Instead, we focus on how the implementation of MaPP and its coordination with other policies can mitigate unintended policy consequences. A prudent approach to MaPP is clearly desirable, but faced with this possible inaction bias, the question remains on how the SARB can limit the probability and severity of financial crises.

In summary, the SARB has performed in a restrained manner when it comes to implementation of macroprudential policies. They have only implemented a handful of policy tools to abide by international standards and deal with potential pressures exerted from domestic and international sources. The main reason for caution is

systems, settlement systems or financial markets, or financial institutions, are able to continue to provide financial products or financial services, or services provided by a market infrastructure".

[10] BIS "Moving forward with macroprudential frameworks" 2018 *BIS Annual Economic Report* 63; additionally, Tucker "What is macroprudential policy for? Making it safe for central bankers" 2017 *BIS Papers* No 91 at 5.

[11] We emphasise market concentration and concentration risk, lending risk, funding liquidity risks, and the intensive and extensive margins of macroprudential policy coordination.

[12] These tools (or policy instruments) are often tailored for specific sectors, regions and institutions, and can be classified as capital-based instruments, asset-side tools and liquidity-based instruments. There is a wide array of tools, with multiple intermediate targets, all charged with the same final objective of financial stability. See Arslan and Upper "Macroprudential frameworks: Implementation and effectiveness" 2017 *BIS Papers* No 94 at 25; Villar "Macroprudential frameworks: Objectives, decisions and policy interactions" 2017 *BIS Papers* No 94 at 7.

[13] Szpunar "Institutional and operational aspects of macroprudential policy in Central and Eastern European EU member states" 2017 *BIS Papers* No 94 at 289.

that macroprudential policy can create perverse incentives between imposed policy instruments, which can conflict with other macroeconomic policies as well. MaPP therefore cannot be seen as a panacea to all financial instability woes and must, in particular, be coordinated with and subordinate to monetary policy. But given the structure, size and international integration of South Africa's financial sector, it is clear that supervision and regulation is needed.

The key to mitigating the probability and severity of financial crises is to reduce the build-up of imbalances (*ie* build resiliency through incentive compatible instruments and effective supervision) and contain financial distress that allows for a dynamic and innovative financial system (*ie* clear crisis management and resolution rules that minimise any implicit or explicit bail-out or too-big-to-fail guarantee). We see little scope, currently, for a strong "leaning against the financial cycle" approach (*ie* the active use of the countercyclical capital buffer) and advocate, instead, for a macro-financial (or "whole-economy") approach to macroprudential policy, which emphasises independent yet close coordination with other macroeconomic policies. Once again, the important caveat is that MaPP should, in general, be subordinate to monetary policy when conflicts arise between their objectives. Finally, financial stability can be a mandate of the central bank, but it cannot be the objective of monetary policy. The purpose of this division is twofold. First, monetary policy is ill-equipped to combat financial instability and its policy objectives can conflict with the promotion of financial stability. Secondly, independent decision-making bodies separate accountability for achieving their respective goals. For example, this independence mitigates the spill-over of credibility erosion on both monetary policy, in the event of a financial crisis, and financial stability, in the event of a recession or temporary inflation. The rest of this chapter outlines the different challenges faced by the regulatory authority and how we believe they should deal with future concerns.

2 Rationale for focus on financial stability

Many central banks across the world have had to shoulder the burden of financial instability, either *de jure* by institutional design or *de facto* through public perception.[14] Legal objectives for central banks with respect to financial stability, however, are "generally vague, do not define success or failure, and say nothing about competing objectives".[15] Decades of research in monetary policy has taught us that central banks need an appropriate, well-defined objective to remain accountable and, by extension, to be regarded as credible institutions.[16] When curating these objectives, it needs to be considered that macroprudential measures available to central banks are meant to deal with financial instability in a preventative sense, rather than trying to manage risk once it manifests.

[14] A common example would be *de jure* deposit insurance schemes purposed to prevent traditional bank runs versus a *de facto* bail-out premium for systemically important financial institutions.

[15] Upper "Macroprudential frameworks, implementation and relationship with other policies: Overview" 2017 *BIS Papers* No 94 at 1–5.

[16] Villar (n 12) 9.

In this section, we first define financial stability to give context for the discussion on the most appropriate way for policy-makers to address financial externalities. After that, we explore the reasons why the current incarnation of monetary policy is not equipped to deal with preventing the build-up of systemic risk in the financial system. Finally, we discuss the role of macroprudential policy in maintaining financial stability. In particular, this section outlines the way in which we believe macroprudential policy can be most effectively framed.

2.1 *Defining financial stability*

At this point there is little convergence in the literature on a true definition of financial stability. Financial stability is most often defined as the lack of financial fragility or systemic risk. Some would argue that this only shifts the burden of definition to a different, similarly vague, notion. Nonetheless, we adopt this approach and define financial stability as the lack of systemic risk. In this setting, systemic risk relates directly to possible impairment of the financial system, and by extension the broader macroeconomy. Systemic risk arises endogenously, for example, in the form of *ex ante* correlated risk choices by agents in the financial and banking system (strategic complementarities) or a coordinated interbank liquidity run (asset fire sales and credit crunches). In addition, it could also be the result of exogenous shocks, such as a surge in foreign capital flows, which originate outside of the system. The financial system here refers primarily to financial intermediaries and financial markets but can extend to any systemically important financial institutions in the economy.

Given our discussion thus far, the best way for policy-makers to think about financial instability is in terms of the externalities that are generated by a build-up of systemic risk. It is also important to realise that there are two dimensions to systemic risk. First, there is systemic risk that evolves over time, normally during periods of increased credit extension, accommodative monetary policies and unsustainable asset price growth. One example of this is the low policy rates of the early 2000s in the US. In this case, these low rates were the result of historically low inflation during the Great Moderation and an active attempt by the Federal Reserve to dispel deflationary concerns in the wake of the mild 2001 recession. In fact, Taylor[17] argues that the policy rate was significantly lower than prescribed by an optimal interest rate setting rule. Such an environment could plausibly induce a risk-taking attitude of investors in several ways, which Borio and Zhu[18] call the "risk-taking channel" of monetary policy.

Secondly, there is a cross-sectional dimension, which captures negative externalities from contagion and spill-over effects.[19] Identifying the source of market failure will help regulatory authorities determine the appropriate policy

17 Taylor "The explanatory power of monetary policy rules" 2007 *NBER Working Paper* 13685.

18 Borio and Zhu "Capital regulation, risk-taking and monetary policy: A missing link in the transmission mechanism?" (2012) 8 *Journal of Financial Stability* 236–251.

19 Freixas, Laeven and Peydró *Systemic Risk, Crises and Macroprudential Regulation* (2015) 1–487.

tool. In the next section, we further develop ideas surrounding these externalities and how prudential authorities can potentially prevent perverse incentives.

Externalities (market failures)

In the economic landscape after the financial crisis, policy-makers have been forced to develop tools that deal with externalities generated from financial activity along both a time series and cross-sectional dimension. Claessens[20] provides a classification of externalities along the following lines in his discussion on financial instability and the role of macroprudential policy. First, we have externalities which are generated by borrowers that are unable to see fully the impact of their borrowing decisions on asset prices. In particular, we are referring to borrowers that leverage in a procyclical fashion. This behaviour can lead to potential fire sales of assets, and derivatives based on these assets, once asset prices stall or start to decline.[21] During a contractionary phase of the financial cycle, collateralised borrowing and financing is adversely affected as a result of the weakened balance sheets of financial intermediaries.

Secondly, we have externalities related to strategic complementarities. Externalities of this kind reflect the strategic interaction between banks and other financial market participants which result in a build-up of risk that correlates with the expansion of the financial cycle. While Claessens reserves strategic complementarities as a different class of externalities, Galati and Moessner[22] argue that strategic complementarities are simply an amplification mechanism once fire sales have started. Thirdly, along with the cross-sectional or structural dimension, we have that externalities related to interconnectedness and contagion are of significance. This reflects how financial shocks transmit to systemic institutions and financial agents through their established connection of networks. In the section that follows we discuss the role that monetary policy plays in addressing these externalities.

2.2 *Are monetary authorities equipped to maintain financial stability?*

Historically, policy-makers were concerned with both price and financial stability. In fact, as argued by Goodhart,[23] central banks were initially created to prevent financial crises and bank failures. Central banks were designed with the unique ability to generate liquidity, in the form of bank reserves, providing them with a monopoly over the issuance of their liabilities.[24] As originally envisaged, the principal role for the central bank is the provision of liquidity to key financial

[20] Claessens "An overview of macroprudential policy tools" (2015) 7 *Annual Review of Financial Economics* 397–422.

[21] Galati and Moessner "Macroprudential policy – a literature review" (2013) 7 *Journal of Economic Surveys* 846–878.

[22] *Ibid.*

[23] Goodhart *The Evolution of Central Banks* (1988) 1–218.

[24] Bank for International Settlements "Re-thinking the lender of last resort" 2014 *BIS Papers* No 79 at 1–10.

institutions in times of crisis, the so-called "lender-of-last-resort" function as first described by Thornton[25] and Bagehot.[26] Framed in this way, achieving financial stability is at the heart of monetary policy. In this instance, financial stability can be viewed as a supply-side constraint in credit markets whereby financial intermediation is potentially interrupted. The apparent solution to this problem is for the central bank to issue liabilities to resolve this disruption in intermediation.

However, during the latter quarter of the 20th century, several arguments arose that lead to muted discussions on the central bank's role in achieving financial stability. First, central banks were too narrowly focused on price stability. As a result, financial sector risk was not taken into account in determining the appropriate stance of monetary policy.[27] It is further widely accepted that the capacity of the central bank to combat the build-up of financial instability with conventional policy tools is limited.[28] For example, to combat housing price increases, the magnitude of the change in the nominal short-term interest rate might either be too large or unnecessary for its inflation objective. Indeed, in their seminal article, Bernanke and Gertler[29] argue that monetary policy should be concerned only with factors that could plausibly influence the future path of inflation. In their study, they found that the central bank gains relatively little from responding to asset prices, and it should consider asset price fluctuations only in its capacity to affect the forecast of inflation, referred to as the "benign neglect" approach. In addition, the increase in the policy interest rate might impact asset classes beyond the one where a bubble is developing.[30]

This means that targeted instruments found in macroprudential regulation would perhaps be more appropriate. In fact, after the financial crisis there was a resurgence in the literature on the interaction of monetary policy and financial stability. In an article by Smets,[31] he argues that "price stability has proven not to be a sufficient condition for financial stability and lack of financial stability can have large negative feedback effects on price stability". In his article he calls for macroprudential regulation to run complementary to monetary policy in dealing with the build-up of financial imbalances. Monetary policy should be able to

25 Thornton *An Enquiry into the Nature and Effects of the Paper Credit of Great Britain* (1802).

26 Bagehot *Lombard Street: A Description of the Money Market* (1873).

27 Borio "Central banking post-crisis: What compass for uncharted waters?" 2011 *BIS Working Papers* No 353 at 2–3.

28 Woodford "Inflation targeting and financial stability" *The Future of Central Banking at the Einaudi Institute for Economics and Finance* (2010) 1–34.

29 Bernanke and Gertler "Should central banks respond to movements in asset prices?" (2001) 91 *American Economic Review* 253–257.

30 The term "bubble" does not necessarily imply irrationality or market failure. Rather, a bubble may be an equilibrium state and even optimal for funding (Martin and Ventura "Managing credit bubbles" (2016) 14 *Journal of the European Economic Association* 753–789; Martin and Ventura "Economic growth with bubbles" (2012) 102 *American Economic Review* 3033–3058). Alternatively, *ex post*, one can characterise a bubble as a misallocation of resources that leads to a build-up of financial imbalances. See Woodford (n 28) 3.

31 Smets "Financial stability and monetary policy: How closely interlinked?" (2014) 10 *International Journal of Central Banking* 263–264.

"lean-against-the-wind" in the short-run, coordinated with macroprudential policy, while focusing on price stability in the medium-term.[32]

Secondly, measurement of the build-up of risk has been problematic. For example, it has proven almost impossible to identify asset price "bubbles" until they have burst. Without a proper method for identifying bubbles, it is not considered worthwhile for the central bank to try and lean against asset price increases through contractionary policy. This has led academics and policy-makers alike to suggest mopping up after the bubble has burst. However, as evidenced by the recent crisis, this might prove too costly. On the other hand, one thing gained from the crisis is that the overvaluation of an asset and the accompanying drop in price is not always the issue that needs to be addressed. The important consideration is the development of systemic risk that poses a threat to the health of the overall financial system, in other words, the joint failure of systemically important financial institutions. In this sense, there have been significant improvements in the measurement of risk to financial stability.[33]

Thirdly, central banks generally adhere to the Tinbergen principle of one instrument (tool) for one target (goal). One tool for two goals creates "conceptual and practical" confusion as to the ultimate objective, with communication becoming increasingly difficult.[34] Rather, the fact that there are cyclical differences in intermediate objectives – such as consumer price inflation, housing price growth and total credit growth – and varying effects and types of instruments speaks to a multifaceted but coordinated approach to the two policies. Before the global financial crisis (GFC) central banks largely adopted an overnight interest rate as the tool of monetary policy, and therefore had no power beyond their lender-of-last-resort function to navigate the financial stability space. Microprudential regulation was thought to complement monetary policy and take care of idiosyncratic financial stability concerns. However, once the crisis had hit, this idea surrounding the tools available to the monetary authority with respect to financial stability changed dramatically, bringing the balance sheet of the central bank and macroprudential policy into contention.[35]

2.3 *Using macroprudential policies to combat financial instability*

Having established that interest rates are a blunt instrument against asset price fluctuations and that monetary authorities should be focusing their policy tool at maintaining price stability, we turn our attention to macroprudential policy. Financial regulation before the financial crisis took a microfocused perspective on risk. There was a focus on the health of the balance sheets of individual financial institutions, rather than a holistic understanding of the financial system and its

[32] Smets (n 31) 267.

[33] Woodford (n 28) 1–34.

[34] Svensson "Monetary policy and macroprudential policy: Different and separate?" (2018) 51 *Canadian Journal of Economics* 802–827.

[35] Blanchard "Macroprudential policies in a global perspective" 2014 *NBER Working Paper Series* No 19967 at 1–38

interconnected web of networks.[36] However, after the crisis, it was ascertained that other externalities, such as those outlined above, were at the heart of financial instability and that microprudential regulation as espoused by the first two Basel Accords would need to be reconsidered. Capital adequacy ratios were considered as sufficient as they generated buffers to protect individual institutions and by extension the entire financial system.

According to the renewed view that financial cycles were the driving force behind the recent crisis,[37] financial instability in an economy is generated as a product of, most commonly, excessive risk-taking. It is therefore possible to take preventative measures to combat this type of behaviour. In other words, risks arise in this setting because of perverse incentives, which can to a certain extent be corrected by a regulatory body. In particular, macroprudential policies are seen as "those policies aiming to reduce systemic risks arising from 'excessive' financial procyclicality and from interconnections and other 'cross-sectional' factors".[38] Policy-makers are then tasked with understanding the sources of increased risk-taking and the build-up of systemic risk in financial markets.

There are two general channels identified in the literature.[39] First, the preference channel, by which asset price bubbles originate from investor behaviour that is explained by the tenets of behavioural finance. Motivation for asset price bubbles in this framework include concepts such as irrational exuberance, which reflects an overoptimistic view of the market in good times, while almost entirely neglecting tail risk. This does not mean that preference shifts are always linked to irrationality. There are various theories, such as those that incorporate habit formation, where agents are considered fully rational. In these types of models, financial market participants are less risk-averse during a boom period.

The second explanation forwarded for the pervasiveness of growth in credit and asset prices is limited liability on the behaviour of financial intermediaries, causing them to become highly leveraged. Conventionally, when yields on safe assets are low, investors substitute toward higher-yielding risky assets, a phenomenon that was recorded in the build-up to the crisis, described as a "search-for-yield".[40] This was compounded by the fact that, as suggested by several measures of implied volatility, perceived risk was at an all-time low during the Great Moderation.[41]

36 Freixas et al (n 19) 212.

37 Schularick and Taylor "Credit booms gone bust: Monetary policy, leverage cycles and financial crises, 1870–2008" (2012) 102 *American Economic Review* 1029–1061.

38 Claessens (n 20) 398.

39 See, in general, Freixas et al (n 19).

40 Rajan "Has financial development made the world riskier?" 2005 *NBER Working Paper Series* No 11728 at 3; Shirakawa "Central banking: Before, during, and after the crisis" (2013) 9 *International Journal of Central Banking* 385.

41 Bean, Paustian, Penalver and Taylor "Monetary policy after the fall" 2010 *Proceedings – Economic Policy Symposium – Jackson Hole* 267–328.

In addition, as argued by Adrian and Shin[42] and Moench, Adrian and Shin[43], the increase in the price of risky assets improves the balance-sheet position of financial intermediaries and encourages them to take on more debt (either through the extension of loans or the acquisition of securities), which in turn fuels further asset price increases. This effect was amplified by the procyclical capital requirements of the Basel II Accord. Over time, owing to the limited number of "safe borrowers" in an economy and the depressed interest margins of commercial and investment banks, increased loan provision translates into increased funding of risky projects, inducing a leverage cycle.[44] This is considered an agency view of risk-taking, which puts ideas such as moral hazard and adverse selection into play. As argued by Freixas et al,[45] in this environment "financial gains are privatised but losses are in large part socialised".

Macroprudential policy can address problems if they originate from financial cycles but will struggle to provide useful assistance if they try to regulate activity in the business cycle. In other words, preventative measures can be used in the case of financial cycles, but these policies are ineffective in managing the business cycle. These types of policies can create significant market distortions, often negating the effects of other policy measures, such as monetary and fiscal policy. Macroprudential policies should then be implemented as a preventative (and subordinate) measure in coordination with other policies to regulate financial market instability. Crucially, this class of policy should not be enacted in a reactive fashion. If implemented reactively, it could deepen the liquidity problems that have to be resolved in times of crisis (*eg* Basel II Accords forced unreasonably high capital requirements during a downturn when value of bank capital was declining).

Regulation and supervision of the financial system can be characterised to that of firefighters and forest fires. In this analogy, policy-makers are firefighters and the sources of financial instability are the fires that they wish to extinguish. Given this setting, should firefighters fight forest fires to minimise immediate costs (reactive policy) or contain them to maximise long-run growth (preventative policy)? At the turn of the century, the strategy of firefighters changed dramatically from trying to prevent every fire and make as small as possible the damage to trees and wildlife (*ie* economic assets) to simply managing and containing fires to limit the build-up of debris (*ie* vulnerabilities) which fosters rejuvenation, growth and resilience. There were high costs associated with the former strategy, and both elevated systemic risk and active, resource-intensive management.

Most policy-makers have taken a more preventative approach, but maintain a significant degree of discretionary power. In fact, describing central bankers as "crisis managers" and "firefighters" implies exactly that.[46] For this reason,

[42] Adrian and Shin "Financial intermediaries, financial stability, and monetary policy" 2008 *Federal Reserve Bank of New York Staff Reports* No 346 at 1–37.

[43] Moench, Adrain and Shin "Macro risk premium and intermediary balance sheet quantities" 2010 *Federal Reserve Bank of New York Staff Reports* No 428 at 1–37.

[44] Bean et al (n 41) 272.

[45] See above (n 19) 85.

[46] Chorafas *The Changing Role of Central Banks* (2013) 1–304.

monetary policy must be clearly delineated from macroprudential policy. The former involves maintaining nominal stability and being a lender-of-last-resort. The latter involves credit policies, crisis management and resolution. A second temptation stems from technological advances that have fast-tracked digitisation and microscopic monitoring of the financial sector. Under this presumption of precision, the temptation to fine-tune policies and to foster centralisation (in terms of financial system concentration and infrastructure) must be avoided. If not, it can delegitimise the hard-fought credibility and institutional independence of monetary policy and create a system with a concentrated point of failure.

The core of this message is that macroprudential policies are inherently distortionary. The role of policy-makers is to weigh the benefits and costs of this distortion. If the financial system is already resilient, it is not clear that these policies prevent instability and losses. On the contrary, these policies might lead to a system that allows for failures and no bail-outs. As argued by Claessens,[47] unless firefighters use the appropriate equipment to extinguish the fires they can "worsen some resource allocations. And by constraining actions of agents, they can increase overall systemic risks." Identifying the precise source of the externality is crucial in this regard and will pose unique challenges for each country that implements these measures.

Theoretical development of issues surrounding macroprudential policy is in its infancy. Discussion surrounding macroprudential policy is in a similar stage of development that monetary policy was during the 1940s. Bean[48] argues that in comparison to our development of thought on issues of monetary and fiscal policy, we are "still in the Stone Age in respect of deploying macroprudential policies". One way to think about macroprudential policy then is to frame it within the same setting as monetary policy. We can start thinking along the dimensions of the primary objective, intermediate targets and instruments required. The objective would be the same for most countries: the prevention of systemic risk (increasing system-wide stability). Contained in this objective is the goal of "limiting macroeconomic costs from financial distress".[49] This definition of the objective clearly delineates the importance of thinking of economic growth being at risk during times of financial instability. This means that financial sector volatility can have real consequences, and by implication, macroprudential policies will be indirectly aimed at promoting growth over the longer run.

The source of systemic risk is not the same for all nations, however, and would then mean different intermediate targets and instruments implemented. Developed nations will tend to consider endogenous sources of risk and therefore try to shield against the build-up of risk by using specific tools that provide a well-capitalised financial sector and the ability to monitor the probability of default

[47] See above (n 20) 398.

[48] Bean "Central banking in boom and slump" 4, speech by Mr Charles Bean, Deputy Governor for Monetary Policy of the Bank of England, at the JSG Wilson Lecture in Economics, University of Hull, Hull, 31 October 2012 (https://www.bis.org/review/r121102e.pdf). See also Galati and Moessner (n 21) 846–847.

[49] Galati and Moessner (n 21) 853.

among institutions. In addition, these countries might place a higher weight on the interconnectedness of financial intermediaries and non-banks in order to prevent contagion. In developing countries, the focus might shift toward external factors that could potentially disrupt financial market activity. In the next section, we will take a deeper look at the risks that are specifically relevant for South Africa and the tools that have been utilised to assuage these concerns.

3 Macroprudential policy in South Africa: Institutional structure, goals and decision-making

In this section, we discuss the SARB in South Africa's post-1994 dispensation and provide a summary of the recent developments surrounding macroprudential policy (MaPP). Section 3.1 describes the institutional evolution of the SARB in response to financial system instability generated before and after the recent financial crisis. Section 3.2 contextualises South Africa's macroprudential framework and implementation in response to the GFC.[50]

3.1 *Institutional structure*

In response to the series of emerging market crises from 1995 through to 2001, global financial stability concerns heightened and a concerted effort began to promote international prudential standards.[51] Yet, unlike the G20 countries, and even peer emerging market economies (EMEs), the SARB has had limited experience in implementing macroprudential policies.[52] Since the turn of the century EMEs have

[50] Non-market (direct) control measures in credit and currency markets, in particular, as well as opaque operating procedures characterised the SARB prior to Dr CL Stals's appointment as Reserve Bank Governor in August 1989. See Mollentze "Monetary policy in South Africa on the threshold of a new era" 2000 *South African Journal of Economic and Management Sciences* SS No 2 at 1–50; SARB "South African Reserve Bank commemorative publication" 2011 *South African Reserve Bank Publications* 1–155. It is not clear how financial stability objectives and intervention influenced monetary policy during this period. Aside, the crisis management and resolution of the 1985 corporate debt crisis was unique and instrumental for policy globally (see Harris "South Africa's external debt crisis" (1986) 8 *Third World Quarterly* 793–817). It is also important to note that direct monetary controls were in use between 1965 and 1980 (Mollentze S-6). These included ceilings on bank credit to the private sector, deposit rate control, foreign exchange control and outright control of hire-purchase and consumer credit. According to Mollentze (S-6) banks were intermittently requested to be selective in their credit extension.

[51] These events led to the creation of the Financial Stability Forum (FSF), the Financial Stability Board's predecessor, in February 1999 by G7 finance officials. In response to mounting legitimacy issues, the Financial Stability Board was established in April 2009 by the Group of 20 (G20) countries. See Helleiner "The Financial Stability Board and International Standards" 2010 *The Centre for International Governance Innovation G20 Papers No 1* at 1–27; Frankel "Monetary policy in emerging markets" in Friedman and Woodford (eds) *Handbook of Monetary Economics* (2011).

[52] Ceruttia, Claessens and Laevenc "The use and effectiveness of macroprudential policies: New evidence" (2017) 28 *Journal of Financial Stability* 203–224; Lombardi and Siklos "Benchmarking macroprudential policies: An initial assessment" (2016) 27 *Journal of Financial Stability* 35–49; Havemann "Counter-cyclical capital buffers and interest-rate policy as complements – the experience of South Africa" 2014 *ERSA Working Paper* 476 at 1–22. Furthermore, unlike its peer

more frequently adopted macroprudential measures related to foreign exchange deposits and credit growth, whereas advanced economies have concentrated more on borrower-based credit constraints such as loan-to-value ratios. Both groups generally favour limits on funding from key borrowers (concentration risks), or more specifically, non-bank to bank funding (wholesale funding exposure), as well as limits to leverage. With regards to international prudential standards, however, the SARB has been at the helm of peer EME countries.

In August 1999, the SARB announced its intention to align its monetary policy framework with global developments. The adoption of an explicit inflation-targeting framework coincided with its efforts to position itself within global regulatory and supervisory standards set out in the Basel Accords. The Banks Act 94 of 1990, along with exchange control regulations, provided financial institutions with a strong buffer to absorb both internal and international shocks.[53] There is some *de facto* evidence that from 2003 to 2006, in response to credit growth concerns, the SARB took measures to raise bank capital adequacy ratios.[54] However, due to the overwhelming nature of the shocks generated by the global financial crisis, it was not possible to completely shield South Africa's financial system.[55] In response to the GFC, the Minister of Finance reaffirmed the SARB's role as the nation's macroprudential supervisor. By 2013, with the National Treasury's publication of the proposed "Twin Peaks" model, the SARB was committed to and had already begun re-orientating its existing regulatory framework to address liquidity (funding) and credit risks, as well as investigating the potential for unintended regulatory arbitrage.[56]

With the FSR Act of 2017 financial stability responsibilities are legally delegated to the SARB. Notably, the internal Financial Stability Committee[57] in the SARB, with the Governor as chair, holds executive power to advise regulatory bodies to implement MaPP and to coordinate financial stability objectives with microprudential and monetary policy.[58] In the case of a systemic event the FSR Act provides the SARB with directive powers over financial regulators. Of 24 surveyed EME central banks, in a 2016 Bank for International Settlements[59] questionnaire,

EMEs (Argentina, Brazil, Chile, China, Colombia, Peru and Turkey) the SARB does not include, or use, monetary policy instruments (*eg* reserve requirements on domestic and foreign deposits) as part of their MaPP toolkit or with the aim to stabilise financial conditions (see Villar (n 12) 11).

53 National Treasury "A safer financial sector to serve South Africa better" 2011 *National Treasury Policy Document* 13–15.

54 Havemann (n 52) 7–8.

55 It is not obvious to what extent South Africa's 2009 recession was linked directly to the GFC and capital flows (financial channel) versus that of global demand for goods and services and commodity prices (trade channel).

56 Havemann (n 52) 7–8; National Treasury (n 52) 24.

57 The common acronym is FSC.

58 Established in 2000, the FSC was recently restructured in accordance with the SARB's enhanced financial stability mandate. Currently, MaPP is subordinate to and supportive of monetary policy and microprudential policy.

59 The common accronym is BIS.

South Africa is one of 13 that have full control over macroprudential tools.[60] Brazil and South Africa are the only two that share decision-making responsibilities with the banking supervisor (the Chief Executive Officer of the Prudential Authority in South Africa's case) and other regulatory bodies. In South Africa, members of the FSC overlap with the Monetary Policy Committee[61] and include senior SARB officials from relevant departments within the SARB.[62] The most common coordination approach taken by central banks are inter-agency committees. Villar[63] identifies 14 out of 24 countries with inter-agency committees in which the central bank governor either chairs (as in South Africa) or takes a lead role.[64]

The Financial Stability Oversight Committee[65] is an advisory committee that includes members from the SARB, National Treasury and financial regulators.[66] The FSOC intends to meet every six months, to facilitate cooperation and collaboration between the financial sector regulators and the SARB in respect of matters relating to financial stability. Finally, 13 of the 24 central banks in the previously mentioned survey, now including South Africa, also have statutory mandates with a financial stability objective. These objectives range from being entirely broad (*eg* "promoting financial stability" or "reducing systemic risk") to narrowly defined objectives (*eg* the "normal functioning of internal and external payments" and "to regulate credit in the financial system").[67]

3.2 *A framework for macroprudential policy decision-making*

With the shift in focus to the prevention of systemic risk, macroprudential policy adopted the mechanism-design approach of monetary policy. This entails, as discussed more generally in section 2, first identifying the goal(s) of MaPP, then the related intermediate target(s), and finally the relevant instrument(s). The SARB, specifically, has adopted a clear three-step process to identify, motivate and respond to financial sector developments. First, it assesses systemic risk. Thereafter, it builds a case for MaPP intervention. Finally, the SARB selects and decides whether to implement the relevant MaPP instrument(s). Within this context, the SARB views MaPP as subordinate to and supportive of monetary policy and microprudential policy. In what follows, we discuss how the SARB defines its MaPP objectives (goals) and how this relates to "systemic risk" (step 1). We then contextualise its approach to identifying MaPP intervention within current realities (step 2). Our discussion on step 3 follows in section 4 below.

60 These include, amongst others: countercyclical capital buffers and capital requirements; margins and haircuts; sector-specific capital requirements for the banking sector; and debt service-to-income and loan-to-value ratios. See Villar (n 12) 7.

61 The common acronym is MPC.

62 SARB "Macroprudential frameworks, implementation and relationships with other policies" *BIS Papers* No 94 at 331.

63 See above (n 12) 12.

64 The effectiveness of such committees is not uniform and difficult to quantify.

65 The common acronym is FSOC.

66 See SARB (n 62) 331 fn 6.

67 Villar (n 12) 7.

The SARB's definition of financial stability stresses the "resilience" of and "confidence" in financial institutions and market infrastructures:

> "Financial stability refers to a financial system that is resilient to systemic shocks, facilitates efficient financial intermediation, and mitigates the macroeconomic costs of disruptions in such a way that confidence in the system is maintained."[68]

Maintaining the general provision and performance of services matters as well. That is, despite a changing environment, the SARB endeavours not only to maintain the functioning (capability) of the financial system but also to ensure confidence in its ability to do so. Notably, the SARB identifies the macroeconomic costs associated with *financial disruptions* as its ultimate welfare objective. This potential *growth-at-risk* is borne out with the SARB's emphasis on *systemic risk* as the focus of macroprudential policy:

> "'Systemic risk' is defined here as the risk of a disruption(s) to the provision of any of the key financial services that is caused by an impairment of a part(s) of the financial system or the financial system as a whole, and which can have serious consequences for the real economy."[69]

It is of importance to notice the distinction (as discussed in section 2 above) between *systemic risk* – which is wholly or in part unobserved and tends to build-up during the expansionary phase of the business cycle – and *financial disruptions* – which are realised outcomes in the financial sector and the real economy from the (endogenous) response of economic agents to externalities. In this light, the SARB recognises the origination of risks – both domestic and international – but emphasises MaPP instruments which target impediments to the provision of financial services that put economic growth at risk.

As we will discuss below, the SARB implicitly acknowledges that both systemic risk and the potential impact it may have on the real economy are difficult to identify *ex ante* (*ie* are typically only observed when they materialise). And, even more so than monetary policy, it is difficult to establish a stable link between instruments (on cross-sectional and time dimensions), intermediate objectives (the financial system, individual sectors, indicators or measurements) and final goals (the macroeconomic costs of financial disruptions). As such, the SARB focuses on prudence by taking a *preventative approach* to limiting systemic risks and mitigating externalities:

> "Two broad aims that are not mutually exclusive: first, strengthening the resilience of the financial system to economic downturns and other adverse aggregate shocks, and second, leaning against the financial cycle to limit both the accumulation of financial risks and the likelihood or the extent of a financial crisis."[70]

Here it is important to note that "strengthening the resilience of the financial system" does not imply that the central bank needs to identify systemic risks specifically or build a case for intervention with a specific MaPP instrument in mind. In fact,

[68] SARB *Financial Stability Review: First Edition* (2017) D.
[69] SARB (n 68) 33.
[70] *Ibid*.

Villar[71] points out that "central banks have more instruments at their disposal to strengthen the resilience of the financial system than to rein in financial booms".[72] This is not only an important reality but a desirable one. The SARB's attempts to identify financial cycle vulnerabilities, motivate policy intervention, and select effective instruments will need to be guided, at least initially, by a significant amount of discretion.[73]

The SARB is, however, clearly proactive with its surveillance of the financial system, and it actively seeks a high level of compliance with international standards outlined by the Basel Committee on Banking Supervision[74] and the Financial Stability Board. Indeed, existing relatively high capital and liquidity buffers for large banks and insurers, above Basel III requirements, has fostered a robust financial system and mitigated any pressures on the SARB actively to enforce any regulatory tools, both domestic and international. In section 4, we highlight key risks and vulnerabilities that the SARB faces and how it intends to respond to them (step 3). We also raise important risks and vulnerabilities not identified and/or clearly dealt with by the SARB, to which we hope to contribute to the SARB paving the way forward.

4 The way forward

In the second section of the paper we considered the role that macroprudential authorities could play in facilitating an environment conducive to financial stability. In what follows, we discuss how policy-makers within the South African context could potentially approach concerns of financial instability. Section 4.1 details unique characteristics of South Africa's financial system. In section 4.2, we specifically identify key risks and vulnerabilities in the South African economy that may lead to or currently justify a MaPP response. In section 4.3, we first highlight some of the unintended policy consequences of MaPP, and then consider how the SARB should manage these risks and vulnerabilities. We then contrast these realities with the current approach being followed by the SARB, as outlined in the previous section, and provide a critical evaluation of their actions.

71 See above (n 12) 11.

72 According to a BIS survey response (see Arslan and Upper (n 12) 41) the SARB measures vulnerabilities using the following tools: "(1) Risks in institutions identified as systemically important, shadow banks, asset markets and the non-financial sector. (2) Level of leverage, and general credit market conditions. (3) Maturity and currency mismatches. (4) Changes to lending standards. (5) Stress tests. (6) House prices, commercial property prices and asset valuations in equity markets. (7) Government and corporate bond spreads, credit default swap spreads and measures of risk premia. (8) Underwriting standards, and asset quality and credit conditions."

73 SARB (n 62).

74 The common acronym is BCBS.

4.1 *Financial markets and institutions: An overview of recent developments*

Market size, market innovation and international integration

A well-established regulatory framework goes hand-in-hand with a large, sophisticated and globally integrated financial sector.[75] In this regard, South Africa is well-placed with its peer emerging market economies. South African total financial sector assets amount to 305% of GDP (as of December 2017) where total banking assets make up 108% of GDP and total assets for non-bank financial institutions make up 197% of GDP.[76] In addition, total off-balance sheet activities of banks amount to 27% of GDP. Finally, the gross external position of the private sector at 283% of GDP, measured as the sum of total foreign assets and liabilities, highlights the degree of global integration.[77] The exposure of the banking sector to external positions are, however, muted. Most banking assets are domestic, long-term, and a mix of commercial and retail credit facilities and loans. Most banking liabilities are domestic, short-term and deposit financed.

Since 2013, non-bank financial institutions (henceforth NBFIs or non-banks) account for two-thirds of total assets. NBFIs are categorised as insurance companies, pension funds, public financial enterprises and other financial intermediaries (OFIs). NBFIs typically include so-called "shadow banking" activities sub-categorised into money market funds (MMFs), fixed income, multi-asset, funds of funds, hedge funds, finance companies, insurance and securitisation. Shadow banking is a term used to describe the services that NBFIs provide similar to that of "traditional" deposit-taking banks which fall outside banking regulations.[78] These shadow-banking activities amounted to R2.208 billion in the third quarter of 2016.[79] Notably, collective investment schemes, identified as being susceptible to funding liquidity shortfalls (*ie* "runs"), make up approximately 80% of this total figure. It is worth pointing out here that the global financial crisis of 2007–2008 predominately involved risk-taking behaviour in such market-based finance of NBFIs.[80] Taken

[75] Ceruttia et al (n 52); Lombardi and Siklos (n 52).

[76] We derive this value from SARB data for total bank and non-bank assets (KBP1132M and KBP2637K). This value is different from the inferred shares of total assets ascribed from each sector in SARB (2018b). Using the ratios of total financial assets for banks and non-banks (29.2% and 66.7% respectively) we find total assets to be between 295% and 370% of GDP.

[77] As of December 2017, South Africa's foreign liabilities were 138% of GDP, while the country's foreign assets amounted to 145% of GDP. The country's (positive) net international investment position was 7.4% of GDP at the end of 2017.

[78] See Kemp "Measuring shadow banking activities and exploring its interconnectedness with banks in South Africa" 2017 *South African Reserve Bank Occasional Paper Series OP/17/01* at 13–19 for an explanation of the SARB's narrow measure of shadow banking. The multi-asset category currently dominates shadow-banking activities at 47%, followed by funds of funds (13%), MMFs (13%), finance companies (12%) and fixed income (11%). It is important to note that the majority of shadow-banking entities or activities in South Africa are indeed regulated, and not all NBFIs activities are considered shadow-banking activities. In fact, some traditional bank activities fall under this definition of shadow banking.

[79] Kemp (n 78) 16.

[80] Adrian and Jones "Shadow banking and market-based finance" 2018 *International Monetary Fund, Monetary and Capital Markets Departmental Paper* No 18/14 at 1.

as given, however, these credit intermediation innovations reflect the needs and preferences of South African borrowers and lenders.[81] From this perspective, the South African economy exhibits a modern and innovative financial sector.

Market structure

The South African banking system is dominated by Standard Bank, Barclays/ABSA, Old Mutual (Nedbank), FirstRand Bank and Investec. These financial conglomerates maintain a 90% market share of total bank assets. The high concentration within the banking sector can be attributed to the high barriers to entry imposed by the Banks Act 94 of 1990. Notably, however, the traditional banking subsidiaries have seen a marked decline in the share of total financial assets from 37.6% for 2003 to 29.2% for 2017.[82] This market share decline, including the declining shares of insurance companies, pension funds and public financial enterprises, has been taken up by non-banks referred to by the SARB as "other financial institutions" (OFIs). Their share of total financial assets has risen from 8.4% for 2003 to 21.1% for 2017.[83] This rise in market-based finance mirrors that of the global trend before and after the global financial crisis, wherein shadow banking exhibits the weakest resilience.[84]

This phenomenon has occurred in South Africa, in particular, with the commensurate rise in non-bank (wholesale) funding to the banking sector. Non-bank claims on banks as a share of total non-bank assets is 20.1% as of 2016. This statistic places South Africa as the second largest wholesale funded banking sector out of the 27 advanced and emerging economies considered in the SARB's Financial Stability Report.[85] Recent work by IMF[86] and Kemp[87] document this high degree of interconnectedness between banks and non-bank financial institutions – with money market funds (MMFs) taking the predominant exposure.[88] Banks' increasing reliance on MMFs for short-term wholesale funding, as well as the general rise in off-balance sheet and shadow-banking activities, is likely both as a result of tighter regulation in the traditional banking sector and the search for yield of financial conglomerates – that is, to maintain an attractive return on equity (ROE). Notably, this robust increase in financial activity has persisted through a weakly performing economy. For example, the average ROE for all banks from 2001 to 2007 was 13% (over a period of rapid global and local economic growth) and approximately 16.5% from 2015 to 2018 (over a period of weak global and local economic growth). At the same time, there has been a marked increase in over-the-counter (OTC) foreign exchange (FX) and interest rate derivative trading.

[81] As our earlier discussion on market failures suggests, an inefficient allocation of financial products can lead to a net social welfare loss. That is, the needs and preferences of South African borrowers and lenders likely do not coincide with some more-efficient social outcome.

[82] SARB *Financial Stability Report: Second Edition* (2018b) 15.

[83] SARB (n 82) 15.

[84] Adrian and Jones (n 80).

[85] SARB (n 7) 11.

[86] IMF "South Africa: Financial system stability assessment" 2014 *IMF Country Report* No 14/340.

[87] See above (n 78).

[88] SARB (n 7) 10.

With respect to the provision of domestic banking services to households and non-financial firms, South Africa faces several unique structural pressures. Most notably, high unemployment and inequality make access to credit and even basic financial services provision difficult for the un-banked and under-banked (typically individuals living in non-urban areas and/or who are dependent on the informal sector). More generally, the stagnant economic performance of the country over the last decade has tightened credit conditions and produced an excess demand for funding.[89]

This apparent demand for banking services and various forms of financing has led to the proliferation of the micro-lending sector. As a result, this sector has seen at least two major financial distress episodes since 2000.[90] The first major episode occurred over the period 2000 to 2002 with the insolvencies and even voluntary relinquishment of bank licenses of several medium to small banks.[91] The failures of the 7th largest bank (Saambou) and then the 6th largest bank (Board of Executors) at the time were the most notable. From 1999 to 2003, the total number of registered banks operating in South Africa dwindled from 60 to 38.[92] Fissures in the domestic financial system appeared again in August 2014 when African Bank was placed under curatorship due to significant wholesale funding shortages. More recently, in March 2018, VBS Mutual Bank was placed under curatorship for similar lacklustre (and allegedly fraudulant) lending standards. On both counts, the SARB's decisive action limited contagion to the sector and the wider financial system. By not simply bailing out these institutions the SARB reduced any implicit too-big-to-fail (or *de facto* bail-out) premium. Overall, immediate financial sector risks and vulnerabilities appear isolated to this sector.

89 For large-, medium- and small-sized firms in urban areas, access to credit and banking services is less of an issue than structural issues related to electricity provision and perceptions related to corruption and the legal system (see, for example, World Bank Enterprise Surveys: South Africa Country Profile 2007 at http://www.enterprisesurveys.org). The stagnant economic performance of the country over the last decade, in particular, can be attributed to the erosion of business and consumer confidence. These factors stem from inefficient infrastructure investment and maintenance, political uncertainty, and the malfunctioning of key institutions and state-owned enterprises (see Bureau of Economic Research Stellenbosch at https://www.ber.ac.za/).

90 Schoombee (n 5) 586–587; Havemann (n 52) 21.

91 *Ibid*.

92 This was only in part due to the local banking crisis. Over this time, the banking sector was also consolidating under the so-called "four-pillar" policy (see Mboweni *The South African Banking Sector: An Overview of the Past 10 Years* (2004) 1 at 4). The idea is that a concentrated banking sector, with a minimum of four banks, makes prudential supervision easier, promotes resilience and limits the spread of risk.

4.2 *Risks, vulnerabilities and policy trade-offs*

Market concentration and concentration risk: Financial stability versus consumer welfare

South Africa's concentrated banking sector has been fostered by the so-called "four-pillar" policy,[93] the idea being that a concentrated banking sector of at least four "big banks" makes prudential supervision easier, promotes resilience and limits the spread of risk. Naturally, however, there are concerns around concentration risk in a bank's portfolio and high market concentration.

High concentration risk, whether on the asset side to a specific sector or on the liability side to a particular wholesale funding counterparty, implies low bank portfolio diversity, highly correlated returns, and therefore greater risk of a systemic event. In this sense, competition raises consumer welfare through financial sector diversification and minimising systemic externalities from financial institutions. High market concentration implies high barriers to entry, which tends to limit fruitful competition in the financial system. Reducing entry barriers reduces costs to the supply of services and funding to households and firms, and it incentivises the provision of financial technologies that broaden access. Another concern is that with market concentration comes greater market power (monopolistic competition) which may lead to unintended consequences for policy effectiveness and consumers of financial services. For example, market power can stifle monetary policy by limiting the pass-through of policy rate changes,[94] or the costs associated with funding a deposit insurance scheme could more easily be pushed onto consumers.

Therefore, without the appropriate amount of competition, key institutions become too large and the potential for rent-seeking and moral hazard is increased. That said, bank competition can also induce excessive risk-taking due to risk-shifting.[95] It is not clear which of these two dimensions (or possibly both, when one takes the view that banks operate globally) dominated during the 2008–2009 global financial crisis. Nevertheless, with a few large commercial banks in South Africa, policy-makers would have no choice but to rescue these systemically important institutions if they were to experience sudden liquidity shortages or become severely undercapitalised. Indeed, the SARB recognises the need for a clear resolution framework for designated financial institutions (*ie* designated resolution institutions). At this time, the SARB, with National Treasury oversight,

[93] Mboweni (n 92) 4.

[94] Hollander and Liu "Credit spread variability in the U.S. business cycle: The Great Moderation" (2016) 67 *Journal of Banking & Finance* 37–52; Hollander and Van Lill (forthcoming) "On the estimation and application of structural decompositions of the South African business cycle" in Boshoff (ed) *Business Cycles and Structural Change in South Africa: An Integrated View.*

[95] Because bank competition can lower the franchise value of a bank, higher volatility in asset returns can become more attractive. Feng ("Bank competition, risk taking and their consequences: Evidence from the U.S. mortgage and labor markets" 2018 *IMF Working Paper* WP/18/157 at 1–46) uses micro-level US mortgage data to show how banks operating in competitive mortgage markets lowered lending standards (*eg* the loan-to-income ratio and acceptence rate) twice as much from 2000 through 2005.

intends to draft a Special Resolution Bill which would cover registered banks,[96] non-bank financial institutions (including insurance companies), financial market infrastructures, and financial conglomerates.[97]

Lending risk: Financial stability versus access to credit

Access to credit is an important facet of financial inclusion in South Africa. Yet, under the auspices of financial inclusion, an excess demand for credit has led to the proliferation of micro-lending and unsecured loans. These short-term loans are normally provided to individuals and firms with below-average credit ratings, which implies higher risk premia priced into interest rates and higher probabilities of borrower defaults. As a result, banks inherit greater liquidity risk if actual loan losses significantly exceed loan loss provisions, and are therefore more likely to default (solvency risk) as well. In 2014, for example, African Bank was placed under curatorship because of its exposure to these risky loan portfolios. While these externalities generated little systemic risk in the form of contagion,[98] they did initiate a narrative in South Africa, as the special resolution framework suggests, around the role of the SARB in crisis management and resolution.

One widely implemented tool to prevent bank runs, especially for those institutions exposed to these unsecured loans, is that of an industry-funded deposit insurance scheme.[99] In fact, the SARB is currently underway with the designing of a DIS.[100] There are, however, well-known unintended consequences from implementing such a scheme.[101] For example, with an explicit deposit guarantee, depositors do not have an incentive to monitor the riskiness of their bank's assets. And given that deposit losses are "covered", both banks and depositors have an incentive to increase the aggregate level of risk in their portfolios.[102]

Turning to the general provision of credit for consumption and production activities, we observe exacerbated private domestic debt burdens. The deleveraging

[96] Registered banks refer to any bank registered in terms of the Banks Act 94 of 1990, a cooperative bank registered in terms of the Cooperative Banks Act 40 of 2007 or a mutual bank registered in terms of the Mutual Banks Act 124 of 1993.

[97] National Treasury "Financial Sector Laws Amendment Bill for public comment" 2018 *Strengthening South Africa's Resolution Framework for Financial Institutions* (http://www.treasury.gov.za/twinpeaks/ (accessed 10 May 2019)).

[98] Havemann "Can creditor bail-in trigger contagion? The experience of an emerging market" 2018 *Review of Finance* rfy023 (https://doi.org/10.1093/rof/rfy023).

[99] The common acronym is DIS. See Demirgüç-Kunt, Kane and Laeven "Deposit insurance database" 2014 *IMF Working Papers* 14/118, International Monetary Fund 1–43.

[100] SARB "Designing a deposit insurance scheme for South Africa – a discussion paper" 2017 Financial Stability Department, South African Reserve Bank 1–60.

[101] Anginer and Demirgüç-Kunt "Bank runs and moral hazard: A review of deposit insurance" 2018 *World Bank Group, Policy Research Working Paper* 8589 at 1–31.

[102] The SARB has proposed a DIS fund equivalent to 5% of "covered deposits", which, in 2016, amounted to approximately R17 billion (see SARB (n 100) 35). The perceived credibility of institutions is therefore still important for financial stability. And an important unintended consequence is to incentivise risk-shifting behaviour: most notably, an even greater skewed distribution toward a concentrated group of wholesale (non-bank) funders.

process by both borrowers and banks in response to the global financial crisis has led to markedly weak average credit growth of 6% (from January 2010 to December 2017) from a pre-crisis average of 20.6% (from March 2003 to December 2008). As a result, the banking sector is also exposed to demand-side (borrower) credit risk for two reasons. First, households and non-financial corporations (firms) are highly indebted and face rising debt servicing costs (for example, domestic and international upward pressure on interest rates can emanate from US monetary policy, sovereign debt downgrades, exchange rate uncertainty and higher domestic inflation). Secondly, there has been an erosion of non-financial sector collateral and creditworthiness (for example, weaker house price growth and weaker household incomes and firm profits). These demand-side factors put significant pressure on the whole economy. Current banking sector funding trends, however, suggest movement away from risk exposure on assets in the retail sector to assets held off-balance sheet or in wholesale markets.

Funding liquidity risk: Internal and external drains (or, the interplay between domestic financial stability and currency stability)

South Africa's financial system depends on access to external financing and over-the-counter[103] markets and is highly integrated with the global financial system, which means that cross-border capital flows have a large impact on the liquidity position of local institutions. As such, the South African economy can experience several shocks along the international dimension that constrain the ease with which financial institutions can obtain funding. This external funding risk is especially prominent in capital markets and OTC foreign exchange and interest rate derivatives, which can lead to episodes when access to foreign financing tightens considerably.[104] Equally important, and in contrast to typical concerns about "sudden stops" of capital inflows,[105] are risks associated with systemic liquidity drains, both internal and external. Internal drains occur during bank deposit runs as short-term obligations are converted into currency. External drains occur in capital flight episodes when domestic assets are converted into foreign assets. The threat of a "double drain scenario", as documented by Obstfeld, Shambaugh and Taylor,[106] sees foreign exchange reserves drained as residents use domestic bank deposits

103 The common acronym is OTC.

104 Funding liquidity risks need not relate to fundamental factors, but can emanate from contagion: *eg* from political instability in a systemic middle-income country like Turkey, which does not have a major trade or financial link with South Africa.

105 See Calvo and Reinhart "When capital inflows suddenly stop: Consequences and policy options" in Swoboda (ed) *Reforming the International Monetary and Financial System* (2000) 175–201. The literature on sudden stops typically highlight external short-term debt and trade openness as important predictors of currency crises. This led to what is known as the Guidotti-Greenspan rule for adequate foreign exchange reserves holdings for central banks. However, the high levels of international reserves we currently observe in emerging markets far exceed what these predictors would deem adequate. Obstfeld, Shambaugh and Taylor ("Financial stability, the trilemma, and international reserves" (2010) 2 *American Economic Journal: Macroeconomics* 57–94) show that financial stability and financial openness can account for this global reserve accumulation puzzle.

106 See above (n 105) 63.

(internal drain) to finance domestic capital flight (external drain). In their own words "domestic financial stability is inescapably a central consideration in reserve management policy [the Central Bank's function as lender-of-last-resort]" and "it is the threat of this type of drain that most worries emerging market policymakers".

Obstfeld, Shambaugh and Taylor[107] attribute the continued shoring up of international reserves by emerging markets to buffers against internal and external drains (bank deposit runs and capital flight). Essentially, the lender-of-last-resort[108] function of the central bank together with the size and openness of the domestic banking sector drives reserve accumulation – as opposed to the traditional trade channel.[109] In economies not operating within a fixed exchange rate regime, the rationale to shore up foreign liquidity buffers extends predominantly to public insolvency risks. On the one hand, Rodrik[110] argues that countries can avoid costly pecuniary externalities from reserve accumulation by implementing capital controls on short-term capital inflows.[111] On the other hand, Obstfeld et al[112] point out the difficulty of implementing such a policy and that foreign exchange reserve accumulation may be the intermittent social welfare improving insurance that emerging markets need in today's level of financial globalisation.

The SARB does not include, or use, reserve requirements on domestic and foreign deposits (that is, *monetary policy* instruments) as part of their MaPP toolkit or with the aim to stabilise financial conditions. That said, the reserve bank does maintain adequate foreign exchange reserves (as measured by the Guidotti-Greenspan ratio of a one-to-one ratio between reserves and short-term foreign debt obligations). But it is unclear how useful this will be for active MaPP given its direct conflict with monetary policy implementation – disentangling, in particular, macroprudential policy from monetary policy's LLR function. It is also unclear whether the SARB's continued use of the Guidotti-Greenspan ratio to maintain an "adequate" level of foreign reserves suggests either an implicit guarantee (a limited "tolerance of risk") or, simply, institutional inertia related to operational requirements and investment.[113]

107 See above (n 105).

108 The common abbreviation is LLR.

109 See also Rodrik "The social cost of foreign exchange reserves" (2006) 20 *International Economic Journal* 253–266.

110 See above (n 109).

111 Taxes of the Chilean-type in the 1990s are typically the "go-to" example (see also, Forbes "One cost of the Chilean capital controls: Increased financial constraints for smaller traded firms" (2007) 71 *Journal of International Economics* 294–323).

112 See above (n 105).

113 The SARB allocates its foreign-exchange reserve holdings into three functions: (1) domestic liquidity management for the "timely availability of reserves to meet commitments without incurring significant penalties"; (2) a capital preservation buffer such that "risks are controlled in a prudent manner to ensure the security of reserves"; and (3) income generation from reserve holdings (investments) that provide a "market-related total return within a framework of acceptable risk" (see SARB (n 50) 35). Somewhat more concerning is the discretionary leeway given to itself in defining what "adequate" means: "The level of foreign reserves may be described as adequate when a central bank feels that it can achieve its selected objectives." See SARB (n 50) 33.

From a national level and given the country's low savings rate and high dependence on international capital inflows, the realisation of these risks can lead to a current account reversal. Political uncertainty and a chronically weak fiscal position only raise the probability of such events. In this respect, the banking sector's share of high-quality liquid assets[114] has been rising steadily since the phasing-in of the Basel III liquidity coverage ratios, over half of which are Rand-denominated government debt securities. How exactly the SARB can integrate sovereign default (downgrade) risks, and its associated corporate spill-over risks, into the prudential framework is also unclear given the banking sector's heavy reliance on government debt securities for HQLA.

4.3 *Unintended policy consequences*

Conflicting macroeconomic policies and regulatory arbitrage

We define policy coordination failures on the extensive margin to be conflicting outcomes between macroprudential policy and monetary or fiscal policy. Policy coordination failures on the intensive margin describe regulatory arbitrage between macroprudential instruments.

Policy coordination on the extensive margin has received the most attention in the literature to date.[115] These studies either look at the interaction of specific macroprudential policy instruments, such as loan-to-value rules and capital adequacy rules, with monetary policy and fiscal policy, or on the impact of various macroprudential policy tools on the broader economy. The performance of these policy coordination exercises is typically measured by the minimisation of welfare losses. For example, the success of macroprudential policy is measured by its ability to reduce the procyclicality of the financial system.[116] Here, variables like house prices, equity prices, bank leverage and credit spreads serve as measures of financial stability, and the risk of financial instability can be related to the distance of observed bank leverage from a regulatory leverage ratio or excessive maturity mismatches between assets and liabilities. Furthermore, these regulatory requirements can be set to adjust to financial and/or business cycle fluctuations such as the credit-to-GDP ratio. Notably, monetary policy and MaPP instruments may also have positive and negative spill-over effects on each other's objectives.[117] For example, in an economic expansion, tighter monetary policy (aimed to reduce inflation) can reinforce financial system resilience by constraining credit expansion. In contrast, if bank liquidity or capital requirements become binding in a recession, this can constrain the countercyclical effectiveness of monetary policy. Regarding fiscal policy, the relationship between its stance and the activation of

[114] The common abbreviation is HQLA.

[115] Galati and Moessner (n 21); Hollander "Macroprudential policy with convertible debt" (2017) 54 *Journal of Macroeconomics* 285–305.

[116] Borio "Rediscovering the macroeconomic roots of financial stability policy: Journey, challenges and a way forward" (2011) 3 *Annual Review of Financial Economics* 17.

[117] Arslan and Upper (n 12) 32.

macroprudential instruments is much less clear. Arslan and Upper[118] suggest that MaPP can limit the ability of low-income earners and SMEs to access finance, which conflicts with redistributive (fiscal) policies. Policy-makers cannot combat these spill-over effects, but need to take into account the effect of each measure on the whole economy.

Requiring monetary policy and financial stability policy coordination under the oversight of the central bank, as with the SARB, can conflict with the credibility and independence of monetary policy. If South Africa experiences a systemic financial crisis and public perception is such that the episode is viewed as a financial stability policy failure, it is unclear to what degree the credibility of monetary policy decisions will remain unaffected. The independence of the monetary authority can then come under disrepute. There is also growing concern over consolidated (unelected) power within these institutions and the trade-off it faces with political interference.[119]

Analysis of regulatory arbitrage on the intensive margin (that is, policy coordination failures between MaPP instruments) has received much less attention in the literature. We have already touched on perverse incentives, such as risk-shifting from deposit insurance, which leads to the substitution from bank-based to non-banking intermediation. Incentive incompatibilities can further arise from profit-maximising behaviour, competition and regulations. A more subtle problem is when macroprudential instruments impact financial risk indictors without dealing with the underlying systemic risk.[120] A good example is the well-documented unintended procyclical effect of Basel II regulations on the business cycle. Here, research shows how Basel II altered its own measure of resiliency due to its risk-weighted approach to capital adequacy requirements. Once a MaPP policy instrument is activated or implemented, financial institutions tend to allocate time and resources to target that requirement, whether it be a systemic risk measure or a financial stability stress test simulation. These unintended consequences are only compounded when two or more macroprudential policy instruments become binding in the financial system.

4.4 *Dealing with the unintended consequences of policy (in)action*

Measurement, infrequent instrument activation, and rules versus discretion

The first point to note is that the SARB faces a trade-off between correctly measuring the likelihood and cost of financial distress with a sufficient lead (*ie* missing the build-up of financial imbalances) and being confident about the desired effect from taking a specific preventative action (*ie* activating an instrument(s) that is not needed or inappropriate). Most institutions in charge of financial stability measure systemic risk with historical and real-time data, across institutions and across time (see section 2 above). And most institutions measure the vulnerability

[118] See above (n 12) 38–39.
[119] Tucker (n 10).
[120] Arslan and Upper (n 12).

of the financial system to risks by simulating stress tests.[121] Financial stability stress tests provide forward-looking counterfactual scenarios to determine whether policy intervention is currently necessary. It therefore follows that measuring systemic risk and financial system vulnerabilities are sensitive to methodological approaches: any over-weighted single measure or under-weighted discrepancy can have a sizeable influence on the assessment of systemic risk (step 1 in section 3). Real-time data and quantitative methods only compound the likelihood of measurement error. Furthermore, there may be systematic biases in the underlying approach: inference errors from standardised or prescribed stress tests can lead to severe consequences when it matters most (the recent global financial crisis being a clear example). And as policy communication improves and markets internalise policy decisions more rapidly (that is, policy becomes more endogenous and operates with long and variable leads) estimating the effects of macroprudential policy becomes increasingly difficult. These realities will have a non-negligible influence on making a case for MaPP intervention (step 2) – especially if selecting and applying the relevant MaPP instrument (step 3) requires experimentation and informed discretion.[122]

In South Africa, the major banks are well above key metrics such as the leverage ratio, liquidity coverage ratio and capital requirements. The SARB therefore has had little need to use their instruments. Currently, and in line with international standards (Basel III), the SARB only uses its countercyclical capital buffer (CCB) as an indicator of the stance of MaPP.[123] But even if they decide to activate the CCB during a credit expansion (the main early-warning indicator being a high credit-to-GDP ratio) overall capital ratios may remain unchanged if banks prefer to reduce their precautionary capital holdings. This short-run ineffectiveness of the CCB on the financial cycle brings into question what level of capital requirements is appropriate for an economy like South Africa, and whether the CCB can even mitigate the financial cycle. Given this, the SARB has taken a prudent approach, and emphasises the usage of the CCB on building up resilience aimed at long-run growth stability. As such, the CCB could represent an acceptable risk tolerance, or so-called "standard of resilience".[124]

An unfortunate unintended consequence of prudence is inaction. One way to deal with inaction bias is with a rule-based approach. For example, a predetermined response of the CCB to the ratio of credit to GDP, akin to those adopted for monetary policy, works well in constrained model environments. In reality, however, it requires not only a good understanding of the transmission mechanisms, but a stable relationship between the instrument and the objective. More generally, the CCB reaction function may involve following a systematic rule or a process of "guided discretion" whereby the SARB sets its instrument (the CCB) to target the forecasts of its target variables (the financial cycle) to show how policy should

[121] SARB financial stability stress tests follow both bottom-up and top-down approaches, and involve both banks and non-banks (*ie* insurers and financial conglomerates). See Arslan and Upper (n 12) 27.
[122] SARB (n 62).
[123] SARB (n 7) 27–28.
[124] Tucker (n 10) 10.

be made to hit their objective over the medium- to long-run. The communication and interpretation of stress tests and the CCB should emphasise the expected path of a policy intervention given the available information fed into the model(s) and given the CCB reaction function. These quantitative results should provide an informative range of counterfactual paths of the economy. It is also important to emphasise that the path of policy decisions is conditional on a certain decision-making process: whether adopting a strict rule or guided discretion, it is crucial for both policy-makers and the public to not be lulled into false expectations of the central bank's ability to fine-tune the financial cycle. A rule-based approach constrains this temptation, and can effectively leverage the communication of MaPP even if its regulatory requirement is not binding.[125]

At the institutional level, structural reform that formalises clear guidelines and rules can enhance financial system resilience. The SARB has made significant headway in this regard with the drafting of a Special Resolution Bill.[126] The bill intends to establish the SARB as the sole resolution authority with clear governance guidelines and rules. The bill highlights the establishment of a uniform definition of a trigger for entry into resolution, open resolution procedures to restore and maintain critical functions of a designated resolution institution, transparency and cooperation with other jurisdictions, an industry funded DIS, and more certainty for creditors and investors. We have detailed some of the unintended consequences of a DIS scheme, and in a similar vein, a uniform trigger would likely create some distortions on the balance sheets of financial institutions. The distortions and administrative costs of these additional rules need to be carefully weighed against the benefits of guided discretion and market-based outcomes.

Transparent and credible *ex post* measurement of stress episodes can also help ensure the accountability of the SARB to its mandate, thus building credibility in markets and with the public.

Policy coordination: A "whole-economy" approach

Policy coordination failures on both the intensive and extensive margins require a prudent approach to macroprudential policy. Because it is clear that the SARB should step in given an episode of financial distress or crisis, the key question that arises is on the appropriate response to the build-up of systemic risk. As suggested above, we see little scope for the active use of the countercyclical capital buffer to mitigate the financial cycle. We advocate, instead, for a macro-financial (or "whole-economy") approach to macroprudential policy, which emphasises independent yet close coordination with other macroeconomic policies.[127] This macro-financial stability framework encompasses policy coordination with microprudential regulation and supervision (to which there is often overlap or no clear distinction), monetary policy, fiscal policy and structural policies.

[125] Broadly, the instruments of MaPP are supervision, regulation and communication. See also Svensson (n 34).

[126] National Treasury (n 97).

[127] BIS (n 10); Tucker (n 10) 5.

Given the historically conflated responsibilities of monetary policy and macroprudential policy, we highlight two important caveats for paving the way forward. The first caveat is that financial stability can be a mandate of the central bank, but it cannot be the objective of monetary policy. The purpose of this division is twofold. First, monetary policy is ill-equipped to combat financial instability and its policy objectives can conflict with the promotion of financial stability. For example, if monetary policy responds to heightened credit risk indicators by raising its policy rate, the reduction in inflation below anchored expectations will erode real incomes and raise real debt burdens. This rise in the cost of servicing debt can reduce financial stability through a rise in non-performing loans and risk-taking. Secondly, independent decision-making bodies separate accountability for achieving their respective goals. For example, this independence mitigates the spill-over of credibility erosion on both monetary policy, in the event of a financial crisis, and financial stability, in the event of a recession or temporary inflation.

The second caveat is that MaPP should be subordinate to monetary policy when conflicts arise between their objectives. The simple reason is that the practice of MaPP is still in its infancy to that of monetary policy: its mandate is difficult to measure and/or define, and the transmission mechanisms of its instruments are not well understood. For example, consider a sharp and persistent rise in inflation above the monetary authority's objective that requires an increase in the monetary policy instrument (the policy rate). This monetary policy response can reduce the yield curve on government debt (which can create a maturity mismatch on bank balance sheets) or tighten the net interest rate margin of financial institutions (which reduces effective profits). A reactionary response of MaPP to loosen financial conditions can reverse monetary policy's restraint on inflation. The net effect will result in policy ineffectiveness, undue volatility of financial and economic variables, which may result in a ratchetting effect by which the level of the policy instruments becomes distortionary for prolonged periods. MaPP therefore cannot be seen as a panacea to all financial instability woes and must, in particular, coordinate with monetary policy decision-making and be subordinate to monetary policy in achieving its target. Given the structure, size, and international integration of South Africa's financial sector, it is clear that the rules and guidelines for crisis management and resolution must be efficient and effective.

5 Conclusion

In conclusion, it seems appropriate to use the current implementation protocol adopted by the SARB in applying macroprudential policies as a baseline for evaluation. As previously discussed, the SARB uses a three-step procedure, which bears some likeness to the structure imposed in determining how to conduct monetary policy effectively. In the first step, the governing authority attempts to identify the nature of the systemic risk in question. This might be the most important, and often overlooked, consideration for the regulating body. Applying the incorrect tool could potentially exacerbate distortions and generate a type of government failure that deepens the financial imbalance in the economy. Problems along this dimension are exacerbated by the fact that measurement of financial instability and stress testing is difficult to perform. In South Africa, the development of tools

to identify sources of risk are still in their infancy, which makes it difficult for policy-makers to prepare for potential perturbations. This has prompted authorities to take more of a stilted stance to policy-making at present and thus providing the appropriate buffers against shocks.

The second step is to explore the case for macroprudential intervention once the source of risk has been identified. This ties into the central theme of this paper, which centres on the extent to which policy-makers need to intervene (given they can correctly identify the source of systemic risk). The firefighters analogy points to the fact that while intervention might equilibrate the economy in the short run, it could come at the cost of longer-run instability. Some argue that only when financial procyclicality is considered excessive or increased interconnections between systemic institutions can produce catastrophic failure should prudential authorities intervene. We argue that policy-makers should realise that their actions are inherently distortionary. In this regard, they should only intervene in key areas that they find to be of utmost importance and not micromanage each individual institution in a financial sector. In this regard, the SARB has been successful. They have thus far met the minimum criteria for the Basel III Accords while not imposing too many restrictive measures.

The final step is selecting and applying relevant macroprudential instruments to achieve the stated goal of decreasing the build-up of systemic risk. The range of tools available to policy-makers is ever expanding after the crisis; however, only a few tools have been implemented across a wide spectrum of countries. The reason for this is that while selection of the instrument is considered vital for the task at hand, misuse can lead to worse outcomes than abstaining from implementation (or using only a limited subset of tools available). South Africa has been particularly restrained in the active usage of tools to combat the potential build-up of systemic risk or any realised financial distress (which have so far been contained to smaller institutions). We regard this as a productive approach that could potentially foster growth in the medium to long run, the reason being that using a barrage of policy tools to address a singular problem can lead to conflicting results and coordination problems. Finally, we stress the importance of taking the whole economy into consideration before implementing policy intervention. This requires subordinating *active* macroprudential policy (to monetary, fiscal and other structural policies) in favour of *preventative* measures (such as the Special Resolution Bill, a well-capitalised banking sector and accountable supervision) that build financial system resilience.

Cost of credit in terms of the National Credit Act: "On the road fees", administrative fees and/or handling fees

CORLIA VAN HEERDEN*
STEFAN RENKE**

1 Introduction

In the 2004 Policy Framework that preceded the subsequent introduction of the National Credit Act 34 of 2005 (NCA) it was observed:

> "The credit market that developed over the past 40 years is inappropriate for the present and future political, economic and social context of South Africa. It is also a market that both reflects, but also reinforces, the two economies of South Africa – one economy that is modern, globally integrated and producing most of the country's wealth, the other characterised by underdevelopment and structurally disconnected from the first and the global economy. It is furthermore a market that is characterised by a lack of transparency, limited competition in the *high cost of credit* and limited consumer protection. For all these reasons a *fundamental review* of the credit market and its regulation is necessary."[1]

The 2004 Policy Framework in particular stressed the need to balance access to credit with the need to avoid consumer over-indebtedness.[2] The objectives of the new consumer credit Policy Framework was stated to be inter alia "making the *credit market* function more *cost-effectively* and competitively and promoting a fair, competitive and sustainable credit market".[3]

Insofar as cost of credit was concerned, the Policy Framework proposed the standardisation of charges across all service providers, into three categories, namely loan origination fees, monthly service fees and interest. It was envisaged that the new credit legislation would provide the Minister of Trade and Industry with powers to introduce, if necessary, limitations for all three categories of fees.[4] It was also indicated that very few consumers were aware of the total costs, including

* Professor of Mercantile Law, University of Pretoria; ABSA Chair in Banking Law in Africa.
** Associate Professor of Mercantile Law, University of Pretoria.

1 DTI Policy Framework 2004, Chapter 2 – authors' emphasis. See also Chapter 4 where it was stated that "[t]he credit market at present is a dysfunctional market that underserves the historically disadvantaged and is characterised by a lack of effective competition, inadequate transparency and the high cost of credit".

2 *Ibid* – authors' emphasis.

3 *Ibid*.

4 DTI Policy Framework 2004 par 4.9. It was envisaged that the discretion of the Minister in setting these fees would be limited and had to be based on recommendations by the National Credit Regulator pursuant to research conducted to minimise the distorting effect of finance charge regulation.

ABLU 2019 (Part ONE).indb 79 2019/10/09 9:49 AM

the fees, charges and "add-ons", of an item bought on credit and that the disclosure requirements of the existing credit legislation (the Credit Agreements Act 75 of 1980 and the Usury Act 68 of 1973) were outdated and ineffective. Consequently, the new credit policy would require credit providers to disclose costs in a standard manner with prescriptions around the cost that may be charged. Pre-contractual disclosure in the form of a compulsory written quote which would be binding on the credit provider for a minimum period would therefore also be introduced.[5]

In the Memorandum on the National Credit Bill it was subsequently stated that the Bill regulates the "cost of credit" in the following manner:[6]

> "*(a)* Allowing for the Minister to establish an interest cap and other cost controls, and prohibiting interest or other costs in excess of those prescribed rates.
>
> *(b)* Prohibiting any costs other than the principal sum borrowed, interest, an initiation fee (which may not be charged unless a credit agreement results from the application), periodic or transaction-based service fees, insurance premiums for credit insurance, and delivery, installation and like charges, and collection costs. Each of the categories of fees, premiums and charges is subject to regulatory maximums or standards.
>
> *(c)* Prohibition of surcharges of amounts spent on account of the consumer for insurance and incidental costs.
>
> *(d)* Modification of the *in duplum* rule, as set out in section 103(5)."

In addition, the Memorandum indicated that the Bill required advance disclosure of all costs, and that the consumer has the right to arrange incidental matters directly, rather than pay the credit provider to do so, and may choose to arrange his or her own insurance policies.[7]

Accordingly, when the National Credit Act 34 of 2005 came into full, effective operation on 1 June 2007 it comprehensively transformed the regulation of the South African credit market, including the aspect of cost of credit.

The purpose of this contribution is to consider which costs the NCA allows to be charged to the consumer under an instalment agreement in respect of the acquisition of a motor vehicle. The focus will be on so-called "on the road fees" and "administrative and/or handling fees" that are charged by motor dealerships. It will be accepted that the agreements concerned will either be intermediate or large credit agreements, meaning that the principal debt under the said agreement is in excess of R15 000.[8]

2 "On the road fees", administrative and/or handling fees

The concept "on the road fees" is not defined in the NCA, and neither is the concept "administrative and/or handling" fees. In fact, the NCA does not mention these

[5] DTI Policy Framework 2004 par 5.1, 5.4 and 5.5. It was envisaged that the quote would allow consumers time and provide them with information needed so that they could shop around. It would also help consumers to make better choices between cash and credit purchases and between different credit providers.

[6] Memorandum on the National Credit Bill 2005 par 2.11 – authors' emphasis.

[7] *Ibid*.

[8] Section 9(3)*(b)* and 9(4)*(b)*.

types of fees at all. In practice these fees are charged by a motor dealership when a vehicle is bought. If the consumer is in the fortunate position to be able to afford to pay cash for the vehicle, he or she will pay these fees directly to the motor dealership in cash. Where, however, as is the case with most consumers who wish to acquire vehicles, he or she is not in a position to enter into a cash transaction with the dealer, the consumer will have the transaction financed by a credit provider. In such instance the dealership will then invoice the credit provider who will then, on behalf of the consumer, effect payment of the purchase price of the vehicle and also the "on the road fee", administrative and/or handling fee charged by the dealership.[9]

There is no standard document in the public domain that explains the concept of "on the road fees", nor are these fees dealt with or prescribed by the Motor Industry Association of South Africa (MIOSA).[10] As explained by Pretorius, typical "on the road fees" involve the following items:[11]

- pre-delivery inspection/safety check;
- certificate of roadworthiness;
- delivery fuel;
- initial fuel;
- hire purchase information clearance (HPI clearance);
- administration fee;
- FSB fees;[12] and
- cleaning or valet costs.

From the above it appears that "on the road fees" are actually a mix of expenses and fees. These expenses and fees will not necessarily all apply in the case of a brand-new vehicle, which may for instance not need a valet before delivery, but generally all apply in the event of pre-owned cars. These expenses and fees represent costs that are actually incurred, and someone has to pay for it. That someone is the consumer who wants to become the owner of the car, be it by way of a cash transaction or eventually after he or she has paid all his or her obligations in terms of a credit agreement whereby the purchase of the car is financed.

[9] See *Volkswagen Financial Services v National Consumer Commission* NCT/94937/2017/56(1) par 9 where this was explained by the Tribunal.

[10] See, however, "The road ahead" *Miosa Newsletter* Issue 3 August 2018 at 4 (https://www.miosa.co.za>register.newsdocs.pdf (accessed 6 August 2019)) where a brief and unclear reference is made to the National Credit Regulator's investigation into "on the road fees".

[11] Pretorius "'On the road fees': What does it really mean?" (https://www.autotrader.co.za/cars/news-and-advice/automotive-news/on-the-road-fees-what-does-it-really-mean/3040 (accessed 6 August 2019)).

[12] Given that the Financial Services Board (FSB) has been replaced by the Financial Sector Conduct Authority (FSCA) in accordance with the Financial Sector Regulation Act 9 of 2017, introducing the South African Twin Peaks model of financial regulation, these fees will be referred to as FSCA fees.

ABLU 2019 (Part ONE).indb 81 2019/10/09 9:49 AM

Pretorius further explains that the pre-delivery inspection (safety check) is enforced by the franchise agreements between the dealership owners and manufacturers requiring dealers to inspect every car they sell to ensure that it conforms to the basic roadworthiness criteria.[13] He points out that every time a car's ownership changes such change has to be accompanied by a certificate of roadworthiness, involving an outside testing station.[14] The delivery fuel is the fuel in the car's tank when it gets delivered to the consumer. This is different from the initial fuel, which refers to the fuel in the tank that enables the consumer to test drive the vehicle and that allows the car to be taken off-site for other purposes as well. The HPI clearance is necessary because before a dealership can sell a vehicle to a consumer, it must first ascertain that the vehicle "is free of outstanding debt and that it hasn't been involved in a serious accident". As regards the administration costs Pretorius remarks:

> "This is where the OTR cost saga gets juicy. Licensing and registration alone costs from R500 and up, depending on the vehicle – big SUVs and bakkies may even reach double that amount. If new license plates are needed, there's an extra R350+ out of the buyer's pocket, and remember that the staff or outside contractors tasked with the licensing process also do not work for free …"

Pretorius explains that FSB fees refer to fees payable to the FSB (now FSCA)[15] in relation to inhouse F&I (finance and insurance) service licences. The necessity for cleaning or valet services of course speaks for itself, especially where the car concerned was pre-owned.

Administrative and/or handling fees appear to refer to some of the fees included above, such as the fees for obtaining roadworthy certificates and preparing the dealer's invoice. "On the road", administrative and/or handling fees must further be distinguished from dealer "add-ons". For dealer "add-ons" such as, for example, aftermarket wheels or fitting a tracker system to the vehicle, permission of the consumer is obtained – thus he or she can actually say no to these "add-ons".

3 The National Credit Act and cost of credit

Cost of credit is dealt with in Part C of Chapter 5 of the NCA, which is titled "Consumer's liability, interest, charges and fees" and contains section 100

[13] Pretorius (n 11) indicates that in the case of pre-owned cars this duty goes even deeper because those cars have to be inspected by relevant agents and certified to conform to the manufacturers' specifications. He points out that this expense increases if a car is sold on through an outside dealership, especially if the car in question is still under warranty or if its maintenance or service plan is still active.

[14] Pretorius (n 11) further points out that the acquisition of a certificate of roadworthiness applies to private sales as well (where a dealership is not involved) except that in such instance the responsibility for obtaining a certificate of roadworthiness rests with either the seller or buyer. However, where a car is sold by a dealership the dealership is obliged to take care of the registration, mainly to ensure that the vehicle is removed from its system and to safeguard the dealership against traffic fines and infringements.

[15] In terms of the Financial Sector Regulation Act 9 of 2017 the FSB ceased to exist and the FSCA, established in terms of s 56, is the new market conduct regulator.

(prohibited charges), section 101 (cost of credit), section 102 (fees and charges), section 103 (interest), section 104 (changes to interest, credit fees or charges), section 105 (maximum rates of interest, fees and charges) and section 106 (credit insurance). Part C of Chapter 5 has to be read with regulations 39 to 48.[16]

For purposes of this discussion the focus will be on sections 100, 101 and 102.

3.1 *Section 100: Prohibited charges*

In terms of section 100(1),

> "a credit provider *must not charge* an amount to, or *impose a monetary liability on*, the consumer in respect of—
>
> *(a)* a credit fee or charge prohibited by the NCA;
>
> *(b)* an amount of a fee or charge exceeding the amount that may be charged consistent with the NCA;
>
> *(c)* an interest charge under a credit agreement exceeding the amount that may be charged consistent with the NCA; or
>
> *(d)* any fee, charge, commission, expense or other amount payable by the credit provider *to any third party* in respect of a credit agreement, *except as contemplated in section 102 or elsewhere in the NCA*."[17]

Failure to comply with the provisions of section 100 is viewed in a very serious light and constitutes an offence.[18]

3.2 *Section 101: Cost of credit*

In order to determine the nature and scope of the cost of credit that is allowed by the Act regard must be had to section 101 (read with section 102 where relevant for purposes of determining the extent of the principal debt).

In terms of section 101(1):

> "A credit agreement must not require payment by the consumer of any money or other consideration, except—
>
> *(a)* the *principal debt*, being the *amount deferred* in terms of the agreement, *plus the value of any item contemplated in section 102*;
>
> *(b)* an *initiation fee*, which—
>
> (i) may not exceed the prescribed amount relative to the principal debt; and
>
> (ii) must not be applied unless the application results in the establishment of a credit agreement with that consumer;
>
> *(c)* a *service fee*, which—
>
> (i) in the case of a credit facility, may be payable monthly, annually, on a per transaction basis or on a combination of periodic and transaction basis; or

[16] National Credit Regulations, 2006 (GN R489 in *GG* 28864 of 31 May 2006) ("Regulations").

[17] Authors' emphasis.

[18] Section 100(3) as added by s 29 of the National Credit Amendment Act 9 of 2014. Section 161 provides for penalties for offences and stipulates that, except for offences relating to a contravention of s 160(1), a person convicted of an offence in terms of the NCA is subject to a fine or imprisonment for a period not exceeding 12 months, or to both a fine and imprisonment.

(ii) in any other case, may be payable monthly or annually; and

(iii) must not exceed the prescribed amount relative to the principal debt;

(d) *interest*, which—

(i) must be expressed in percentage terms as an annual rate calculated in the prescribed manner; and

(ii) must not exceed the applicable maximum prescribed rate determined in terms of section 105;

(e) cost of any *credit insurance* provided in accordance with section 106;

(f) *default administration charges*, which—

(i) may not exceed the prescribed maximum for the category of credit agreement concerned; and

(ii) may be imposed only if the consumer has defaulted on a payment obligation under the credit agreement, and only to the extent permitted by Part C of Chapter 6; and

(g) *collection costs*, which may not exceed the prescribed maximum for the category of credit agreement concerned and may be imposed only to the extent permitted by Part C of Chapter 6."[19]

Section 101 thus prohibits a credit provider from charging amounts as cost of credit other than the costs stated in section 101(1)*(b)* to *(g)*. Although the principal debt is mentioned in sub-section *(a)* of this section, which is titled "cost of credit", the principal debt cannot strictly be regarded as cost of credit but rather the amount on which costs are subsequently levied by the credit provider.

Where a credit provider charges costs other than that allowed by section 101, he or she thus contravenes the Act by engaging in prohibited conduct. It is, however, to be noted that section 101 merely prohibits the charging of these fees and it is not stated that charging such fees constitutes unlawful conduct as envisaged in section 90. This can be interpreted to mean that if these costs are charged, the credit provider can be sanctioned for engaging in prohibited conduct, for example, by means of an administrative fine in terms of sections 150 and 151, but the credit agreement itself will not be invalid.[20]

3.3 *Concepts relevant to the interpretation of section 101*

Various concepts appear in sections 100 and 101 that require further clarification. Only those concepts relevant to this contribution will be dealt with here. Given that section 100 prohibits a credit provider to "charge" or "impose a monetary liability on" a consumer who enters into a credit agreement in the circumstances set out in section 100(1)*(a)* to *(d)* the meaning of the word "charge" as used in section 100 is significant. The same can be said of the phrase "impose a monetary liability on".

[19] Authors' emphasis.

[20] Section 150 gives the Tribunal the power to declare conduct prohibited and to interdict such conduct and/or impose an administrative fine on the credit provider who engages in such conduct. Section 151 deals with the power of the Tribunal to impose administrative fines.

The word "charge" is not defined in the NCA, nor is any explanation thereof given in any of the other sections of the Act. In accordance with the well-established principles of statutory interpretation, "charge" therefore has to be afforded its ordinary grammatical meaning, bearing in mind that it is used in the context of credit granting. The Oxford Dictionary defines the verb "charge" inter alia as to "demand (an amount) as a price for a service rendered or goods supplied".[21] The word "impose" is also not defined in the NCA, nor is the concept "monetary liability". No explanation is given of the phrase "impose a monetary obligation on". Again, these words must be afforded their ordinary grammatical interpretation: "impose" inter alia means "to officially force a rule … to be obeyed".[22] "Monetary" means "relating to money or currency",[23] and "liability" means "the state of being legally responsible for something"[24] and is used in relation to debts. Thus, monetary liability would refer to liability for a debt sounding in money. The gist of section 100 is thus that it prohibits a credit provider from demanding that a consumer pay certain amounts as specified in the section.

The concept "principal debt" is defined in section 1 of the NCA as "the amount calculated in section 101(1)*(a)*". Section 101(1)*(a)* refers to the principal debt as the "amount deferred" *plus* section 102 costs. Interestingly, the NCA contains no definition in section 1 of the concept "amount deferred/deferred amount" or of the concept "defer" although the concept "deferred amount" is defined in regulation 39(1), as indicated in more detail below. The ordinary grammatical meaning of "defer" is to "delay something until a later time".[25] It is trite, however, that credit involves the deferral of payment of a debt and that certain costs are usually levied as "compensation" for such deferral. In the context of the principal debt as mentioned in section 101(1)*(a)*, it is the payment of the said principal debt that is being deferred when credit is granted.

Chapter 5 of the Regulations to the National Credit Act deals with interest and fees and contains certain definitions as set out in regulation 39. As such, regulation 39 provides that "deferred amount" means

> "any amount payable in terms of a credit agreement the payment of which is deferred and upon which interest is calculated, or any fee, charge or increased price is payable by reason of the deferment, and
>
> *(a)* the deferred amount includes
>
> (i) any obligation of the consumer that is deferred as per section 8(3) and 8(4) of the Act;
>
> …
>
> (iii) the amounts referred to in section 101(1)*(b)* to 101 (1)*(g)* inclusive;
>
> (iv) the amounts referred to in section 102(1)*(b)* to section 102(11)*(f)* …"[26]

[21] Available at https://www.oxforddictionaries.com (accessed 29 July 2019).
[22] Available at https://www.dictionary.cambridge.org (accessed 29 July 2019).
[23] Available at https://www.oxforddictionaries.com (accessed 29 July 2019).
[24] Available at https://www.dictionary.cambridge.org (accessed 29 July 2019).
[25] Available at https://www.oxforddictionaries.com (accessed 29 July 2019).
[26] The reference to s 102(11)*(f)* is wrong and should be to s 102(1)*(f)*.

Thus, practically interpreted, the "deferred amount" in regulation 39 includes:

- any obligation of the consumer that is deferred in terms of a credit facility or a credit transaction;
- an initiation fee, a service fee, interest, cost of any credit insurance, default administration charges, and collection costs; and
- cost of an extended warranty agreement; delivery, installation and initial fuelling charges; connection fees, levies or charges; taxes, license or registration fees; or credit insurance premiums (being the costs mentioned in section 102 as dealt with in more detail below).

It thus appears that the definition of "deferred amount" in regulation 39 is broader than the concept of "deferred amount" as envisaged by section 101 as the deferred amount in the latter section is used to refer to the part of the principal debt before any section 102 costs are added, whereas for purposes of regulation 39 it appears to include the whole caboodle, *ie* the principal debt (deferred amount plus section 102 fees) as well as section 101(1)*(a)* to *(g)* costs. There thus appears to be a contradiction between the term "deferred amount" as used in the Act and as used in the Regulations in that the Act uses the concept of "deferred amount" to refer to an amount before any costs are added whereas the Regulations refer to an amount that includes costs. Be that as it may, it is also trite that the Regulations are subordinate legislation and cannot be used to interpret provisions of the enabling Act in terms whereof they were issued if the Act itself lacks certain provisions such as a definition of the concept "deferred amount".[27] The concept of "deferred amount" as it is used in regulation 39 surfaces again where the calculation of interest is explained in regulation 40. It also appears earlier in the Regulations, namely in regulation 31, which deals with the requirements for intermediate and large credit agreements.

It is further submitted that the use of the word "except" in section 101(1) indicates that the legislature intended the costs mentioned in section 101(1) to constitute a "closed list", *ie* not allowing for the addition of any other costs of credit *except* those mentioned in section 101 read with section 102 (which is included in section 101(1) by virtue of the reference thereto in section 101(1)*(a)*).

Finally, it also needs to be pointed out that the costs mentioned in section 101(1) *(b)* to *(f)* are all typical costs that are *charged by the credit provider in relation to the extension of credit to the consumer*. This is clear from the meaning that the Act attributes to each of these concepts, namely:

- "Initiation fee" means "a fee in respect of costs of initiating a credit agreement, and (i) charged to the consumer by the credit provider; or (ii) paid to the credit provider by the consumer upon entering into the credit agreement".[28] Regulation 42, Table B sets out the maximum initiation fee that may be charged in respect of mortgage agreements, credit facilities, unsecured credit transactions, developmental credit agreements, short-term credit transactions, other credit agreements and incidental credit agreements.

27 *Rossouw and Another v FirstRand Bank* 2010 (6) SA 439 (SCA).
28 Section 1.

- "Service fee" means "a fee that may be charged periodically by a credit provider in connection with the routine administration cost of maintaining a credit agreement".[29]
- "Interest" is not defined in the Act or in the Regulations. It is submitted that interest has a narrower meaning than "finance charges", which is also not defined in the Act but which, it is submitted, can be equated with the broader cumulative concept "cost of credit". Calculation of interest is dealt with in regulation 40. Regulation 42, Table A further stipulates the maximum prescribed interest rates for mortgage agreements, credit facilities, unsecured credit transactions, developmental credit agreements, short-term transactions, other credit agreements and incidental credit agreements.
- "Cost of credit insurance" is provided for extensively in section 106 of the NCA and the credit insurance guidelines and will not be dealt with in any further detail here.
- "Default administration charge" is defined to mean a "charge that may be imposed by a credit provider to cover administration costs incurred as a result of a consumer on an obligation under a credit agreement".[30] In terms of section 101*(f)*(i) and (ii) the costs of default administration charges may not exceed the prescribed maximum for the category of credit agreement concerned. Also, default administration charges may be imposed only if the consumer has defaulted on a payment obligation under the credit agreement, and only to the extent permitted by Part C of Chapter 6. Regulation 46 further stipulates:

  > "The credit provider may require payment by the consumer of default administration charges in respect of each letter necessarily written in terms of Part C of Chapter 6 of the Act. Such payment may not exceed the amount payable in respect of a registered letter of demand in an undefended action in terms of the Magistrates Court Act, 1944 in addition to any reasonable and necessary expenses incurred to deliver such letter."

- The Act defines "collection costs" to mean "an amount that may be charged by a credit provider in respect of enforcement of a consumer's monetary obligations under a credit agreement, but does not include a default administration charge".[31] Section 101(1)*(g)* allows collection costs to be charged as "costs of credit" but stipulates that it may not exceed the prescribed maximum for the category of credit agreement concerned and may be imposed only to the extent permitted by Part C of Chapter 6.

The gist is, however, that all these costs mentioned in section 101(1)*(b)* to *(g)* are costs that are *directly related to the granting of credit* and that are *charged specifically by the credit provider* in exchange for deferring payment of the principal

29 *Ibid*.
30 *Ibid*.
31 *Ibid*.

debt that is financed in terms of a credit agreement. In particular, it is these fees that attract the application of the NCA to the transaction.[32]

3.4 *Section 102: Fees or charges*

Section 102 deals with fees or charges that may be included *in* the principal debt deferred under the agreement (as also mentioned in section 101(1)*(a)* above). Notably, the heading of section 102 is merely "fees or charges" and not "cost of credit". The fact that these fees and charges are set out in a separate section, points to the fact that they are fees and charges that can be *included in the principal debt* (*ie* added to the "deferred amount" as provided for in section 101(1)*(a)*) – on which cost of credit is then charged or levied as set out in section 101(1)*(b)* to *(g)*.

Section 102(1) provides:

> "If a credit agreement is an instalment agreement, a mortgage agreement, a secured loan or a lease, the credit provider may *include in the principal debt deferred* under the agreement any of the following items to the extent that they are applicable in respect of any goods that are the subject of the agreement—
>
> *(a)* an initiation fee as contemplated in section 101(1)*(b)*, if the consumer has been offered and declined the option of paying that fee separately;
>
> *(b)* the cost of an extended warranty agreement;
>
> *(c)* delivery, installation and initial fuelling charges;
>
> *(d)* connection fees, levies or charges;
>
> *(e)* taxes, licence or registration fees; or
>
> *(f)* subject to section 106, the premiums of any credit insurance payable in respect of that credit agreement."

Section 102(2)*(a)* to *(c)* further stipulates that a credit provider must not charge an amount in terms of section 102(1) unless the consumer chooses to have the credit provider act as the consumer's "agent" in arranging for the service concerned; require the consumer to appoint the credit provider as the consumer's agent for the purpose of arranging any service mentioned in section 102(1); or charge the consumer an amount under section 102(1) in excess of the actual amount payable by the credit provider for the service, as determined after taking into account any discount or other rebate or other applicable allowance received or receivable by the credit provider, or the fair market value of a service contemplated in section 102(1), if the credit provider delivers that service directly without paying a charge to a third party.

Section 102 deals with amounts that the credit provider may include in the principal debt, in other words, amounts that may be added to the "deferred amount" and on which the credit provider can charge the cost of credit as envisaged in

[32] If no interest or fee or charge is levied the mere fact that an amount is deferred will generally not attract the application of the NCA. See the definitions of the various types of credit agreements in s 8(3) and s 8(4) read with the various applicable definitions in s 1 of the Act. Although the definition of "mortgage agreement" does not specifically mention the charging of interest or other fees, it is trite that such agreements also entail the charging of interest.

section 101(1)*(b)* to *(g)*. It appears that the legislature wanted to give the credit provider the option of including these amounts as part of the principal debt but that it was not the intention of the legislature that these fees will always have to be ("must" be) included in the principal debt. It would thus appear that where a consumer prefers paying these fees directly to the dealer, the credit provider will not be able to insist that these fees must be included in the principal debt.

Given the manner in which section 102(1) is phrased, it appears that the items mentioned therein also constitute a "closed list". Section 102 creates the impression that these costs (save for the initiation fee) are generally not costs that are charged by the credit provider specifically as *quid pro quo* for extending credit as it does not pertain to services that the credit provider provides himself or herself, hence the requirement that the credit provider has to be appointed as the "agent" of the consumer in procuring these services.

4 Where do "on the road fees" and administrative and/or handling charges fit in?

It may be asked whether, given that it is not specifically mentioned by name in section 101 or 102, "on the road" and administrative and/or handling fees may be charged to consumers at all?

One may possibly argue that section 102(1)*(c)*, which refers to delivery fees and initial fuelling charges, as well as section 102(1)*(e)*, which refers to taxes, licence or registration fees, can be interpreted broadly to encompass most of the "on the road fees" mentioned above in paragraph 2. As such, a broad interpretation of the concept "delivery fee" can arguably include all fees that are necessary to enable delivery of the vehicle concerned, namely the pre-delivery inspection, the certificate of roadworthiness, the delivery fuel and the cleaning and valet of the car (which is a necessity if the vehicle was pre-owned). "Taxes" broadly construed may refer to costs imposed for the HPI clearance as well as FSB fees. As section 102(1)*(c)* provides for charging of "initial fuelling charges" it appears that it will cover the initial fuelling charges that are levied by the motor dealer as part of the "on the road fee". A broad interpretation of section 102 may thus allow these fees – some expressly and others at least by implication.

If, however, this broad interpretation of section 102 is not upheld, one would have to look "elsewhere in the Act" as stated in section 100(1)*(d)* to determine whether "on the road", administrative and/or handling fees are allowed. In the latter context, one then has to determine whether these fees are either expressly or by implication allowed by other sections of the Act (excluding section 102).

As mentioned, there are no sections in the Act (apart from section 102 broadly interpreted) that expressly allow these fees. It may consequently be asked whether they are allowed by implication in the sense that they are not specifically or by implication prohibited or declared unlawful by other sections in the NCA. To determine this, regard must be had to section 90 (unlawful provisions of credit agreements); section 91 (prohibition of unlawful provisions in credit agreements and supplementary agreements); section 92 (pre-agreement disclosure) read with regulation 29 (pre-agreement statements and quotation for intermediate and large

agreements); and section 93 (form of credit agreements) read with regulation 31 (requirements for intermediate or large agreements).

Section 90 of the NCA states that a credit agreement must not contain an unlawful provision. A provision in a credit agreement is unlawful if its general purpose or effect is to defeat the purposes or policies of the NCA, deceive the consumer, or subject the consumer to fraudulent conduct.[33] A provision in a credit agreement is also unlawful if it directly or indirectly purports to waive or deprive a consumer of a right set out in the NCA, avoids a credit provider's obligation or duty in terms of the Act, sets aside or overrides the effect of any provision of the Act, authorises the credit provider to do anything that is unlawful in terms of the NCA, or fails to do anything that is required in terms of the Act.[34] Having regard to the rest of the provisions of section 90, *ie* section 90(2)*(c)* to 90(2)*(o)*, it appears that none of these provisions specifically states that charging an "on the road fee", administrative and/or handling fees (either by using these collective terms or by mentioning specific fees that would qualify as "on the road fees", administrative and/or handling fees) is unlawful. Applying the more general provisions of section 90(1) and (2) to the issue in question, it is submitted that it cannot be said that charging or levying these fees, that have *de facto* been incurred by the dealer on behalf of the consumer to ensure the proper and valid delivery of the vehicle to the consumer, defeats the purpose or policies of the Act or deceives the consumer or subjects him or her to fraudulent conduct. In the same vein it can also not be said that these fees waive or deprive the consumer of a right as set out in the NCA or that it avoids the credit provider's obligation or duty in terms of the Act or authorises him or her to do something that is unlawful in terms of the Act or to fail to do anything that is required in terms of the NCA.

Section 91 further deals with unlawful provisions in credit agreements and supplementary agreements. It bars a credit provider from directly or indirectly, by false pretences or with the intent to defraud, offer, require or induce a consumer to enter into or sign a credit agreement that contains an unlawful provision as contemplated in section 90. As it has been pointed out that section 90 does not specifically mark "on the road fees", administrative and/or handling fees charged by the motor dealer as unlawful, it is submitted that section 91 would find no application even where these fees are reflected in a supplementary agreement.

Regard may be had to section 92, which deals with pre-agreement disclosure. Section 92 does not specifically refer to "on the road", administrative and/or handling fees. It states that a credit provider must not enter into an intermediate or large credit agreement unless the credit provider has given the consumer a pre-agreement statement in the form of the proposed agreement, or in another form addressing all matters required in terms of section 93.[35] In addition, prior to entry into such an agreement, the credit provider must give the consumer a quotation in the prescribed form, setting out *the principal debt*, *the proposed distribution of that amount*, the interest rate and other credit costs, the total cost of the proposed

[33] Section 91(2)*(a)–(c)*.
[34] Section 90(2).
[35] Section 92(2)*(a)*.

agreement, and the basis of any costs that may be assessed under section 121(3) if the consumer rescinds the contract.[36]

Section 92 has to be read with regulation 29, which also does not specifically mention "on the road", administrative and/or handling fees. Regulation 29 inter alia provides that the information required to be disclosed in the quotation is:[37]

> "(i) principal debt;
> (ii) proposed distribution of principal debt with reference to items listed in section 102(1*)(b)* to *(f)* of the Act *and specify any other*;
> (iii) other ongoing credit costs;
> (iv) service fee and whether it is paid monthly, annually or on any other basis as prescribed in section 101(1)*(c)* of the Act;
> (v) initiation fee;
> (vi) rand value of interest;
> (vii) residual or final amount payable (if any);
> (viii) total cost of the proposed agreement;
> (ix) annual interest rate;
> (x) … the basis for any costs payable under section 121(3)*(b)*(i) of the Act, if applicable;
> (xi) … the reasonable rental to be charged in terms of section 121(3)(b)(ii) of the Act, if applicable;
> (xii) the number of instalments to be paid;
> (xiii) instalment amount."

Here, again, we seem to have an example of the Act and Regulations not being aligned: what are the words "and specify any other" and "other ongoing credit costs" doing in regulation 29 if section 102 is to be regarded as a closed list of costs that may be added to the principal debt?

The prescribed form of the quotation is captured in Form 20.1. Under Part A (amount advanced) the credit provider must reflect the following items:

- credit advanced or value of goods or services provided on credit;
- initiation fees (if the consumer declined the offer to pay these fees separately);
- total of additional charges (as reflected in Part E of the quotation);
- minus deposit (if any).[38]

This would then add up to the "total amount deferred per credit agreement" which has to be reflected at the end of Part A.

Part B (instalment payable) has to reflect:

- the instalment in respect of the amount deferred;
- monthly service fee monthly credit life insurance premium; and
- number of instalments and frequency.

[36] Section 92(2)*(b)*. Section 121 provides the consumer with a cooling-off right in certain instances.

[37] Reg 29(1)*(d)*.

[38] The payment of a deposit is not required by the NCA as was previously the case under the Credit Agreements Act 75 of 1980.

At the end of Part B the total instalment then has to be reflected.

Part C (total cost and interest rate) has to reflect:

- total amount deferred per credit agreement;
- total interest, fees and credit life insurance (although not expressly stated, these appear to be the amounts mentioned in section 101(1)*(b)* to *(g)*);
- total amount repayable (this is stated to be the total of all instalments excluding optional insurance); and
- annual interest rate.

Part D (optional items) makes provision for two columns, namely *optional items that will be added to the instalment* (column one) and *other optional items* (column two).[39] Column one should reflect (obviously only if applicable) the additional monthly premium for optional insurance and a description of such optional insurance. Column two, on the other hand, is blank and appears to suggest that the credit provider may insert certain optional items there that are not specifically referred to in the Act.

Part E (additional charges added to credit agreement) requires the credit provider to reflect additional charges as per section 102(1)*(b)* to *(f)* that will be added to the amount of credit (the items applicable and the amount of each item must be listed). The total of these charges must then be calculated and reflected in Part E and, as indicated above, this total also gets reflected under Part A.

Part F (security provided) requires a description of the security concerned and the conditions under which possession would occur. Part G (repayment arrangements) requires the quotation to reflect information regarding the frequency of payments, including the method of payment, date of the first payment and date of the last payment. Part H (further information on rights and obligations) requires the credit provider to

> "add further information on material aspects of the rights and obligations of the consumer and credit provider in respect of the proposed credit agreement, as required: Where a transaction fee is charged, indicate 'transaction fee' in service fee above, and describe fees and basis for levying such fees in this section; Include further disclosure required by legislation in respect of any item above, where applicable. Consider in particular disclosure requirements of section 106 and 121(3)."

Part I (further information on features of credit product) requires the credit provider to reflect further information on material features or attributes of the credit products or proposed credit agreement, as required.

From Form 20.2 it thus appears that the fees or charges mentioned in section 102 have to be reflected in Part E and the total of these charges also have to be reflected in Part A, as one of the amounts added up to calculate the "total amount deferred per credit agreement". Again, one must ask what the "other optional items" under Part D refer to if neither section 101 nor section 102 allows for any costs, fees or charges other than those listed in the aforesaid sections.

[39] Authors' emphasis.

Section 93 deals with the form of credit agreements. It states that intermediate and large agreements must be in the prescribed form, if any; alternatively it may in in a form determined by the credit provider and must comply with any prescribed requirements for the category or type of credit agreement concerned. No form is, however, prescribed in the Regulations for intermediate and large agreements and these types of agreements therefore have to meet the requirements of regulation 31. Section 93 does not mention "on the road", administrative and/or handling fees. It sheds no light on the admissibility of the aforesaid fees but inter alia requires the contents of intermediate and large credit agreements to comply with regulation 31. Regulation 31(2) contains a long list of information that must be contained in intermediate or large credit agreements and inter alia states the following in respect of the cost of credit:

> "*(c)* Cost of credit reflecting the following:
>
> (i) The amount of the principal debt, including the amount deferred in terms of the credit agreement as well as the nature and amount of the following fees and charges where they have been included in the principal debt in terms of the agreement:
>
> *(aa)* the cost of an extended warranty agreement;
>
> *(bb)* delivery, installation and initial fuelling charges, limited to the actual cost of these items;
>
> *(cc)* connection fees, levies or charges;
>
> *(dd)* taxes, licence or registration fees;
>
> (ii) If the amount deferred in terms of the credit agreement is not ascertainable, the maximum amount deferrable;
>
> (iii) The proposed distribution of the principal debt and to whom each amount is to be paid …"

These provisions basically mirror aspects of sections 101(1)*(a)* and 102, except that it appears that the words "deferred amount" are ascribed a broader meaning than in section 101(1)*(a)* of the Act and also include the costs mentioned in section 101(1)*(b)* to *(g)* (*ie* as contemplated in the broad definition of "deferred amount" in regulation 39). Regulation 31(2)*(c)*(iii) requires that the proposed distribution of the principal debt be reflected in the credit agreement as well as to whom each amount is paid. Regulation 31(2)*(c)*(vi) to (xxii) deals with the items mentioned in section 101(1)*(b)* to *(g)*: initiation fees, service fees, credit life insurance costs, default administration charges and collection costs. Interestingly, unlike regulation 29, no reference is made in regulation 31 to "other costs or ongoing costs".

The gist here is that the NCA nowhere explicitly prohibits the charging of "on the road", administration and/or handling fees or labels the charging of such fees as unlawful. If the argument that these fees are chargeable under section 102 is not accepted on the basis that section 102 does not cater for these items, then what inference may one draw from the fact that these fees are not mentioned in the Act at all? As explained below, case law does not seem to yield an appropriate answer.

5 Case law before *Volkswagen*

The fees and charges that can be charged in respect of a credit agreement in terms of section 102 came under the spotlight in 2017 in the decisions of the National Consumer Tribunal in the matters of *National Credit Regulator v Edcon Holdings Ltd*[40] and *National Credit Regulator v Lewis Stores (Pty) Ltd.*[41] In *Edcon* the Tribunal considered whether the Act allows a credit agreement to contain any fee or charge other than that permitted by the Act. Edcon submitted that the "club fees" it charged were linked to a purchase consideration of a product with benefits to consumers. The Tribunal, however, found that, apart from the fees or charges mentioned in section 102, *the Act does not allow for any other fees, charges or costs to be reflected in the credit agreement or credit agreement documents, irrespective of the nature of the charge, fee or cost*. It consequently held that any "purchase" of any product must be an *entirely separate transaction which does not in any way form part of the credit agreement* or credit facility application process.[42]

In particular the following remarks by the Tribunal are pertinent:[43]

> "The NCA goes to great lengths to describe the fees and charges which are permitted to form part of a credit agreement. The intention of the legislature is clearly to ensure that consumers are not misled in any way and know exactly the fees, charges and costs associated with the credit agreement. These specific costs, fees and charges are described in section 101 of the NCA. *There is no section of the NCA which provides for a club fee or anything similar. Any fee appearing on a credit agreement which is not a fee or charge as described in the NCA would therefore not be permitted.* …
>
> *Although section 101 of the NCA contains the word 'require' this does not necessarily allow a credit provider to add a charge or fee as long as the consumer is provided with an option to refuse the fee or payment.* It is irrelevant whether or not the fee is presented as a tick box option, can be cancelled, is optional, is not amortized or whether there is a clause specifically excluding the fee from the credit agreement. Section 101 seeks to prescribe the fees and charges that may appear in the credit agreement. *The NCA does not allow for or provide for any exceptions to the fees or charges which can be levied.*"

Subsequent to *Edcon*, in the *Lewis* matter, the Tribunal again had the opportunity to consider club fees charged by a credit provider.[44] In this decision, consumers entered into separate agreements each time they wished to join the relevant club. The NCR argued that the words "cost of credit", although not defined in the Act, require a contextual reading, which, it is submitted, is the correct view. The Tribunal in *Lewis* eventually held that the Act does not prevent a credit provider

[40] [2017] ZANCT 58 (24 April 2017).

[41] [2017] ZANCT 78 (25 May 2017).

[42] Authors' emphasis. Van Zyl in Scholtz et al *Guide to the National Credit Act* (2014 *et seq*) par 10.7 remarks that it appears that the Tribunal accepted that a credit provider may sell products separately from any credit agreement it enters into, as long as the processes and agreements are completely separate. She points out that the Tribunal did not consider whether a separate agreement would constitute a prohibited supplementary agreement as contemplated in s 91.

[43] Par 41 and 42 of the *Edcon* case (n 40) – authors' emphasis.

[44] *Lewis Stores* case (n 41). This matter also dealt with fees allegedly charged by Lewis for the cost of an extended warranty in circumstances where the original manufacturer's warranty had not yet expired.

from offering the services of a club to consumers provided that these services are not part of the cost of credit.

The Tribunal's judgment in the *Edcon* matter was taken on appeal to the High Court where the judgment was reversed.[45] Two observations by the High Court are of note. First, the High Court stated that in the context of section 101, the word "require" can only be interpreted to demand from a consumer who applies for credit, or to impose an obligation on such consumer, to pay for something which is not permitted in terms of the section. Secondly, it stated that the "cost of credit" is the cost of lending money or extending a credit facility.

However, it is submitted that in the aforementioned cases the Tribunal was concerned with a product or service that was created by the credit provider itself and charged to consumers who could acquire it at their option. It was definitely not a necessary product or service. This is markedly different from the situation where the credit provider is invoiced for an amount that is made up of the sale price of a vehicle plus the dealer's "on the road, administration and handling fees" even before one cent is added as credit costs by the credit provider. These "on the road fees" relate to vital services undertaken by dealers to ensure that the car can eventually be delivered to the consumer – not services that credit providers undertake in the ordinary course of their business.

6 *Volkswagen Financial Services v the National Credit Regulator*

6.1 *Background*

The Tribunal recently pronounced on the issue of "on the road", administrative and/or handling fees in the matter of *Volkswagen Financial Services v The National Credit Regulator*.[46] The facts were that the NCR issued a compliance notice in terms of section 55(1) of the NCA against Volkswagen on 23 October 2017. The NCR alleged that it undertook an investigation, which revealed that Volkswagen charged consumers certain "on the road", administrative and/or handling fees on credit agreements, which fees were disguised as service and delivery fees. It alleged that the charging of these fees in credit agreements were in contravention of sections 3*(e)*, 89(2)*(c)*, 90(1), 90(2)*(b)*(iv)*(aa)*, 90(2)*(e)*, 90(2)*(f)*, 91(2), 100(1)*(a)*, 101(1)*(a)*, and 102(1) and (2) of the NCA in that:

- the "on the road", administrative and/or handling fees were credit fees or charges prohibited by section 100(1)*(a)* of the NCA;

[45] *Edcon Holdings Ltd v The National Consumer Tribunal* (unreported, GNP case no A237/2017 (23 May 2018)). The judgment was delivered by Louw J (Mdalana AJ concurring).

[46] NCT/94937/2017/56(1). The NCR also referred a largely similar complaint against BMW Financial Services to the Tribunal whereupon BMW applied for consolidation of the proceedings with that of Volkswagen Financial Services. This consolidation application was, however, rejected on the basis that the complaints against BMW and Volkswagen were not similar in all respects. See *BMW Financial Services (SA) (Pty) Ltd v National Credit Regulator; In Re BMW Financial Services v National Credit Regulator; In Re Volkswagen Financial Services (SA) (Pty) Ltd v National Credit Regulator* [2018] ZANCT 49 (5 June 2018).

- the "on the road", administrative and/or handling fees were not credit fees or charges permitted to be charged on a credit agreement in terms of section 101(1) of the NCA;
- the "on the road", administrative and/or handling fees were not credit fees or charges that could be included in the principal debt deferred in respect of an instalment or a lease agreement in terms of section 102(1) of the NCA;
- the dealer invoices containing the "on the road", administrative and/or handling fees were supplementary agreements or documents that contained provisions relating to fees prohibited by sections 100(1)*(a)*, 101(1), 102(1) and 101(2)*(a)* of the NCA;
- the "on the road", administrative and/or handling fees were disguised or inaccurately disclosed as service and delivery fees in credit agreements in contravention of section 3*(e)* read with section 92(2) of the NCA; and
- Volkswagen had charged consumers "on the road", administrative and/or handling fees on credit agreements despite not having been chosen by the consumers to act as their agent and arrange for these services in contravention of section 102(2)*(a)* of the NCA.[47]

In terms of the compliance notice the NCR required Volkswagen to stop charging consumers the impugned fees by 24 October 2017 and to provide written confirmation to the NCR by 2 November 2017 that it had acted accordingly. It also required Volkswagen to submit a list of all the consumers to whom these fees were charged from 2007, specifically the number of consumers involved, and the amounts charged to the consumers. Volkswagen was further required to refund the consumers and to submit a report to the NCR setting out details regarding how many consumers were refunded and the total amounts of refunds effected.[48]

Volkswagen, however, denied having contravened the Act as alleged, and was of the view that it was not obliged to take the steps set out in the compliance notice and that the compliance notice should be set aside. At the time of the hearing, it was inter alia placed on record that the parties agreed that the NCR would not persist with its case that the dealer's invoice is a supplementary agreement as contemplated by section 91(2), which regulates impermissible fees in contravention of sections 100(1)*(a)*, 101(1), and 102(1) and (2)*(a)* of the NCA.[49]

It was further common cause between the parties that Volkswagen finances motor vehicles for consumers based on invoices it received from dealers; that such financing was done by means of credit agreements, specifically instalment agreements; and that the *invoice from the dealer* included an "on the road fee", an administrative fee and/or a handling fee. It was also recorded to be common cause between the parties that the aforementioned charges were "incorrectly/wrongly/

[47] Par 8.
[48] Par 9.
[49] Par 13.4.

confusingly disclosed, described and/or labelled" as a "service & delivery charge" in the credit agreement.[50]

6.2 *The parties' submissions*

Volkswagen denied that it charged consumers fees in contravention of the NCA but contended that it was the dealer who charged the impugned amounts in terms of a separate agreement concluded between the consumer and the dealer. Volkswagen submitted that these fees:[51]

- formed part of a separate cash purchase and sale agreement between the dealer and the consumer to which Volkswagen was not a party;
- formed part of an amount invoiced by the dealer and, as such, formed part of the principal debt, being the deferred amount, as envisaged in section 101(1)*(a)* of the NCA; and
- had been incorrectly and separately reflected in Part E of the credit agreement as a "service and delivery charge" instead of in Part A as part of the principal debt.

Volkswagen submitted that section 102(1) and (2) was not applicable as Volkswagen, as credit provider, does not charge the fees in accordance with section 102 of the NCA. It also submitted that these "on the road", administrative and/or handling fees were not negotiated by Volkswagen as the consumer's agent, and that Volkswagen did not render the services billed for by the dealer and did not get paid for them.

Volkswagen's managing director testified that the consumer identifies the product and services he wants to purchase from the dealer. The consumer thereafter signs an offer to purchase which is similar to a quotation. The consumer then decides whether or not to have the purchase of the goods and services financed by a financier (credit provider). If the consumer decides to pay cash for the goods and services, he or she gets invoiced directly. However, if the transaction is financed, the invoice is submitted to the credit provider. Volkswagen, as credit provider, is not involved in the discussions between the consumer and the dealer or in the conclusion of the offer to purchase. In fact, it often does not have sight of the offer to purchase. Volkswagen merely gets invoiced by the dealer, and upon receipt of the invoice and proof that the consumer has received the goods and services, Volkswagen then pays the dealer on behalf of the consumer.[52]

Volkswagen therefore took the view that it merely pays the dealer what the dealer charged the consumer (for services that the dealer rendered) and, thus, that it does not "charge" the consumer as envisaged in the NCA. Accordingly, Volkswagen contended that it could not be held responsible for charging the said fees in contravention of the NCA.[53]

[50] Par 13.6.
[51] Par 15.
[52] Par 17.
[53] Par 18.

The NCR, however, submitted that:[54]

- The only agreement relevant to the matter before the Tribunal was the credit agreement concluded between Volkswagen and the consumer, and the NCA applied only to that credit agreement and not to the purchase and sale agreement entered into between the consumer and the dealer.
- Since the credit agreement included an "on the road fee", administrative fee and/or a handling fee described as a "service and delivery charge" not listed in sections 101 and 102 of the NCA, Volkswagen charged amounts beyond what is allowed in those sections and in contravention of the Act.
- Volkswagen purchased the vehicles from the dealer for cash at the price set out in the dealer's invoice, which was issued to Volkswagen.
- Volkswagen then sold the vehicle to the consumer on credit, with the purchase price repayable to Volkswagen (as credit provider) by the consumer in instalments; took delivery of the vehicle from the dealer through the consumer, acting as Volkswagen's agent; and Volkswagen became the owner of the vehicle and retained such ownership of the vehicle until the consumer had paid all the instalments due in terms of the credit agreement.

6.3 *Issues for determination*

The Tribunal indicated that in order to determine the issues between the parties, it had to consider the following:

1. the nature and legal consequences of the transactions between all the parties involved; and
2. whether Volkswagen charged consumers "on the road", administrative and/or handling fees in contravention of the provisions of the NCA. In order to address this question, findings had to be made regarding: (1) the permissible fees, costs and charges pertaining to instalment agreements as provided for by the NCA; (2) the practical application of the meaning of "principal debt", being the deferred amount, as per the relevant provisions of the NCA; (3) the application and implications of section 102(1) and (2) to the issues between the parties; (4) whether Volkswagen was allowed to recoup fees or charges invoiced by dealers from consumers, in terms of instalment agreements but not provided for in the Act; and (5) whether the "on the road", administrative and/or handling fees were disguised or inaccurately disclosed by Volkswagen as service and delivery fees in credit agreements in contravention of section 3*(e)* read with section 92(2) of the NCA.

[54] Par 19.

6.4 *The Tribunal's judgment*

The Tribunal referred to a number of provisions of the Act before handing down its judgment. First, it regurgitated the definition of instalment sale agreement in section 1 of the NCA.[55] It then indicated that section 3 provides that the purposes of the NCA are to

> "promote and advance the social and economic welfare of South Africans, promote a fair, transparent, competitive, sustainable, responsible, efficient, effective and accessible credit market and industry, and to protect consumers, by—
>
> …
>
> *(e)* addressing and correcting imbalances in negotiating power between consumers and credit providers by—
>
> (i) providing consumers with education about credit and consumer rights;
>
> (ii) providing consumers with adequate disclosure of standardised information in order to make informed choices …"

Thereafter the Tribunal referred to section 8(4),[56] which sets out the types of credit transactions, and sections 92(2), 100, 101 and102.

According to the Tribunal it was apparent that a number of transactions are entered into before a credit agreement is concluded, namely:

> "The first and initial transaction is a purchase and sale agreement between the consumer and the dealer when they agree on the goods and services bought and the prices thereof. Pursuant to this agreement the dealer prepares an invoice which is then issued to the consumer, in the event of a cash sale, alternatively to the credit provider in the event of the acquisition of the vehicle by the consumer being financed in terms of a credit transaction. If the sale is for cash the matter ends there except for insurance and so forth which fall outside the scope of the issues to be determined by the Tribunal

[55] In terms of s 1 of the NCA an "instalment agreement" means "a sale of movable property in terms of which—
(a) all or part of the price is deferred and is to be paid by periodic instalments;
(b) possession and use of the property is transferred to the consumer;
(c) ownership of the property either—
(i) passes to the consumer only when the agreement is fully complied with; or
(ii) passes to the consumer immediately subject to a right of the credit provider to repossess the property if the consumer fails to satisfy all of the consumer's financial obligations under the agreement; and
(d) interest, fees and other charges are payable to the credit provider in respect of the agreement, or the amount that has been deferred …"

[56] Section 8(4) provides that "an agreement, irrespective of its form but not including an agreement contemplated in subsection (2), constitutes a credit transaction if it is—
(a) a pawn transaction or discount transaction;
(b) an incidental credit agreement, subject to section 5(2);
(c) an instalment agreement;
(d) a mortgage agreement or secured loan;
(e) a lease; or
(f) any other agreement, other than a credit facility or credit guarantee, in terms of which payment of an amount owed by one person to another is deferred, and any charge, fee or interest is payable to the credit provider in respect of—
(i) the agreement; or
(ii) the amount that has been deferred."

> in this case. Where the sale is to be financed the dealer issues the invoice to the credit provider. Thereupon the *consumer as agent of the credit provider* takes delivery of the goods and services and the credit provider then becomes owner of the goods and services upon paying the dealer's invoice. This constitutes the second transaction which is a cash transaction that falls outside the scope of the NCA. The credit provider remains the owner of the goods and services until the consumer has made full payment in accordance with the credit agreement."[57]

The Tribunal indicated that the credit provider then "on sells" the goods and services to the consumer and remarked:

> "For the financier to be repaid for the cash outlay it made on behalf of the consumer, for the goods and services, as per the dealer invoice, a *third agreement comes into existence between the consumer and the Applicant/financier*. In the matter before the Tribunal this is the credit agreement, more specifically an instalment agreement."[58]

The Tribunal thereupon referred to the applicable provisions of Volkswagen's credit agreement which bears out this process.[59] This agreement inter alia contained clause 20.1.1, which read: "You choose us as your agent for arranging the services as set out in section 102 of the Act and you agree that the value of those services are included in the principal debt." It also contained clause 1.14 that defined "principal debt" as the "amount that we will provide you as set out in Part A of the quotation/ cost of credit".[60]

The Tribunal remarked that the agreements that Volkswagen entered into with consumers were instalment agreements as defined in and regulated by the NCA and that Volkswagen thus had to adhere to the spirit and provisions of the Act and specifically to the provisions that the NCR alleged that Volkswagen had contravened.[61]

Thereafter the Tribunal turned its attention to the permissible fees, costs and charges in terms of the NCA pertaining to instalment agreements. It remarked that these permissible fees, costs and charges are carefully structured in the Act. It stated that section 100(1)*(a)* sets out a broad prohibition on the credit provider requiring a consumer to pay "a credit fee or charge prohibited by this Act" and is followed by section 100(1)*(b)* to *(d)* that provides for the quantum of the fees, costs and charges beyond which the credit provider may not go.[62] The Tribunal further indicated that section 101 contains a *closed list* of seven items, outside of which the credit provider may not require payment from the consumer. It pointed out that the "on the road", administrative and/or handling fees are not included in this closed list of items.[63]

The Tribunal then turned to the concept of "principal debt", remarking that the principal debt is one of the items referred to in the closed list contained in section 101.

57 Par 28 – authors' emphasis.
58 Par 29 – authors' emphasis.
59 Par 30.
60 Par 31.
61 Par 32.
62 Par 33 and 34.
63 Par 35.

It referred to the definition of "principal debt" in section 1 of the NCA read with sections 101(1)*(a)* and 102(1) and stated:

> "It is defined as 'the amount calculated in accordance with section 101(1)*(a)* of the NCA.' Section 101(1)*(a)* provides '… *the principal debt, being the amount deferred* in terms of the agreement, *plus the value of any item contemplated* in section 102'. The term 'amount deferred' is not defined in the NCA."[64]

The Tribunal indicated that the principal debt is in essence "an amount owed by one person to another". It remarked that payment is deferred and instalments are paid in terms of an instalment agreement (as defined in the Act), which also forms a basis for levying charges, fees or interest payable to the credit provider.[65] The Tribunal then referred to section 102 and remarked that to "calculate" the principal debt in accordance with section 1 read with section 101(1)*(a)* and section 102(1) therefore requires adding the items allowed in terms of section 102(1)*(a)* to *(f)* to the "principal debt". It again pointed out that the "on the road fee", administrative and/or handling fees are not included in the closed list in section 102(1)*(a)* to *(f)* that the credit provider may include in the principal debt and charge consumers for.[66]

The next aspect the Tribunal dealt with was whether Volkswagen "charged" the consumer as envisaged by the NCA. The Tribunal did not go into any depth in this question other than to indicate it was of the view that when Volkswagen sells the goods and services to the consumer, and the consumer and Volkswagen enter into an agreement regarding the payment of the deferred amount and the charges, fees and interest in respect of the deferred amount, Volkswagen "charges" the consumer within the meaning of section 102(1) and (2) of the Act. It stated: "[T]his is moreover so as the charges, fees and interest in respect of the deferred amount do not follow from the invoice but flow from the credit/instalment entered into between the consumer and the Applicant/financier."[67]

The next issue dealt with by the Tribunal was what constitutes the principal debt/deferred amount? Here the Tribunal sought to answer the following two questions:[68]

1. If an item is *not* listed in section 102(1)*(a)* to *(f)* and provided the credit provider meets the requirements of section 102(2), may the credit provider charge the consumer for such items as part of the principal debt/deferred amount?
2. Does the credit provider have the *discretion* to include any *expense item* in the principal debt, as long as it advanced the money to the consumer and deferred the payment thereof?

64 Par 36 – Tribunal's emphasis.
65 Par 37.
66 Par 39 and 40.
67 Par 44.
68 Par 45.1 and 45.2.

Volkswagen's main contention was that the principal debt

> "normally include[s] the cash amount advanced to the consumer or the cash price for the goods and services … the *total purchase price agreed upon between the dealer and the consumer, which will include all the items chosen by the consumer from the dealer …*"

It indicated that the "on the road fee", administrative and/or handling fees are included in the principal debt/invoiced amount, and the *invoiced amount is based on the agreement between the dealer and the customer*.[69]

The NCR, on the other hand, persisted with its view that the fees charged by the dealer to Volkswagen, in terms of the dealer's invoice, *are paid by Volkswagen* when it acquires ownership of the vehicle from the dealer. It stated that *Volkswagen then charges the same amount to the consumer*, including the charges that fall outside of the closed list of charges allowed in contravention of sections 101 and 102 of the NCA.[70]

The Tribunal indicated that in determining this issue, it considered the fact that the legislature deemed it fit to list specific items that may be included in the principal debt (once specific requirements have been met) *and as costs of credit*. In the Tribunal's view this was done to give effect to the intent of the NCA, as captured to some extent in section 3*(c)*, which refers to the object of the Act to promote responsibility in the credit market by

> "(i) encouraging responsible borrowing, avoidance of over-indebtedness and fulfilment of financial obligations by consumers; and
>
> (ii) discouraging reckless credit granting by credit providers and contractual default by consumers …"[71]

Volkswagen further submitted that it was under no obligation, and the NCA does not impose a duty on it, to interrogate fees and charges on the dealer's invoices. It therefore submitted that the NCR was acting arbitrarily and *ultra vires* by imputing these duties on credit providers. The Tribunal agreed with this contention that the legislature, through the NCA, "does not impose an obligation on credit providers to 'police' consumers and impose an obligation on credit providers to veto what consumers may or may not buy and may and may not finance in terms of a credit agreement". Having stated this, the Tribunal remarked that the NCA does, however, impose an obligation on a registered credit provider to adhere to the provisions of the Act and, in terms of section 100(1)*(a)*, to not charge an amount to, or impose a monetary liability on, a consumer in respect of a credit fee or charge prohibited by the NCA.[72] It stated that in this instance, the provisions of sections 101 and 102 specifically are pertinent to this matter in respect of the two main monetary components of a credit agreement, namely the principal debt or deferred amount and the costs of credit. Given that these are specifically enumerated, the Tribunal indicated that Volkswagen as a credit provider is obliged to keep the principal

69 Par 46.
70 Par 47.
71 Par 49.
72 Par 50–52.

debt and cost of credit within the parameters of this section. It remarked that it is within the power and ability of Volkswagen as credit provider to do so because it is Volkswagen as credit provider who gets invoiced the amounts and not the consumer, and accordingly it is *Volkswagen, at the point of invoicing, who could interrogate the items on the invoice and not the consumer*. Further, it is incumbent on the credit provider to ensure that it does not pay a dealer upon invoice for costs, fees and charges proscribed under the NCA, and then as a result being forced into recouping them from the consumer in the credit agreement in contravention of the NCA.[73]

The Tribunal remarked that the evidence before it (and conceded by Volkswagen) was that it had not included the "on the road," administrative and/or handling fees in the principal debt but had incorrectly reflected it separately in the credit agreement as a "service & delivery charge" instead as part of the deferred amount. The Tribunal stated that the admitted fact that these fees had been disclosed in Part E of the instalment sale (sic) agreement, and not in Part A of the agreement as part of the principal debt, *"*gave credence to the NCR's assertions that these fees did in fact not form part of the principal debt*"*.[74]

The Tribunal noted the submissions by Volkswagen that it did not, as credit provider, derive any financial benefit from the payments of the "on the road" and administrative and/or handling fees. However, it disagreed with this contention because although Volkswagen paid these amounts over to the dealer, the Tribunal remarked that it "derives interest income from these deferred amounts over the life of the instalment agreement. Aggregating these amounts over to its loan book, the profit to the credit provider could amount to substantial amounts of money."[75]

The Tribunal remarked that there is a danger that the purpose of the NCA may not be realised if the credit provider may include any or all amounts invoiced by the dealer, pay the dealer and then include these amounts in the credit agreement, requiring the consumer to pay. It remarked that, first, the danger may come from the potential perverse incentive for the credit provider to *not interrogate* the amounts it gets invoiced for by the dealer because the inclusion of the impugned fees in the principal debt increases the credit provider's profit from the "charges, fees and interest's income". Secondly, it indicated that the danger comes from these items increasing the capital sum owing and deferred and as a result increases the costs of credit to consumers,

> "albeit that in the overall scheme of the credit agreement these items may be small. However, with charges, fees or interest levied thereon over the term, generally 60 to 72 months, the amounts owing can become substantial. The impact on consumers' indebtedness could potentially even be more severe if those consumers go into debt counselling and get debts re-arranged over a longer period of time."[76]

Thus, the Tribunal held that it could not have been the intention of the legislature to allow the credit provider carte blanche *to add into the principal debts items not listed in section 102(1)*. Accordingly, the Tribunal took the view that the credit

[73] Par 53 and 54.
[74] Par 55.
[75] Par 56.
[76] Par 57.

provider is obligated to ensure that it meets the prescripts of the NCA and that it does not include items beyond those mentioned in section 102(1) in the principal debt or deferred amount.[77] It stated that allowing credit providers carte blanche to include any items in the credit agreement on the basis that they had been invoiced for those items by the dealer, and then passing it on to the consumer in the credit agreement, manifestly runs counter to the purpose of the NCA as reflected in section 3*(c)*. It remarked that "restricting the items that may be included in the principal debt/deferred amount and payments the credit agreement (and by extension the credit provider) may not require from consumers … begins to give content to this purpose of the NCA set out in section 3*(c)* of the NCA".[78]

The Tribunal subsequently dealt with the allegation that Volkswagen had charged consumers the "on the road", administrative and/or handling fees in credit agreements despite not having been chosen by consumers to act as their agents to arrange for services for which these fees are charged as required by section 102(2)*(a)* of the NCA. It pointed out that where a credit provider charges a consumer fees in terms of section 102(1), then it should meet the requirements of section 102(2) as alluded to above in paragraph 3. However, in light of its finding that the charging of the "on the road" and administrative and/or handling fees constituted a contravention of the Act, the Tribunal indicated that it would not deal with the question whether Volkswagen acted as the agent of consumers in procuring these services.[79]

The next issue the Tribunal dealt with was that the "on the road", administrative and/or handling fees were disguised or inaccurately disclosed as "service and delivery fees" in credit agreements in contravention of section 3*(e)* read with section 92(2) of the NCA. The Tribunal stated that it was common cause between the parties that:[80]

- a "service & delivery" fee is actually reflected in Part E of the agreement;
- the "service & delivery" item reflected in Part E of the instalment agreement consists of the "on the road fee" and administrative and/or handling fees; and
- Part A of the "quotation/cost of credit" document excludes the "service & delivery" fees.

As indicated above, Volkswagen submitted that the disclosure of these fees was made in error in Part E of the instalment agreement instead of in Part A. However, the NCR claimed that irrespective of the aforementioned, the "on the road", administrative and/or handling fees were disguised or inaccurately disclosed as "service and delivery fees" in credit agreements in contravention of section 3*(e)* read with section 92(2) of the NCA, and that this was misleading as it did not

[77] Par 58.
[78] Par 60.
[79] Par 61–63.
[80] Par 64.

provide adequate disclosure and standardised information to protect consumers against deception.[81]

The Tribunal eventually found that:

- Volkswagen had inaccurately and in a misleading way disclosed costs and charges in the instalment agreement in contravention of section 92(2) of the NCA. It stated that the "service and delivery charge" should have been reflected for what it was, namely "on the road", administrative and/or handling fees.[82]
- The "on the road" administrative and/or handling fees were credit fees or charges prohibited by section 100(1)*(a)* of the NCA.
- The "on the road", administrative and/or handling fees were not credit fees or charges permitted to be charged on a credit agreement in terms of section 101(1) of the NCA.
- The "on the road", administrative and/or handling fees were not credit fees or charges that could be included in the principal debt deferred in terms of an instalment agreement or a lease agreement according to section 102(1) of the NCA.
- The "on the road", administrative and/or handling fees were disguised or inaccurately disclosed as "service and delivery fees" in credit agreements in contravention of section 3*(e)* read with section 92(2) of the NCA.

The Tribunal ordered, in accordance with section 56(2), the modification of the compliance notice to delete references to contraventions of sections 90(2)*(e)* and 91(2) given that the NCR had abandoned its allegations that the dealer invoices reflecting the "on the road fees" were supplementary agreements.[83] With regard to the steps that the NCR required Volkswagen to take in terms of the compliance notice, the Tribunal indicated that it could not confirm an order for compliance with a retrospective date, due to the impossibility adhering to such an order would impose on the relevant party. The Tribunal thus modified the time within which Volkswagen had to comply with the compliance notice.[84] It further stated that *if it were to order a refund of the fees paid, the consumer would derive a benefit from an amount paid over by Volkswagen without the consumer paying for it*. However, it held that at the same time, Volkswagen should not be deriving financial benefit out of its unlawful levying of these fees and thus it modified paragraph 3 of the compliance notice to read:

[81] Par 65 and 66. Volkswagen's view was that s 3*(e)* of the NCA relates to statement of purpose and that it was not capable of being contravened. The Tribunal agreed with this (par 67 and 69).

[82] Par 68.

[83] Par 73–75.

[84] Par 78.3. Volkswagen was required from 10 April 2019 to cease charging "on the road", administrative and/or handling fees and submit written confirmation to this effect by no later than 25 April 2019.

> "Applicant is required to, in respect of all the consumers identified in B2, calculate the total amount of charges, fees or interest levied on 'on the road', admin and/or handling fees and refund all those consumers those charges fees or interest levied and submit to the NCR a report by an independent auditor setting out – (a) The number of consumers who were charged those charges, fees or interests; (b) The number of consumers who were refunded those charges, fees or interests; and (c) The total amount of charges, fees or interest refunded to consumers."

7 The true nature of "on the road fees"

It strikes one as rather odd that in a 30-page judgment that focuses on "on the road", administrative and/or handling fees there is not once any indication of what exactly these fees entail. The nature of these fees was simply never considered by the Tribunal. It is submitted that this is a grave mistake as it could have led to quite a different outcome in the *Volkswagen* matter.

Bearing in mind the various transactions at play when a consumer buys a vehicle from a dealer is pivotal as it illuminates the various items and amounts that are being charged in this process, by whom they are charged and what their nature is. As explained in paragraph 2 above, when a consumer buys a car from a dealership, the vehicle will be sold at a specific retail price. To this price then gets added amounts that are referred to as "on the road", administrative and/or handling fees that cover a wide array of expenses that the dealership incurred in order to make the vehicle available for delivery to the consumer. These items include necessary steps, such as doing pre-delivery inspections, obtaining roadworthy certificates, paying fees to the conduct regulator for F&I services, and so forth. In particular these items *are charged by the dealer.* The bottom line is, however, that the "on the road", administrative and/or handling fees are expenses that were *actually incurred* by the dealer. These expenses were incurred *for the benefit of the consumer*. What is more: a consumer will have to pay these expenses regardless of whether he or she buys a car for cash or on credit.

It should then become clear that these expenses in fact have nothing to do with "cost of credit" because they have to be paid even if it is a cash transaction. Furthermore, it is clear that where a vehicle is subsequently financed, these expenses and fees are not "charged" by the credit provider and are not latched onto the transaction as cost of credit – they were already present *even before any credit was granted*. In fact, they are already charged by and reflected on the dealer's invoice as part of the "deferred amount" as referred to in section 101(1)*(a)* – that is, the part of the principal debt *before* the section 102 costs are added. These fees are never "charged" by the credit provider and the credit provider does not, in principle, derive any benefit from them except to the extent that they form part and parcel of the principal debt (being part of the "deferred amount"). The credit provider is then duly entitled to be compensated by means of being able to charge the *costs of credit* mentioned in section 101(1)*(b)* to *(g)*, because he or she pays these expenses on behalf of the consumer and defers repayment thereof by the consumer.

It is by this time well known that the NCA calls for a purposive but balanced interpretation aligned with the purposes of the Act as stated in section 3. In the context of this discussion, the aim of the Act to protect consumers deserves its rightful place. As pointed out in the discussion of the 2004 Policy Framework and

Memorandum on the National Credit Bill in paragraph 1 above, the objective was to inter alia protect consumers by regulating the "costs of credit". This, it is submitted, ties in with the purpose of the Act to promote the development of an accessible credit market (as per section 3*(a)*); to promote responsibility in the credit market by not allowing costs that are so high that it leads to consumer over-indebtedness (as per section 3*(c)* and *(g)*); and to promote disclosure as a means to enable informed choices by consumers (as per section 3*(e)*). Sight should, however, not be lost of the fact that the intention of the legislature was to regulate "cost of credit". The intention was never to draw aspects that do not qualify as "cost of credit" into the regulatory realm or to regulate aspects that fall outside the purview of credit legislation. As such, it was never the intention, nor would it have been competent for a credit regulator to also attempt to regulate prices (*ie* "deferred amounts"). The regulatory net of the NCA covers credit providers, consumers, credit bureaux and payment distribution agents. It also covers certain instances of conduct by persons or entities who undertake the regulated activities of the Act whilst being unregulated. The Act does not give the NCR authority to also start querying the prices that manufacturers ask for their goods or the services rendered by motor dealerships. That is for the markets to sort out or for the Competition Commission to interrogate should it raise any concerns under the Competition Act 89 of 1998.[85]

Why should the credit provider be penalised for expenses it did not charge? Necessary expenses that were actually incurred on behalf of the consumer – not by the credit provider but by the dealership? Expenses that have nothing to do with the granting of credit? Maybe this also explains why the Act does not define the concept "deferred amount" and why it does not prescribe which amounts make up the deferred amount but only that the principal debt consists of the deferred amount plus section 102 costs. It is simply not within the remit of the NCA, and by extension the NCR, to dictate what must be included in the "deferred amount" precisely because it is not a price regulator.

The Tribunal appears to have been so intent on the noble pursuit of protecting consumers that it seems to have been blinded by the very simple facts underlying the charging of "on the road", administrative and/or handling fees. It agreed that the credit provider does not have to police the amounts levied by dealers; it acknowledged that these amounts form part of the principal debt, although it failed to appreciate that it forms part of the "deferred amount" and not of the section 102 costs. This is, however, not, as viewed by the Tribunal, a matter where the credit provider exercises a "discretion" to add amounts to the principal debt – this is the amount for which the dealership makes the vehicle available to anyone who wants to buy it, whether a sale is for cash or on credit. In any event, chances are that if the Tribunal's judgment is upheld, which it should not be, car prices will merely increase to absorb these dealer expenses and fees.

The next question that may be asked is if a credit provider is not in contravention of the NCA when it includes these dealer expenses and fees as part of the deferred amount which, together with permitted section 102 fees, constitute the "principal

[85] See s 8(1)*(a)* of the Competition Act 89 of 1998 that deals with excessive pricing.

debt", does it then matter at all where these amounts are disclosed and should they even be disclosed? As pointed out, the Regulations cannot generally be used to interpret provisions in the Act. However, in the context of the quotation for and contents of intermediate and large credit agreements, the Act specifically indicates that the prescribed form of the quotation in regulation 29 read with Form 20.1 as well as the list of contents set out in regulation 31 have to be complied with. As indicated in paragraph 4 above, regulation 29 requires the proposed distribution of the principal debt with reference to items listed in section 102(1)*(b)* to *(f)* of the Act *"*and [to] specify any other*"*. Regulation 31(2)*(c)*(i) requires that the agreement reflects the amount of the principal debt, including the amount deferred in terms of the agreement as well as the nature and amount of fees and charges related to the cost of an extended warranty agreement, delivery, installation, initial fuelling charges, connection fees, levies or charges, taxes, and licence or registration fees. However, regulation 31(2)(iii) requires reflection of "the proposed distribution of the principal debt and to whom each amount is paid". Form 20.1, which sets out the format of the quotation, requires Part A to reflect the "credit advanced or value of goods or services provided on credit", whilst Part E requires a reflection of the section 102 charges and fees (which is also reflected as one of the items under Part A). The "optional items" to be reflected in Part D would most likely be "dealer add-ons" but from section 101 and 102 it would not seem as if these extra amounts ("dealer add-ons") are not expressly allowed by the Act. The bottom line is, however, that it does not appear that the Act and the Regulations require the credit provider to specifically set out the "on the road", handling and/or administration fees and the amounts they entail.

Given that these fees represent expenses and fees actually incurred by the dealer on behalf of the consumer in order to facilitate delivery of the vehicle concerned and that it is not dependent on the transaction being cash or credit, coupled with the fact that the legislature's intention was to regulate cost of credit, it is submitted that any attempts by the Act to regulate such fees would be *ultra vires*. From this it would then follow that even the regulation of the fees mentioned in section 102 is *ultra vires* to the extent that the legislature attempts to meddle in the type of charges or fees that can be added to the principal debt. It is consequently submitted that the most that the legislature would legitimately have been able to do was not to try and limit these amounts that can be charged to the few specific items mentioned in section 102. The legislature could only require that any extra fees and charges added to the deferred amount as part of the principal debt should be disclosed so that the consumer would be able to see exactly what the principal debt on which he or she then has to pay the cost of credit in section 101(1)*(b)* to *(g)* eventually amounts to.

This argument may probably not sit so well with those who take the view that the NCA should be a consumer nanny at all times. The reality is, however, that consumers wish to buy vehicles and this process entails certain necessary processes, such as incurring the fees and expenses that are collectively referred to as "on the road", administrative and/or handling fees. Where a consumer decides to finance a vehicle, it is still the consumer who is the beneficiary of these services because it is the consumer who is earmarked to become the eventual owner of the car albeit only

after he or she has paid the amounts due to the credit provider under an instalment agreement. *De facto*, even in such a case, it is the consumer who uses the vehicle on a daily basis and who benefits from the vehicle having passed a safety inspection, being roadworthy and not being subject to eviction by a previous owner – peace of mind that is delivered via "on the road fees", and administrative and/or handling expenses.

If this view does not find favour with the courts, then as pointed out in paragraph 3 above, it is submitted that section 102 should be broadly construed to encompass these fees that are collectively referred to as "on the road", administrative and/or handling fees. Should this be embraced as the fall-back position, it is further submitted that section 102 should actually not be interpreted as constituting a closed list, to the exclusion of some items for which the dealer has actually incurred expenses or validly charged fees that are not mentioned by name in section 102. The reason for this is because – in addition to the argument that it is not for the legislature who seeks to regulate credit to dictate that certain necessary expenses and fees in making the vehicle available for delivery may not be charged, despite them actually having been incurred – it would appear that in its attempts to regulate these items, the legislature also failed to consider items that are essential in the context of sales of vehicles. As such, one may ask why section 102(1)*(b)* allows for the cost of an extended warranty agreement but not for the cost of a service plan? It would also be apt to point out here that the remark by the Tribunal in the *Volkswagen* case that the fact that Volkswagen disclosed the "on the road", administrative and/or handling fees in Part E instead of Part A of the quotation "gives *credence* to the NCR's assertions that these fees did in fact not form part of the principal debt" is fallacious. This is because it has been pointed out in paragraph 4 above that the items to be listed in Part E are section 102 items and that Form 20.2 requires the items in Part E to also be reflected in Part A as part of the principal debt.

8 Conclusion

It may thus be argued that the reason why "on the road", administrative and/or handling fees are not specifically mentioned in the NCA is because that by their very nature they do not fall to be regulated in an Act that seeks to regulate the cost of credit. Alternatively, these "collective terms" should be broken down to reflect what they really are, and in such an event they may then be grouped under section 102, which should be broadly interpreted to accommodate all such expenses and fees that are necessary – not only those listed in the section. If one accepts this latter view, then it does not at all seem strange that Volkswagen reflected these items as "service and delivery fees" in Part E of the quotation. In any event, the discrepancies between the items listed section 102 and regulation 29, which also requires the credit provider to "specify any other" possibly indicates that whoever drafted the Regulations may have appreciated that there are other expenses and fees apart from those specifically mentioned by name in section 102 that would also qualify as amounts to be allowed under section 102. The Act also does not mention "dealer add-ons" but Part D of Form 20.2 coincidentally seems to provide for these – is this evidence that the drafters of the Regulations had a better appreciation of the reality of the landscape of car sales than the legislature responsible for the drafting of the Act?

ABLU 2019 (Part ONE).indb 109 2019/10/09 9:49 AM

Finally, credit providers should not be punished merely because they provide credit and because the fact that they are regulated under credit legislation makes them an easy target on whom to offload some of the liability that should be attributed to other parties that are also involved in the initial stages of the lifecycle of a credit agreement. Just as consumers, for whose benefit the "on the road", administrative and/or handling fees were incurred by dealerships, cannot be expected to be refunded amounts that were paid on their behalf (as the Tribunal duly acknowledged) it can also not be expected that credit providers have to pay back interest they legitimately earned on credit they provided to consumers who wanted to acquire motor vehicles where the deferred amount included dealer charges that were actually incurred and over which the credit provider has no control. It is trite that consumers who apply for credit generally do not want to do so in a piecemeal fashion by paying for certain items in cash and financing the rest. Arguably many of them would not have the cash to pay the "on the road", administrative and/or handling fees upfront before entering into a credit agreement, which would mean that they would then be prevented from accessing credit to finance the acquisition of the vehicle. Maybe it is time to recognise that credit providers cannot be expected to police the legitimate business of others and that consumers also have the responsibility to check these charges themselves and negotiate better deals with dealers where possible if they are of the opinion that these expenses and fees are higher than normal. After all, it is the consumer who wants the car.

In addition, one may also ask whether the Tribunal was in fact the correct forum to hear this matter, given that section 100(3) stipulates that a person who contravenes section 100 is guilty of an offence. It appears that the legislature intended contravention of section 100 to constitute more of a transgression than just being prohibited conduct, hence labelling it as an "offence" which attracts criminal sanction.[86]

Section 27*(a)*(ii)[87] read with section 150*(a)*[88] makes it clear that the Tribunal can hear complaints of prohibited conduct whereas offences fall within the jurisdictional remit of the National Prosecuting Authority. Initially section 1 of the NCA provided that "prohibited conduct" means

> "[a]n act or omission in contravention of this Act, *other than an act or omission that constitutes an offence under this Act*, by—
>
> *(a)* an unregistered person who is required to be registered to engage in such act; or
>
> *(b)* a credit provider, credit bureau or debt counsellor."[89]

[86] In terms of s 161 any person convicted of an offence other than a contravention of s 160(1), which deals with failure to comply with a Tribunal order, is liable to a fine or imprisonment for a period not exceeding 12 months, or both.

[87] Section 27*(a)*(ii) indicates that the Tribunal may adjudicate in relation to any "allegations of prohibited conduct by determining whether prohibited conduct has occurred and, if so, by imposing a remedy provided for in this Act".

[88] Section 150*(a)* provides that the Tribunal may make an order declaring conduct to be prohibited in terms of the NCA. It can then, in terms of s 150*(c)* read with s 151, impose an administrative fine on a person who has engaged in such prohibited conduct.

[89] Authors' emphasis.

This definition was subsequently amended so that it now merely reads that prohibited conduct means "an act or omission in contravention of this Act". This apparently means that certain conduct may constitute both prohibited conduct as well as an offence. It further appears to mean that where conduct constitutes both prohibited conduct and an offence, the Tribunal would be able to hear complaints relating to such conduct as the bar relating to offences has been removed from the definition of prohibited conduct.

However, section 140 of the NCA, which indicates how the National Credit Regulator has to act in relation to the outcome of a complaint, has not been amended. This section inter alia provides that after completing an investigation, the National Credit Regulator may make a referral to the Tribunal if it believes that the person investigated has engaged in prohibited conduct (section 140(1)*(b)*) *or* it may refer the matter to the National Prosecuting Authority if the complaint concerns an offence under the NCA (section 140(1)*(d)*). Interestingly, section 141, which provides for direct referral to the Tribunal by a consumer in the event that the Regulator has issued a non-referral in respect of a complaint, reads that such direct referral is only competent if the Regulator issued a non-referral in response to a complaint "other than a complaint concerning section 61 or an offence in terms of this Act". Section 141 has thus not been amended by the 2014 National Credit Amendment Act to align it with the definition of "prohibited conduct" or whatever it was that the legislature attempted to effect by that amended definition. The effect now possibly may be that not only is the Regulator authorised to decide whether to refer a matter to the Tribunal or the National Prosecuting Authority (given that section 140 provides no guidelines as to the basis on which such discretion should be exercised) but consumers who complain about conduct that is at the same time prohibited conduct and an offence are excluded, subsequent to a non-referral by the Regulator, from approaching the very entity that was created to provide them with speedy and cheap redress, whereas the Regulator may approach such entity to pursue a finding of prohibited conduct where it decides to refer a complaint itself. Given that the Regulator is not provided the option to both refer a complaint to the Tribunal *and also* refer it to the National Prosecuting Authority, it is submitted that the Regulator will very likely, without fail, choose to refer these types of matters to the Tribunal to obtain a speedy outcome and generate an administrative penalty. It would mean that labelling the conduct in section 100 as "offences" would become meaningless as these offences will never be referred for criminal prosecution. Given that it is a well-known principle of statutory interpretation that the legislature does not intend to make meaningless legislation, it may be argued that the obligation on the Regulator to refer conduct that is labelled to be an offence to the National Prosecuting Authority, as per section 140(1)*(d)*, still remains, and thus the jurisdiction that the Tribunal now assumes in relation to contraventions of section 100 is questionable.

Twin Peaks and the impact of the Consumer Protection Act on financial products and services

MONICA LAURA VESSIO*

1 Introduction

The Financial Sector Regulation Act[1] came into force on 29 March 2018.[2] A number of provisions came into operation in April, May, September and October 2018, other provisions between January and April 2019, while still others will become operative in April 2020.[3] The object of the Financial Sector Regulation Act is to achieve a stable financial system that works in the interests of financial customers and that supports balanced and sustainable economic growth by establishing, in conjunction with the specific financial sector laws,[4] a regulatory and supervisory framework that promotes financial stability, the safety and soundness of financial institutions, the fair treatment and protection of financial customers, the efficiency and integrity of the financial system, the prevention of financial crime, financial inclusion, transformation of the financial sector and confidence in the financial system.[5]

* Research Associate: Centre for Banking Law, University of Johannesburg.

1 9 of 2017, hereinafter the "Financial Sector Regulation Act" or "Twin Peaks". For a discussion on the Act, see Moorcroft and Vessio in Moorcroft *Banking Law and Practice* (2018 *et seq*) Chapter 39.

2 GN 169 in *GG* 41549 of 29 March 2018, assented to on 21 August 2019.

3 For the various commencement dates, see GN 169 in *GG* 41549 of 29 March 2018. For the date of commencement of s 290 in respect of the amendments to the Financial Markets Act 19 of 2012 published in GN R99 in *GG* 41433 of 9 February 2018, see GN 169 in *GG* 41549 of 29 March 2018. The Financial Sector Regulations, 2018 were published in GN R405 in *GG* 41550 of 29 March 2018 and the Regulations relating to the Levies on Financial Institutions in GN 384 in *GG* 42579 of 12 July 2019.

4 The financial sector laws are the Financial Sector Regulation Act; the legislation listed in schedule 1, which are the Pension Funds Act 24 of 1956, the Friendly Societies Act 25 of 1956, the Banks Act 94 of 1990, the Financial Services Board Act 97 of 1990, the Financial Supervision of the Road Accident Fund Act 8 of 1993, the Mutual Banks Act 124 of 1993, the Long-term Insurance Act 52 of 1998, the Short-term Insurance Act the 53 of 1998, the Financial Institutions (Protection of Funds) Act 28 of 2001, the Financial Advisory and Intermediary Services Act 37 of 2002, the Collective Investment Schemes Control Act 45 of 2002, the Co-operative Banks Act 40 of 2007, the Financial Markets Act 19 of 2012, the Credit Rating Services Act 24 of 2012, the Insurance Act 18 of 2017; and any regulation or regulatory instrument made in terms of Twin Peaks or made in terms of a law referred to in schedule 1.

5 Section 7 of the Financial Sector Regulation Act. In other words, it is the Act that regulates the system within which the financial sector role players and financial customers operate in. It is an overarching piece of legislation meant as a macro-management rather than a micro-management device.

ABLU 2019 (Part ONE).indb 113 2019/10/09 9:49 AM

The Act is made up of 305 sections which are separated into 17 chapters. It is not surprising that it is of this size as the industry it regulates is large and the products and services diverse.[6] Twin Peaks[7] introduces new authorities, councils, working groups, committees and subcommittees, all of which are obligated to cooperate with each other. It has overhauled the financial services and products sector and repealed[8] and amended[9] several pieces of legislation.

The effects of the Financial Sector Regulation Act on the Consumer Protection Act[10] can be found in section 10(1) of the Financial Sector Regulation Act.[11] As will become evident in the discussion to follow, the application of the Consumer Protection Act to legislation in the financial sector and the interaction between it and Twin Peaks is not immediately obvious from a cursory reading of section 10. If anything, this section has raised a number of uncertainties, most of which I have only been able to mention here but all of which will have to be canvassed by the courts or reconsidered by the legislature.

[6] The Act contains definitions in s 1. It defines "financial products", "financial services" and "financial stability" individually in s 2, 3 and 4 respectively.

[7] So-called because the model is characterised by two equal and independent peaks: the Prudential Authority and the Financial Sector Conduct Authority. See n 12 and n 13 below. It appears that the Twin Peaks model originated in Australia, which had legislated to incorporate the Australian Securities and Investments Commission and the Australian Prudential Regulation Authority in 1998. The United Kingdom moved to a Twin Peaks model of financial regulation in April 2013. This saw the Financial Services Authority cease to exist and its work split between two new regulatory authorities, the Prudential Regulation Authority, part of the Bank of England, and the Financial Conduct Authority (https://www.lexology.com/library/detail.aspx?g=98c9c22a-63e4-4852-abd9-6210168a21b7 (accessed 4 August 2019)). There are, however, many variations in the design of Twin Peaks models and there is no archetypal model. Choices are made around key elements in areas such as structural design, operational independence and regulatory coordination (Godwin, Howse and Ramsay "A jurisdictional comparison of the Twin Peaks model of financial regulation" (2017) 18 *Journal of Banking Regulation* 103).

[8] Four Acts have been repealed: the Financial Services Board Act 97 of 1990 (this Act is repealed in various stages as follows: as of 1 April 2018, other than in respect of the repeal of the definitions of "appeal board", "financial institution" and "trust property", s 10(3), s 10A, s 14, s 15, s 18–26, s 26A and s 27–30; as of 28 September 2018 in respect of s 1 for the definition of "appeal board" and s 26B; and as of 1 April 2020, s 1 in respect of the repeal of the definitions of "financial institution" and "trust property" and s 15A and s 16); Policy Board for Financial Services and Regulation Act 141 of 1993; Inspection of Financial Institutions Act 80 of 1998; and the Financial Services Ombud Schemes Act 37 of 2004 (as of 1 October 2018). See GN R99 in *GG* 41433 of 9 February 2018, read with GN 169 in *GG* 41549 of 29 March 2018 (schedule 4 of the Financial Sector Regulation Act).

[9] The amended legislation as well as details of these amendments can be found in schedule 4. The amended Acts are: Insolvency Act 24 of 1936; Pension Funds Act 24 of 1956; Friendly Societies Act 25 of 1956; South African Reserve Bank Act 90 of 1989; Banks Act 94 of 1990; Financial Services Board Act 97 of 1990; Financial Supervision of the Road Accident Fund Act 8 of 1993; Mutual Banks Act 124 of 1993; Long-Term Insurance Act 52 of 1998; Short-Term Insurance Act 53 of 1998; Financial Institutions (Protection of Funds) Act 28 of 2001; Financial Intelligence Centre Act 38 of 2001; Financial Advisory and Intermediary Services Act 37 of 2002; Collective Investment Schemes Control Act 45 of 2002; National Credit Act 34 of 2005; Co-operative Banks Act 40 of 2007; Financial Markets Act 19 of 2012 and the Credit Rating Services Act 24 of 2012.

[10] 68 of 2008, hereinafter "the Consumer Protection Act".

[11] As well as the Competition Act 89 of 1998 (s 10(2) of the Financial Sector Regulation Act). The interaction between the Competition Act and Twin Peaks is not dealt with in this article.

2 Application of the Financial Sector Regulation Act

The Financial Sector Regulation Act does not have a section defining the parameters of its own application. As mentioned, its scope and application are, however, far reaching. It regulates a mammoth regulatory regime. The preamble provides that Twin Peaks creates a system of financial regulation by establishing the Prudential Authority[12] and the Financial Sector Conduct Authority[13] and conferring powers on these entities. Its function, inter alia, is to preserve and enhance financial stability by conferring powers on the Reserve Bank.[14] The Act establishes a number of new councils, committees and groups and provides for the coordination, cooperation, collaboration and consultation of all of these and the authorities as well as all relevant organs of state.[15] It regulates and supervises financial product and service providers and aims to improve market conduct in order to protect financial customers.[16] It makes provision for the making of regulatory instruments.[17] It makes provision for the licensing of financial institutions,[18] and for powers to gather information, conduct supervisory on-site inspections and investigations, and enforce financial sector laws, including the imposition of administrative penalties. It provides for the protection and promotion of rights in the financial sector as set out in the Constitution. It establishes the Ombud Council[19] and confers powers on it in relation to ombud schemes,[20] including the handling of financial product and service providers.[21] It establishes the Financial Services Tribunal as an independent

[12] The Act deals with the Prudential Authority in Chapter 3, which chapter is divided into three parts: part 1 deals with its establishment, objectives and functions, part 2 deals with the governance of the Prudential Authority and part 3 deals with matters related to staff, resources and financial management. For a discussion of the powers and functions of the Prudential Authority, see Moorcroft and Vessio (n 1) par 39.2.2.

[13] The Act deals with the Financial Sector Conduct Authority in Chapter 4, which chapter is also divided into 3 parts: part 1 deals with its establishment, objectives and functions, part 2 deals with the governance of the Financial Sector Conduct Authority and part 3 deals with matters related to staff and resources. For a discussion of the powers and functions of the Financial Sector Conduct Authority see Moorcroft and Vessio (n 1) par 39.2.3.

[14] See the preamble and Chapter 2 of the Financial Sector Regulation Act which deals specifically with financial stability.

[15] Refer to Chapter 5 of the Financial Sector Regulation Act.

[16] A "financial customer" is a person to, or for, whom a financial product, a financial instrument, a financial service or a service provided by a market infrastructure is offered or provided, in whatever capacity, and includes a successor in title of the person and the beneficiary of the product, instrument or service (s 1 of the Financial Sector Regulation Act).

[17] A "regulatory instrument" incorporates a prudential standard, a conduct standard, a joint standard, an Ombud Council rule, a determination of fees in terms of s 237(1)*(a)*, an instrument identified as a regulatory instrument in a financial sector law, and an instrument amending or revoking any of them (s 1 of the Financial Sector Regulation Act). The making, publication and the consultative processes related thereto are dealt with in s 97–104 of the Financial Sector Regulation Act. Standards are dealt with in s 105–110 of the Financial Sector Regulation Act.

[18] Licensing is dealt with in Chapter 8, s 111–128 of the Financial Sector Regulation Act.

[19] Established by s 175 specifically but the powers, functions and functioning of the Ombud Council are dealt with in Chapter 14 parts 1 and 3 of the Financial Sector Regulation Act.

[20] See Chapter 14 parts 1 and 3 of the Financial Sector Regulation Act.

[21] See Chapter 14 parts 2 and 4 of the Financial Sector Regulation Act.

tribunal and confers on it powers to reconsider decisions by the financial sector regulators, the Ombud Council and certain market infrastructures.[22] It establishes the Financial Sector Information Register and makes provision for its operation.[23] It provides for information-sharing arrangements and creates offences.[24]

The Minister of Finance is responsible for the administration of the Act,[25] and the object of the Act is to achieve a stable financial system[26] that works in the interests of financial customers and that supports balanced and sustainable economic growth by establishing a regulatory and supervisory framework that promotes financial stability,[27] the safety and soundness of financial institutions, fair treatment and protection of financial customers, the efficiency and integrity of the financial system, the prevention of financial crime, financial inclusion, transformation of the financial sector, and confidence in the financial system.[28]

The Act lays out which authority is responsible for each financial sector law.[29] Here is a schematic representation of the financial sector laws that fall under the responsibility of the Financial Sector Conduct Authority and those that fall under the responsibility of the Prudential Authority:[30]

22 See Chapter 15 of the Financial Sector Regulation Act.

23 See Chapter 17 part 2 of the Financial Sector Regulation Act.

24 See Chapter 17 part 1 of the Financial Sector Regulation Act.

25 Section 8 of the Financial Sector Regulation Act.

26 "Financial system" is defined as "the system of institutions and markets through which financial products, financial instruments and financial services are provided and traded, and includes the operation of a market infrastructure and a payment system" (s 1 of the Financial Sector Regulation Act).

27 The Reserve Bank must, at least every six months, assess the stability of the financial system. This report will be known as the "financial stability review". The Act stipulates what a financial stability review must set out. Notably, any information which, if published, may materially increase the possibility of a systemic event, only needs to be published in a financial stability review after the risk of a systemic event subsides or has been addressed. The responsibilities of the Reserve Bank in light of this duty are delineated in the Act. The review must be tabled in Parliament (s 13 of the Financial Sector Regulation Act).

28 Section 7 of the Financial Sector Regulation Act. The financial sector laws are identified in schedule 1 of the Act. They are also listed below in Diagram 1.

29 Despite the allocations, a financial sector regulator may delegate its functions and powers in relation to a provision of a financial sector law for which it is the responsible authority to another financial sector regulator; the other financial sector regulator is, to the extent of the delegation, the responsible authority for the provision. This is carried out by a s 77 memorandum (s 5(2) of the Financial Sector Regulation Act).

30 The financial sector laws are listed in schedule 1 and the responsible authority for each law in schedule 2.

Diagram 1

FINANCIAL SECTOR CONDUCT AUTHORITY	PRUDENTIAL AUTHORITY
• Pension Funds Act • Friendly Societies Act • Financial Advisory and Intermediary Services Act • Collective Investment Schemes Control Act • Financial Markets Act • Credit Rating Services Act • Long-term Insurance Act • Short-term Insurance Act • A regulatory instrument made by the Financial Sector Conduct Authority • A joint standard, insofar as it relates to matters within the objectives of the Prudential Authority and the Financial Sector Conduct Authority	• Banks Act • Financial Supervision of the Road Accident Fund Act • Mutual Banks Act • Co-operative Banks Act • Insurance Act • A regulatory instrument made by the Prudential Authority • A joint standard, insofar as it relates to matters within the objectives of the Prudential Authority and the Financial Sector Conduct Authority

3 Application of the Consumer Protection Act

Subject to certain exclusions and exemptions, the Consumer Protection Act applies to every transaction occurring within South Africa.[31] It applies to the promotion of any goods or services, or of the supplier of any goods or services, unless those goods or services could not reasonably be the subject of a transaction to which the Consumer Protection Act applies or the promotion of those goods or services has been exempted in terms of the Act itself.[32] The Consumer Protection Act also applies to goods supplied or services performed in terms of a transaction to which the Act applies, irrespective of whether any of those goods or services are offered or supplied in conjunction with any other goods or services, or separate from any other goods or services and goods that are supplied in terms of a transaction that is exempt from the application of the Consumer Protection Act.[33]

If any goods are supplied within South Africa to any person in terms of a transaction that is exempt from the application of the Consumer Protection Act, those goods, and the importer or producer, distributor and retailer of those goods,

[31] Section 5(1)*(a)* of the Consumer Protection Act. For a detailed discussion on the application and scope of the Consumer Protection Act, see Naudé, Eiselen, De Stadler et al *Commentary on the Consumer Protection Act* (2018 *et seq*) Chapter 1 and van Eeden *Consumer Protection Law in South Africa* (2013) Chapter 3.

[32] Section 5(1)*(b)* of the Consumer Protection Act.

[33] Section 5(1)*(c)* and *(d)* of the Consumer Protection Act.

respectively, are nevertheless subject to the sections in the Act dealing with safety monitoring and recall, and liability caused by damaged goods.[34]

The following arrangements have been listed in the Consumer Protection Act as transactions that must be regarded as those between a supplier and consumer, within the meaning of the Act: the supply of any goods or services in the ordinary course of business to any of its members by a club, trade union, association, society or other collectivity, whether corporate or unincorporated, of persons voluntarily associated and organised for a common purpose or purposes, whether for fair value, consideration or otherwise, irrespective of whether there is a charge or economic contribution demanded or expected in order to become or remain a member of that entity; a solicitation of offers to enter into a franchise agreement; an offer by a potential franchisor to enter into a franchise agreement with a potential franchisee; a franchise agreement or an agreement supplementary to a franchise agreement; and the supply of any goods or services to a franchisee in terms of a franchise agreement.[35]

The application of the Consumer Protection Act extends to a matter irrespective of whether the supplier resides or has its principal office within or outside South Africa; operates on a for-profit basis or otherwise; or is an individual, juristic person, partnership, trust, organ of state, an entity owned or directed by an organ of state, a person contracted or licensed by an organ of state to offer or supply any goods or services, or is a public-private partnership or is required or licensed in terms of any public regulation to make the supply of the particular goods or services available to all or part of the public.[36]

The Act does not apply to the following: any transaction in terms of which goods or services are promoted or supplied to the State; any transaction in terms of which the consumer is a juristic person whose asset value or annual turnover, at the time of the transaction, equals or exceeds the threshold value determined by the Minister;[37] if the transaction falls within an exemption granted by the Minister;[38]

[34] Section 5(5) as read with s 60 and 61 of the Consumer Protection Act. See *Eskom Holdings v Halstead-Cleak* 2017 (1) SA 333 (SCA) for a discussion of s 61 of the Consumer Protection Act.

[35] Section 5(6) of the Consumer Protection Act. Despite the monetary exemption in sub-s (2), the Act applies to a transaction contemplated in sub-s (6)*(b)* to *(e)*, that is those sections dealing with franchising, irrespective of whether the size of the juristic person falls above or below the threshold determined (s 5(7) of the Consumer Protection Act).

[36] Section 5(8) of the Consumer Protection Act.

[37] In terms of s 6(1) of the Consumer Protection Act. On the early effective date as determined in accordance with item 2 of schedule 2, and subsequently at intervals of not more than five years, the Minister, by notice in the *Gazette*, must determine a monetary threshold applicable to the size of the juristic person for the purposes of s 5(2)*(b)*.

[38] A regulatory authority may apply to the Minister for an industry-wide exemption from one or more provisions of the Consumer Protection Act on the grounds that those provisions overlap or duplicate a regulatory scheme administered by that regulatory authority in terms of any other national legislation or any treaty, international law, convention or protocol. Upon advice of the Commission, the Minister, by notice in the *Gazette* after receiving the advice of the Commission, may grant an exemption only to the extent that the relevant regulatory scheme ensures the achievement of the purposes of the Act at least, as well as the provisions of the Act and subject to any limits or

if the transaction constitutes a credit agreement under the National Credit Act[39] (however, the goods or services that are the subject of the credit agreement are not excluded from the ambit of the Consumer Protection Act);[40] services to be supplied under an employment contract; giving effect to a collective bargaining agreement within the meaning of section 23 of the Constitution[41] and the Labour Relations Act;[42] or giving effect to a collective agreement as defined in section 213 of the Labour Relations Act.[43] It was provided in terms of section 66 of the Financial Services Laws General Amendment Act,[44] that the Consumer Protection Act would not apply to any function, act, transaction, good or service that is or are subject to Financial Services Board legislation.[45] However, in terms of section 290 read with schedule 4 of the Financial Sector Regulation Act, the greater part of the Financial Services Board Act has been repealed with effect from 1 April 2018, some sections with effect from 1 April 2020, whilst still others with effect from a date to be determined by the Minister.[46] Banks have been exempted from the provisions of section 14[47] of the Act.[48]

4 How Twin Peaks affects other legislation

The sections of the Financial Sector Regulation Act prevail in the event of any inconsistency between the Act and the provisions of another Act that is a financial sector law.[49] Regulations and regulatory instruments made under Twin Peaks do not override the financial sector law.[50] They do, however, override regulations and regulatory instruments made under other financial sector laws.[51] The application

conditions necessary to ensure the achievement of the purposes of the Act (s 5(3) and (4) of the Consumer Protection Act).

39 34 of 2005, hereinafter the "National Credit Act".

40 See also Melville and Palmer "The applicability of the Consumer Protection Act 2008 to credit agreements" 2010 *SA Merc LJ* 272–278; Otto "Verborge gebreke, voetstootsverkope, die Consumer Protection Act en die National Credit Act" 2011 *THRHR* 525, JM Otto and R-L Otto *The National Credit Act Explained* (2010) 135–137; Sharrock "Judicial control of unfair contract terms: the implications of the Consumer Protection Act" 2010 *SA Merc LJ* 295 at 304 and Moorcroft and Vessio (n 1) par 37.2.2.

41 Constitution of the Republic of South Africa, 1996.

42 66 of 1995.

43 Section 5 of the Consumer Protection Act.

44 45 of 2013.

45 See *Miya v Miway* NCT/43934/2016/75(1)(b). We encounter similar terminology (act, transaction, function, financial product, financial service) in s 10 of the Financial Sector Regulation Act, the section that deals with the interplay between the Financial Sector Regulation Act and the Consumer Protection Act. See par 5 below for a discussion.

46 See n 8 above for further detail.

47 Section 14 deals with the expiry and renewal of fixed-term contracts.

48 In terms of s 5(4) of the Consumer Protection Act (GN 532 in *GG* 34399 of 27 June 2011). The pension fund industry, the collective investments schemes industry and the security services industry were exempted in terms of s 5(4) from specified provisions of the Act for a period of 18 months in 2011 – see GN 533 in *GG* 34400 of 27 June 2011.

49 Section 9(1) of the Financial Sector Regulation Act.

50 Section 9(1) of the Financial Sector Regulation Act.

51 Section 9(2) of the Financial Sector Regulation Act.

of the Competition Act[52] to mergers is regulated by section 10(2) of the Financial Sector Regulation Act.

Legislation that is amended or repealed by Twin Peaks is set out in schedule 4.[53] It amends a number of Acts, including the South African Reserve Bank Act, the Banks Act,[54] the Mutual Banks Act,[55] the Co-operative Banks Act[56] and the National Credit Act. The Act has repealed four pieces of legislation. These are the Financial Services Board Act,[57] the Financial Services Ombud Schemes Act,[58] the Policy Board for Financial Services and Regulation Act[59] and the Inspection of Financial Institutions Act.[60]

5 How Twin Peaks affects the Consumer Protection Act

Section 10 of the Financial Sector Regulation Act reads as follows:

> "10 Application of other legislation
>
> (1) The Consumer Protection Act does not apply to, or in relation to—
>
> (*a*) a function, act, transaction, financial product or financial service that is subject to the National Payment System Act or a financial sector law, and which is regulated by the Financial Sector Conduct Authority in terms of a financial sector law; or
>
> (*b*) the Reserve Bank, the Prudential Authority, the Financial Sector Conduct Authority, the Prudential Committee, the Executive Committee, the Chief Executive Officer, the Commissioner or a Deputy Commissioner."

The intention behind section 10 is difficult to comprehend. Had the legislature wanted to oust the effects of the Consumer Protection Act from the financial sector[61] it could easily have done so. However, it did not exclude its application directly and yet the reach of its non-applicability, as will be seen, is very wide.

52 89 of 1998.
53 As read with s 290 of the Financial Sector Regulation Act. See n 8 above for further detail.
54 94 of 1990.
55 124 of 1993.
56 40 of 2007.
57 97 of 1990.
58 37 of 2004.
59 141 of 1993. See n 8 above for further detail.
60 80 of 1998.
61 "Financial sector" has not been defined by Twin Peaks, nor does it find definition in the South African Reserve Bank Act 90 of 1989 or the Banks Act 94 of 1990 for that matter, leaving one to understand that the field is wider than anticipated. This may very well be the conundrum and out-and-out ousting of the Consumer Protection Act would be too risky. The "financial system" is defined in the Twin Peaks, however – see n 26 above.

The words "function,"[62] "act" and "transaction"[63] are not defined in the Act, nor for that matter are they defined in the National Payment System Act.[64] "Financial sector law" includes the Financial Sector Regulation Act itself and any regulation or regulatory instrument made in terms of the Act or in terms of, and including, any laws listed in schedule 1 of the Act.[65] "Financial products"[66] include: participatory interests in collective investment schemes; long-term policies;[67] life insurance policies;[68] short-term policies;[69] non-life insurance policies;[70] benefits provided by pension fund organisations;[71] benefits provided by friendly societies;[72] deposits;[73] health service benefits provided by medical schemes;[74] the provision of credit provided in terms of credit agreements;[75] warranties, guarantees or other credit support arrangements as provided for in a financial sector law; facilities or arrangements designated by regulation as financial products; and ones that include one or more of the financial products referred to here.[76] The Regulations may designate as financial products any facility or arrangement that is not regulated in terms of a specific financial sector law if doing so would further the objects of the Act[77] and the facility or arrangement is one through which, or through the acquisition of which, a person conducts one or more of the following activities:

[62] "Function" is not defined in the Act; however, "control function" is defined. It is submitted that this later definition is presumably not applicable to this section as it refers to the risk management function, the compliance function, the internal audit function and the actuarial function (s 1 of the Financial Sector Regulation Act).

[63] "Transaction" has been defined in the Consumer Protection Act as
"*(a)* in respect of a person acting in the ordinary course of business—
(i) an agreement between or among that person and one or more other persons for the supply or potential supply of any goods or services in exchange for consideration; or
(ii) the supply by that person of any goods to or at the direction of a consumer for consideration; or
(iii) the performance by, or at the direction of, that person of any services for or at the direction of a consumer for consideration; or
(b) an interaction contemplated in section 5(6), irrespective of whether it falls within paragraph (a)".

[64] 78 of 1998, hereinafter the "National Payment System Act". For an introduction to the national payment system and the Act, see Moorcroft (n 1) Chapter 33.

[65] See n 4 and Diagram 1 above.

[66] The definition in s 1 redirects to s 2 of the Financial Sector Regulation Act.

[67] As defined in s 1(1) of the Long-term Insurance Act 52 of 1998. The relationship between the Long-term Insurance Act and the Financial Sector Regulation Act has been defined in s 1A of the Long-term Insurance Act.

[68] As defined in s 1(1) of the Short-term Insurance Act 18 of 2017.

[69] As defined in s 1 of the Insurance Act 53 of 1998. The relationship between the Short-term Insurance Act and the Financial Sector Regulation Act has been defined in s 1A of the Short-term Insurance Act.

[70] As defined in s 1 of the Insurance Act, to a member of the society by virtue of membership.

[71] As defined in s 1(1) of the Pension Funds Act 24 of 1956, to a member of the organisation by virtue of membership.

[72] As defined in s 1(1) of the Friendly Societies Act 25 of 1956.

[73] As defined in s 1(1) of the Banks Act 94 of 1990.

[74] As defined in s 1(1) of the Medical Schemes Act 131 or 1998.

[75] Regulated in terms of the National Credit Act except for the purposes of Chapter 4 and s 106.

[76] Section 2(1) of the Financial Sector Regulation Act.

[77] These are laid out in s 7 of the Act. See introductory par above.

lending, making a financial investment[78] and managing financial risk.[79] Regulations designating a financial product may specify the financial sector regulator that is the responsible authority for that product.[80]

"Financial services" include any activities conducted in South Africa in relation to financial products, foreign financial products, financial instruments,[81] or foreign financial instruments and include offering, promoting, marketing or distributing; providing advice, recommendations or guidance; operating or managing; providing administration services; dealing[82] or making a market[83] in South Africa in a financial product, foreign financial products, financial instruments[84] or foreign financial instruments; payment services; securities services; intermediary services;[85] services related to the buying and selling of foreign exchange; services related to the provision of credit, including a debt collection service but excluding the services of a debt counsellor;[86] payment distribution agents or alternative dispute resolution agents;[87] services provided to financial institutions through outsourcing arrangements; and any other services provided by financial institutions, being services regulated by a specific financial sector law and services designated by the

78 A "financial investment" is made when the "investor" gives a contribution, in money or money's worth, to another person and any of the following apply: the other person uses the contribution to generate a financial return for the investor; the investor intends that the other person will use the contribution to generate a financial return for the investor, even if no return, or a loss, is in fact generated; the other person intends that the contribution be used to generate a financial return for the investor, even if no return, or a loss, is in fact generated and has no day-to-day control over the use of the contribution (s 2(3) of the Financial Sector Regulation Act).

79 A person is considered to be "managing a financial risk" when such person manages the financial consequences to the person of particular events or circumstances occurring or not occurring or avoids or limits the financial consequences of fluctuations in, or in the value of, receipts or costs, including prices and interest rates (s 2(4) of the Financial Sector Regulation Act).

80 Section 2(5) of the Financial Sector Regulation Act.

81 "Financial instruments" mean shares as defined in s 1 of the Companies Act; a depository receipt and other equivalent instruments; a debt instrument such as a debenture or a bond, but not a credit agreement; money market securities or a derivative instrument as defined in s 1(1) of the Financial Markets Act 19 of 2012; or a warrant, certificate, securitisation instrument or other instrument acknowledging, conferring or creating rights to subscribe to, acquire, dispose of or convert any of these (s 1 of the Financial Sector Regulation Act).

82 For the purposes of "financial services", "dealing", whether done as a principal or as an agent, means: in relation to securities or participatory interests in a collective investment scheme, underwriting the securities or interests and the buying or selling of the securities or interests for their own account or on behalf of another person as a business, a part of a business or incidental to conducting a business (s 3(4) of the Financial Sector Regulation Act).

83 "Making a market" in a financial instrument takes place when a person, through a facility, at a place or otherwise, states the prices at which the person offers to acquire or dispose of financial instruments, whether or not on the person's own account, and other persons reasonably expect that they can enter into transactions for those instruments at those prices (s 3(4) of the Financial Sector Regulation Act).

84 See n 81 above.

85 As defined in s 1(1) of the Financial Advisory and Intermediary Services Act 37 of 2002.

86 Registered in terms of s 44 of the National Credit Act, who provides the services of a debt counsellor as contemplated in that Act.

87 As defined in s 1 of the National Credit Act.

Regulations as financial services.[88] Services provided by market infrastructures[89] are not financial services unless so designated by regulation and only if doing so would further the objects of the Act.[90] The Regulations may designate as a financial service any service that is not regulated in terms of a specific financial sector law if the service, provided in South Africa, relates to financial products, foreign financial products, financial instruments or foreign financial instruments; arrangements that are in substance arrangements for lending, making financial investments or managing financial risks[91] or the provision of benchmarks or indices; or services provided by market infrastructures.[92] Regulations designating a financial service may specify the financial sector regulator that is the responsible authority for that service.[93]

If one reads section 10(1)*(a)* in isolation, it would be reasonable to conclude that to determine whether an act, transaction, function, financial product or financial service falls under the auspices of the Consumer Protection Act, one would simply have to determine whether it falls under any of the financial sector laws and, thereafter, whether such law falls under the responsibility of the Financial Sector Conduct Authority. However, once one moves to section 10(1)*(b)*, it appears that the effort by the legislature to emphasise the separation of an act, transaction, function, financial product or financial service that falls under the auspices of a financial sector law *and* the Financial Sector Conduct Authority and those that do not, is merely redundant. This is because in the next section, Twin Peaks prevents the application of the Consumer Protection Act, inter alia, to anything relating to the Prudential Authority and to the Financial Sector Conduct Authority. One is wont to assume that this would include any acts, functions, transactions, financial products and financial services that would fall under a financial sector law as listed in schedule 2 and fall, too, under the responsibility of the Prudential Authority and any acts, transactions, functions, financial products and financial services that may or may not fall under a financial sector law but are otherwise "related to" the Prudential Authority or the Financial Conduct Sector Authority. Alternatively, the intention behind section 10(1)*(b)* and its separation from section 10(1)*(a)* is to limit the applicability of the Consumer Protection Act to the acts, transactions, functions, financial products and financial services that fall under the financial sector laws that are under the responsibility of the Prudential Authority. The drafting separation is either artificial or maladroit.

If the Prudential Authority had not been mentioned in section 10(1)*(b)* one could surmise, for example, that any of the following insurance products would not have the Consumer Protection Act applicable to them: engineering, guarantee, liability,

88 Section 3(1) of the Financial Sector Regulation Act.

89 Market infrastructure means each of the following, as defined in s 1(1) of the Financial Markets Act 19 of 2012: a central counterparty, a central securities depository, a clearing house, an exchange and a trade repository.

90 See s 7 of the Financial Sector Regulation Act.

91 As contemplated in s 2(2)–(4) of the Financial Sector Regulation Act.

92 Section 3(2) and (3) of the Financial Sector Regulation Act.

93 Section 3(5) of the Financial Sector Regulation Act.

motor, accident and health, property, transportation, an assistance, a disability, fund, life or sinking fund policy, or a contract comprising a combination of any of those policies as they fall under the auspices of a financial sector law that falls under the auspices of the Financial Sector Conduct Authority. Life insurance and non-life insurance policies would attract the applicability of the Consumer Protection Act.[94] But then, and because of section 10(1)*(b)*, it would seem that all these financial goods are exempted from the application of the Consumer Protection Act. To find another interpretation – that is, that any acts, functions, transactions, financial products or financial services that fall under a financial sector law falling under the responsibility of the Prudential Authority would have the Consumer Protection Act apply – is possible but very questionable. These financial sector goods are very much related to the Prudential Authority and would therefore be ousted from the application of the Consumer Protection Act by virtue of section 10(1)*(b)*.[95] It remains at this point, at least from the writer's perspective, inexplicable as to why the legislature made this distinction.

What confounds the matter further is that the Financial Sector Regulation Act allows for much fluidity in its definitions; thus, where a product may not have been a financial product as defined, a later regulation may designate as a financial product any facility or arrangement that is not regulated in terms of a specific financial sector law.[96] Similarly, the Regulations may designate as a financial service any service that is not regulated in terms of a specific financial sector law.[97] Thus, a financial product or a financial service that would otherwise fall under the auspices of the Consumer Protection Act may at any time be removed therefrom.

Further complications are created by virtue of the fact that a financial sector regulator may delegate its functions and powers in relation to a provision of a financial sector law for which it is the responsible authority to another financial sector regulator. The other financial sector regulator would then, to the extent of the delegation, be the responsible authority for the provision.[98] With the second interpretation supplied above, it would mean that upon a delegation from one authority to the other, the Consumer Protection Act may or may not apply. This would be an exercise in absurdity and such a system would certainly denigrate the legislation.

[94] Although there would in any event have been a logistical issue here, but for s 10(1)*(b)*, as the definition of a short-term policy as found in s 1 of the Short-term Insurance Act includes the definition of a non-life insurance policy as defined in s 1 of the Insurance Act, and the definition of a long-term policy as defined in s 1 of the Long-term Insurance Act includes the definition of a life insurance policy as defined in s 1 of the Insurance Act. The Short-term Insurance Act and the Long-term Insurance Act are financial sector laws that fall under the responsibility of the Financial Sector Conduct Authority, while the Insurance Act is a sector law that is monitored by the Prudential Authority.

[95] Such an interpretation would in any event be further complicated by the delegatory power of these responsibilities (s 5(2) of the Financial Sector Regulation Act).

[96] Section 2(2) of the Financial Sector Regulation Act.

[97] Section 3(3) of the Financial Sector Regulation Act.

[98] See n 29 above.

Section 10(1)*(a)* has also complicated the relationship between the National Credit Act and the Consumer Protection Act because there is interplay between the National Payment System Act[99] and the National Credit Act. Besides the fact that the National Credit Act amended the National Payment System Act by inserting section 6A into that Act,[100] once an act, transaction, function, financial product or financial service relating to a credit agreement,[101] at any point in its lifecycle, becomes subject to the National Payment System Act, it will not have the Consumer Protection Act apply to the goods and services that are the subject of that credit agreement. This interpretation would mean that goods and services that are the subject of credit agreement would be regulated by the Consumer Protection Act only up to a certain point and thereafter, suddenly, not. This was surely not the legislative intention. Whatever the outcome, it places the practical and interpretative burdens on practitioners and the courts alike, with the costs landing squarely on the shoulders of the consumer.

Another question which remains unanswered is the future of section 5(5) of the Consumer Protection Act to the financial industry. This section states that if any goods are supplied within South Africa to any person in terms of a transaction that is exempt from the application of the Consumer Protection Act, those goods, as well as the importer or producer, distributor and retailer of those goods, respectively, will nevertheless be subject to the sections in the Act dealing with safety monitoring and recall, and liability caused by damaged goods. Are sections 60 and 61 as read with section 5(5) of the Consumer Protection Act superseded by section 10(1) or not? It is submitted that section 10(1) has not created an exemption but a legislative imperative and that the, perhaps unintended, consequence has been to interrupt the applicability of the strict liability sections to the financial sector.[102]

It is difficult to fathom a simple solution. A possible way of dealing with the interaction of the two Acts would have been to oust the application of the Consumer Protection Act from the financial sector arena completely. The risk with this option would, however, have fallen too heavily on the unwitting consumer, especially given that the impact and interpretation of the overhaul of the financial sector laws

[99] 78 of 1998.

[100] In terms of s 6A of the National Payment System Act, a person cannot change, manipulate, maintain or apply a payment system in any manner that provides preferential treatment to a payment instruction over any other payment instruction in that system, unless such preferential treatment is prescribed by law. For a discussion, see Van Zyl in Scholtz et al *Guide to the National Credit Act* (2018 *et seq*) par 18.3.

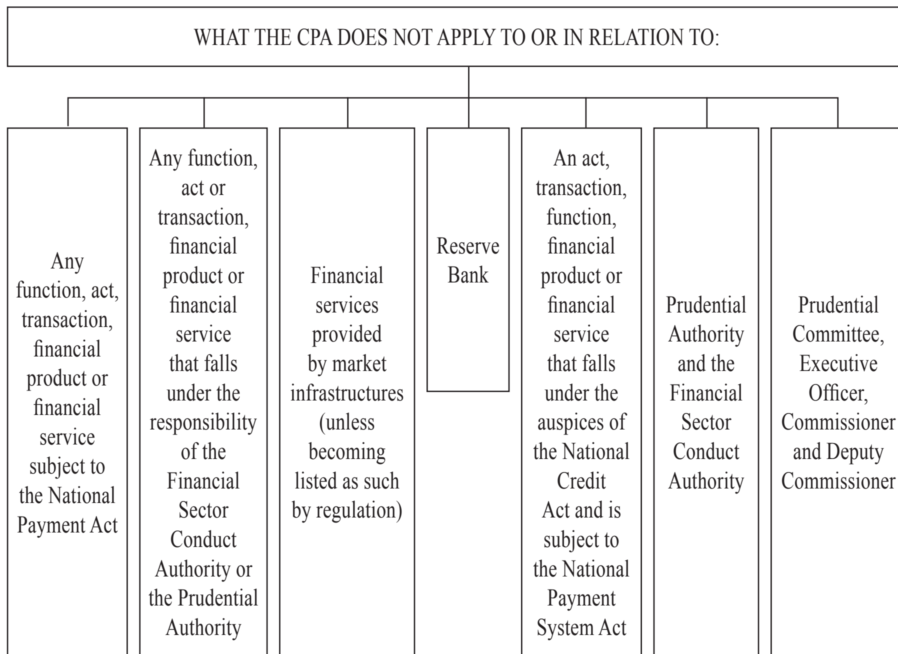
[101] None of these terms are defined in the National Credit Act, except a credit transaction, which is an agreement, irrespective of its form, that constitutes a credit transaction if it is a pawn transaction or discount transaction; an incidental credit agreement; an instalment agreement; a mortgage agreement or secured loan; a lease; or any other agreement, other than a credit facility or credit guarantee, in terms of which payment of an amount owed by one person to another is deferred, and any charge, fee or interest is payable to the credit provider in respect of the agreement or the amount that has been deferred (s 8(4) of the National Credit Act). The effects of s 10(1)*(b)* of Twin Peaks on the relationship between the National Credit Act and the Consumer Protection Act cannot be explored here as it would render this discussion, which is of broader import, too prolix.

[102] To move the discussion much further would require another couple of pages of discussion and thus cannot be tackled in this chapter.

is at this point largely unknown. Another possible solution would have been to adopt a so-called "safety-net" section, that is, a provision stating that in the event of an inconsistency between a financial sector law and the Consumer Protection Act, that the relevant provision of the Act that offers more or better protection for the consumer supersedes the other. The way it reads now, it seems that all sections 10(1)*(a)* and 10(1)*(b)* have done is simply to expose financial consumers.

The following is a schematic representation of the ousting of the products and services and institutions from the application of the Consumer Protection Act by the Financial Sector Regulation Act:

Diagram 2



6 Conclusion

The Financial Sector Regulation Act is regulating a conglomeration of financial sectors, the girth of which is by no means insignificant. It is not surprising therefore that its interaction with any piece of legislation is going to be complex, not least of all when another composite Act like the Consumer Protection Act is involved. However, sometimes simple and pragmatic is better. The intention behind section 10, the section in the Financial Sector Regulation Act dealing with the application of the Consumer Protection Act to the financial sector, is somewhat oblique. The legislative approach has been to make the Consumer Protection Act applicable to all acts, transactions, functions and financial products and services which are not regulated by the National Payment System Act or any of the financial sector laws that are in themselves subject to the Financial Sector Conduct Authority. However, and despite the specificity of section 10(1)*(a)*, 10(1)*(b)* exempts too any matter

related to the Reserve Bank, the Prudential Authority, the Financial Sector Conduct Authority, the Prudential Committee, the Executive Committee, the Chief Executive Officer, the Commissioner or a Deputy Commissioner from the application of the Consumer Protection Act, ultimately making the division between section 10(1)*(a)* and *(b)* redundant.

The National Credit Act is not listed as a financial sector law; however, there is interplay between it and the National Payment System Act and it is therefore not immune to the operations of section 10 of Twin Peaks. The current wording of section 10(1)*(a)* (whether wittingly or unwittingly) appears to oust the application of the Consumer Protection Act to goods and services that are the subject of a credit agreement once any part of the agreement enters the national payment system.

The question remains open as to whether the strict liability and monitoring and recall sections of the Consumer Protection Act would still apply to the supply chain when dealing with the financial sector, and it is unlikely that section 10(1) can be labelled as an exemption. The inclination is therefore to reason that sections 60 and 61 of the Consumer Protection Act no longer apply to the financial sector.

Further difficulties are created by the facility with which products and services may be absorbed under the financial products and financial services definitions, creating uncertainty as to whether the Consumer Protection Act applies or not.

It has been suggested that ousting the applicability of the Consumer Protection Act to the financial sector would leave consumers to risk exposure. The suggestion has also been made that the legislature might have incorporated a clause which gives precedent to the section of the Act, whether a financial sector law or the Consumer Protection Act, that gives more or better protection to the consumer.

The applicability of section 10(1) is not straightforward and practitioners will want to look very closely at the acts, transactions, functions, financial products and financial services which are being affected, and, if litigating, it would be wise to claim in the alternative.

Fair play? The Conduct of Financial Institutions Bill and the new face of the financial services industry

DALEEN MILLARD*

1 Introduction

National Treasury released the draft Conduct of Financial Institutions Bill (COFI Bill) on 11 December 2018. The proposed Bill was open for public comment until 1 April 2019 and is anticipated to be tabled in Parliament during 2019. The Bill represents the next step of legislation aimed at reforming the financial sector, the first step being the Financial Sector Regulation Act[1] (FSRA), which gave effect to the Twin Peaks model that established the Prudential Authority (PA) and the Financial Sector Conduct Authority (FSCA) in 2018.[2] The FSRA defines the roles of these two regulators, and the COFI Bill is focused on the conduct of financial institutions. It is expected to replace the conduct provisions of most existing financial sector laws in an effort to streamline the market conduct framework for all financial sector institutions.[3]

The Bill consists of 118 clauses.[4] It has so far elicited responses ranging from it being a "refreshing and mature piece of legislation"[5] which supports transformation, to a statute that will disempower institutions such as the Council

* Professor of Private law, University of Johannesburg.

1 9 of 2017.

2 Millard "CoFI and T(CF): Further along the road to Twin Peaks and a fair insurance industry" 2018 *THRHR* 374.

3 The objectives of the FSRA are set out in s 57. Thus, the statute aims to "*(a)* enhance and support the efficiency and integrity of financial markets; and
  (b) protect financial customers by—
    (i) promoting fair treatment of financial customers by financial institutions; and
    (ii) providing financial customers and potential financial customers with financial education programmes, and otherwise promoting financial literacy and the ability of financial customers and potential financial customers to make sound financial decisions; and
  (c) assist in maintaining financial stability".

4 The Bill consists of several chapters, namely, Chapter 1, Interpretation, Objects and Application; Chapter 2, Licensing; Chapter 3, Culture and Governance; Chapter 4, Financial Products; Chapter 5, Financial Services; Chapter 6, Promotion, Marketing and Disclosure; Chapter 7, Distribution, Advice and Discretionary Investment Management; Chapter 8, Post-Sale Barriers and Obligations; Chapter 9, Safeguarding Assets and Operational Requirements; Chapter 10, Reporting; Chapter 11, Remedial Actions for Financial Customers; Chapter 12, General Provisions; and Chapter 13, Final Provisions.

5 Vawda "Wake up and smell the black COFI" *Daily Maverick* (15 January 2019) (www.dailymaverick.co.za (accessed 3 August 2019)).

ABLU 2019 (Part ONE).indb 129 2019/10/09 9:49 AM

on Medical Schemes.[6] The purpose of this article is not to provide an analysis of the entire Bill. Rather, it is to evaluate those features that aim to transform the financial services industry, to introduce a measure of fairness and to address those shortcomings that have remained a thorn in the side of South African consumers, such as the cost of financial products, the inability to understand product features due to the use of complex language and the frustrations that accompany the search for redress where harm was suffered. The discussion will include an overview of those features that build on the foundations laid down by the Financial Advisory and Intermediary Services Act[7] (FAIS Act) and, in the final instance, ventures an opinion as to the suitability of the Bill to effectively address persistent, systemic failures of the financial services industry *vis*-á-vis consumers.

However, before the (rather ambitious) objectives of the Bill are discussed, the next part provides a cursory overview of the key definitions in the Bill.

2 Conceptual framework

One should accept that the very first, coordinated approach to market conduct regulation was enabled by the FAIS Act.[8] In terms of this statute, the scope and application of protection of consumers revolved mainly around the definitions of "advice", "intermediary services" and "financial product". The statute currently defines "advice" as any recommendation, guidance or proposal of a financial nature furnished by any means or medium, to any client or group of clients.[9] The advice must pertain to the purchase of any financial product or the investment in any financial product.

The definition of a financial product is rather extensive, confirming that the Act aims to include as many products as possible under the FAIS Act in order to extend protection to consumers. This is a central theme upon which the COFI Bill builds.[10] It includes securities and instruments;[11] a participatory interest in one or more collective investment schemes; insurance policies; benefits provided by pension funds and friendly societies; a foreign currency denominated investment instrument, including a foreign currency deposit; a deposit as defined in section 1(1) of the Banks Act;[12] a health service benefit provided by a medical scheme as

[6] Knoesen "COFI Bill … disappointed, irritated and concerned" *FA News* (3 April 2019) (www.fanews.co.za (accessed 3 August 2019)).

[7] 37 of 2002.

[8] *Ibid*.

[9] Section 1(1)*(a)* of the FAIS Act *sv* "advice". See Moolman, Pillai, Bam and Appasamy *Financial Advisory and Intermediary Services Guide* (2010) 25–26, 197.

[10] See Reinecke, Van Niekerk and Nienaber *South African Insurance Law* (2013) 509–512. See also Millard and Hattingh *The FAIS Act Explained* (2016) 33.

[11] These include shares in a company other than a "share block company" as defined in the Share Blocks Control Act 59 of 1980; debentures and securitised debt; any money market instrument; any warrant, certificate, and other instrument acknowledging, conferring or creating rights to subscribe to, acquire, dispose of or convert securities and instruments; and any "securities" as defined in s 1 of the Financial Markets Act 19 of 2012.

[12] 94 of 1990.

defined in section 1(1) of the Medical Schemes Act;[13] and any other product similar in nature to any of these financial products.

"Advice", as the second key definition, includes any recommendation, guidance or proposal of a financial nature "on the conclusion of any other transaction, including a loan or cession, aimed at the incurring of any liability or the acquisition of any right or benefit in respect of any financial product".[14]

The third key definition in the FAIS Act is "intermediary service", which means any act other than the furnishing of advice that is performed by a person for or on behalf of a client or product supplier that results in the client entering into or offering to enter into any transaction in respect of a financial product with a product supplier. Even if the client may in future enter into any such transaction, it means that an intermediary service had been rendered.[15] It can also mean any act other than the furnishing of advice that is performed by a person for or on behalf of a client or product supplier with a view to one or more of the following, namely, buying, selling or otherwise dealing in, managing, administering, keeping in safe custody, or maintaining or servicing a financial product purchased by a client from a product supplier or in which the client has invested;[16] collecting or accounting for premiums or other monies payable by the client to a product supplier in respect of a financial product;[17] or receiving, submitting or processing the claims of a client against a product supplier.[18]

As the COFI Bill is set to replace the FAIS Act, it is imperative to compare the conceptual frameworks in terms of these two instruments to form an idea of the thinking that informs the future of market conduct in South Africa. The COFI Bill sets out an extensive list of definitions. What is interesting is that "advice" is not defined. Rather, Chapter 7 of the Bill contains provisions on distribution, advice and discretionary investment management. Similarly, the term "intermediary services" is not defined, but the Bill provides a description of those activities by intermediaries that need to be regulated. Furthermore, the term "financial products" has the meaning assigned to it in section 2 of the Financial Sector Regulation Act, and includes a foreign financial product. This section accordingly provides that "financial product" means the following:

> "*(a)* a participatory interest in a collective investment scheme;
> *(b)* a long-term policy as defined in section 1(1) of the Long-Term Insurance Act;
> *(c)* a short-term policy as defined in section 1(1) of the Short-term Insurance Act;
> *(d)* a benefit provided by—
> (i) a pension fund organisation, as defined in section 1(1) of the Pension Funds Act, to a member of the organisation by virtue of membership; or
> (ii) a friendly society, as defined in section 1(1) of the Friendly Societies Act, 5 to a member of the society by virtue of membership;

[13] 131 of 1998.
[14] Millard and Hattingh (n 10) 12.
[15] Section 1(1)*(a)* of the FAIS Act. See also Millard and Hatting (n 10) 39; Moolman, Pillai, Bam and Appasamy (n 9) 26–27, 200.
[16] Section 1(1)*(b)*(i) of the FAIS Act. See also Millard and Hatting (n 10) 39.
[17] Section 1(1)*(b)*(ii).
[18] Section 1(1)*(b)*(iii).

(e) a deposit as defined in section 1(1) of the Banks Act;

(f) a health service benefit provided by a medical scheme as defined in section 1(1) of the Medical Schemes Act;

(g) except for the purposes of chapter 4 and section 106, the provision of credit provided in terms of a credit agreement regulated in terms of the National Credit Act;

(h) a warranty, guarantee or other credit support arrangement as provided for in a financial sector law;

(i) a facility or arrangement designated by Regulations for this section as a financial product; and

(j) a facility or arrangement that includes one or more of the financial products referred to in paragraphs *(a)* to *(i)*."

The FSRA also provides a definition for "financial instrument", which reads as follows:

"*(a)* a share as defined in section 1 of the Companies Act;

(b) a depository receipt and other equivalent instruments;

(c) a debt instrument such as a debenture or a bond, but not a credit agreement;

(d) money market securities as defined in section 1(1) of the Financial Markets Act;

(e) a derivative instrument as defined in section 1(1) of the Financial Markets Act; or

(f) a warrant, certificate, securitisation instrument or other instrument acknowledging, conferring or creating rights to subscribe to, acquire, dispose of, or convert, the financial instruments referred to in paragraphs *(a)* to *(e)*."

The definitions of "financial product" and "financial instrument" must be read together with "financial service", which means:

"*(a)* any of the following activities conducted in the Republic in relation to a financial product, a foreign financial product, a financial instrument, or a foreign financial instrument:

(i) offering, promoting, marketing or distributing;

(ii) providing advice, recommendations or guidance;

(iii) operating or managing;

(iv) providing administration services;

(b) dealing or making a market in the Republic in a financial product, a foreign financial product, a financial instrument or a foreign financial instrument;

(c) a payment service;

(d) securities services;

(e) an intermediary service as defined in section 1(1) of the Financial Advisory and Intermediary Services Act;

(f) a service related to the buying and selling of foreign exchange;

(g) a service related to the provision of credit, including a debt collection service, but excluding the services of—

(i) a debt counsellor registered in terms of section 44 of the National Credit Act who provides the services of a debt counsellor as contemplated in that Act;

(ii) a payment distribution agent as defined in section 1 of the National Credit Act; or

(iii) an alternative dispute resolution agent, as defined in section 1 of the National Credit Act;

(h) a service provided to a financial institution through an outsourcing arrangement;

(i) any other service provided by a financial institution, being a service regulated by a specific financial sector law; and

(j) a service designated by the Regulations for this section as a financial service."[19]

Although the possibility exists that the definition of financial product may change, it is patently clear that the pattern remains the same as that under the FAIS Act, namely, that products are to be defined in order to establish the scope of the regulator's authority. In addition, products may further be deemed to be financial products if the aim is to provide further protection to consumers. In this respect, COFI clearly aims to ensure that everything connected with financial services, products and instruments is regulated.

Part of the regulation of the wide range of products includes "conduct standard", which is defined as "a conduct standard prescribed by the Authority as contemplated in section 106 of the Financial Sector Regulation Act and section 107(1)*(a)* of this Act". Such a standard refers to aspects such as product features and minimum standards that are expected of a financial product.

The COFI Bill also includes a definition of "fit and proper requirements". In terms of the Bill, "fit and proper requirements" means the following:

"*(a)* in relation to a person, requirements relating to—
- (i) honesty and integrity;
- (ii) good standing;
- (iii) competence, including—
  - *(aa)* experience;
  - *(bb)* qualifications;
  - *(cc)* knowledge of financial products, financial instruments, foreign financial products, and financial services;
  - *(dd)* knowledge tested through examinations;
  - *(ee)* continuous professional development; and
  - *(ff)* professional designation or membership;

[19] Section 3(2) further stipulates: "[A] service provided by a market infrastructure is not a financial service unless designated by Regulations in terms of subsection (3).

(3) If doing so will further the object of this Act set out in section 7, the Regulations may designate as a financial service—

(a) any service that is not regulated in terms of a specific financial sector law if the service, that is provided in the Republic, relates to—
- (i) a financial product, a foreign financial product, a financial instrument or a foreign financial instrument;
- (ii) an arrangement that is in substance an arrangement for lending, making a financial investment or managing financial risk, all as contemplated in section 2(2) to (4); or
- (iii) the provision of a benchmark or index; or

(b) a service provided by a market infrastructure."

> *(b)* in relation to a significant owner, requirements relating to—
> (i) honesty and integrity;
> (ii) good standing; and
> (iii) financial standing."

Again, there is a continuation on the theme that was introduced by the FAIS Act, albeit with more detail.[20]

New definitions include that of a "potential financial customer"[21] and a "retail financial customer".[22] Of particular interest is the inclusion of a definition of "transformation of the financial sector". In addition to it being an objective of the COFI Bill, it also means "transformation as envisaged by the Financial Sector Code for Broad-Based Black Economic Empowerment issued in terms of section 9(1) of the Broad-Based Black Economic Empowerment Act, 2003".[23]

A cursory overview of the conceptual framework demonstrates that a concerted effort has been made to ensure that all financial products, instruments and services are regulated. The next part provides an overview of the objectives of the Bill.

3 Objectives of the Bill

3.1 *Introduction*

The short title of the Bill reveals several objectives.[24] Before these can be discussed, it is necessary to remind ourselves of the objectives of the FAIS Act. This statute aimed to "regulate the rendering of certain financial advisory and intermediary services to clients, to repeal or amend certain laws, and to provide for matters incidental thereto". Unassuming as it appears, this statute brought about a revolution in the financial services industry by professionalising financial services and stipulating that advisors and intermediaries needed to comply with the fit and proper requirements, which included having (or obtaining, by a specified date) appropriate qualifications.[25] The FAIS Act, therefore, ensured that financial services were professionalised, the idea being that properly trained service providers are less likely to abuse clients.

At the dawn of the COFI era, and in line with the Twin Peaks model of regulation, the object of the Bill is to "establish a consolidated, comprehensive and consistent

[20] Millard and Hattingh (n 10) 62; Moolman, Pillai, Bam and Appasamy (n 9) 96, 166.

[21] This refers to a person who has "*(a)* applied to or otherwise approached a financial institution or intermediary to become a financial customer; *(b)* been solicited by a financial institution to become a financial customer; or *(c)* received advertising in relation to any financial product or financial service".

[22] This definition denotes a financial customer that is "*(a)* a natural person; or *(b)* a juristic person, whose asset value or annual turnover is less that the threshold value as determined by the Minister after consideration of any similar threshold values determined under the Consumer Protection Act 68 of 2008".

[23] Act 53 of 2003. This aspect is further discussed in par 3.6 below.

[24] The overriding objective is to provide for the establishment of a consolidated, comprehensive and consistent regulatory framework for the conduct of financial institutions. Par 3 below provides further detail on the remaining objectives.

[25] Millard and Hattingh (n 10) 102; Moolman, Pillai, Bam and Appasamy (n 9) 32, 39–41.

regulatory framework for the conduct of financial institutions" with a number of very specific aims, and the FSCA must, in performing their functions, promote the object of COFI. It is thought that an improved legal and regulatory framework can also support broader objectives in the financial sector, which includes growing a more transformed and inclusive industry. To best understand the thinking that informed the drafting of the COFI Bill, it is informative to read the Explanatory Policy Paper.[26]

According to the Policy Paper, the improved regulatory environment is expected to better support the entry of new institutions into the market, while at the same time facilitating the growth and development of existing institutions.[27] This, in turn, supports the transformative effect of the financial sector in the lives of South African customers by providing greater access to appropriate and suitable financial products and services to more South Africans.[28] This is achieved by protecting and promoting the fair treatment of customers, as well as promoting transformation, financial inclusion and innovation in the sector. The Policy Paper further explains that the law will apply across the financial sector to ensure a level playing field and to "reduce the risk of regulatory arbitrage".[29] In addition, it states that the FSCA will be able to appropriately accommodate smaller players in the financial sector (for example, burial societies and *stokvels*) which "pose low conduct risk and should not be subject to onerous regulatory compliance costs".[30] The Policy Paper elaborates that, subject to certain criteria, such entities should be exempted or excluded and the ability to exempt certain financial institutions from provisions, including "for developmental, transformation, and inclusion purposes" will enable the regulator to support these outcomes in the financial sector.[31]

Against this background, the main objectives of the Bill are now discussed.

3.2 *Protection of financial customers*

According to the Policy Paper, the protection of customers in the financial sector, and meaningful financial inclusion in South Africa, are mutually reinforcing objectives.[32] This is why it is foreseen that higher standards of customer protection "can drive greater inclusion as customers feel more secure in their participation in the financial sector".[33] The integration between market conduct and prudential

[26] National Treasury *Explanatory Policy Paper Accompanying the Conduct of Financial Institutions Bill* (2018) (Policy Paper) (www.treasury.gov.za (accessed 2 August 2019)).

[27] *Ibid* 7.

[28] *Ibid*.

[29] *Ibid* 16.

[30] *Ibid*.

[31] *Ibid.* The Policy Paper (n 26) 16 further explains that the FSCA will be required to "set standards, develop and implement its supervisory approach, and enforce requirements, in a manner that is proportionate to the nature, size, scale and complexity of the risks associated with a type of activity or financial institution, and is proportionate to achieving the purpose of the requirement".

[32] Policy Paper (n 26) 3.

[33] *Ibid.* The Policy Paper (n 26) 6 provides examples of market abuse that is behind the drive for increased protection and states as follows: "There have been a number of high profile cases of poor practices in the financial sector. In 2005, for example, after extensive engagement, a Statement of

regulation, therefore, is thought to bring about a better, more consolidated approach by closing the gaps between product standards, on the one hand, and how these products are sold to the public, on the other.[34] Notably, the COFI Bill aims to "significantly streamline the legal landscape for conduct regulation in the financial sector"[35] by having a "single comprehensive market conduct law"[36] to ensure the consistent application of consumer protection principles across the financial services sector, while at the same time providing "more flexibility and better tools to the regulator to support emerging new financial institutions, including black-owned businesses and non-traditional financial institutions (*eg* financial services provided by technology companies through digital innovation)".[37] As it will apply to all financial institutions in the sector, the COFI Bill is also best placed to give legal effect to transformation requirements in support of targets agreed through the Financial Sector Transformation Council and specified in the Financial Sector Code.

As the FSCA has scope of jurisdiction over all financial institutions in South Africa, industries such as the credit sector, the national payments system, the financial markets, pension funds, collective investment schemes and medical schemes also form part of the landscape, and the controversy surrounding inclusion under the COFI Bill should not be underestimated. Financial products are very diverse and, for instance, there is a significant difference between consumer credit and membership of a pension fund. It therefore is no surprise that the Policy Paper focuses on a number of these sectors and explains exactly what is envisaged by the COFI Bill.

For the *credit sector*, the National Credit Act[38] (NCA) already provides standards for credit products. However, it is envisaged that the National Credit Regulator will cooperate with the FSCA to ensure that there is no duplication and that the latter will regulate credit providers on matters pertaining to marketing and promotion, the provision of advice, and the distribution and disclosure of information relating to credit. Any new conduct standards set by the FSCA pertaining to these aspects must take into account requirements already in place under the NCA.[39] Payment systems are equally complex as the total payment process, from payer to beneficiary, enables transacting parties to efficiently exchange value to conclude financial transactions.[40] Although payment systems are very diverse in nature, COFI aims to

Intent was signed between the Minister of Finance and the long-term insurance industry, committing to reducing early termination penalty fees on savings policies. This was subsequently entrenched in legislation. In 2006, the Competition Commission began its investigation into the retail banking sector, with its findings being published in 2008, noting a number of poor customer outcomes in the sector. In the investment management sector, the Fidentia and Sharemax scandals, in 2007 and 2010 respectively, highlighted the need for better protection for investors."

[34] Policy Paper (n 26) 7.

[35] *Ibid* 14.

[36] *Ibid.*

[37] *Ibid*.

[38] 34 of 2005.

[39] Policy Paper (n 26) 22.

[40] *Ibid* 23.

license payment service providers who have a direct relationship with a financial customer, both retail and non-retail, and aims to set standards on aspects such as fees charged.[41]

For *financial markets*, the FSCA already licenses market infrastructure in terms of the Financial Markets Act[42] (FMA), and the buying and selling of foreign exchange and services related thereto are expected to be subjected to further regulation in terms of COFI.[43] As far as the collective schemes are concerned, COFI aims to repeal the Collective Investments Scheme Control Act[44] (CISCA) so as to in future cover a potentially wider range of investment product providers than those currently regulated by CISCA.[45]

For *pension funds*, it is proposed that retirement funds will be required to be licensed under both the Pension Funds Act[46] and COFI to ensure consistency in the treatment of members of retirement funds.[47]

In the final instance, the FSCA's full powers and duties under the FSRA apply in respect of *medical schemes*. However, the Minister of Finance has determined that the Council for Medical Schemes (CMS) will exercise these powers until 31 March 2021, albeit with the concurrence of the FSCA.[48] The main objective is to ensure that consumers of medical schemes enjoy adequate protection, and it is submitted that the medical schemes industry is in dire need of reform as far as market conduct is concerned.

Therefore, it is evident that while the FAIS Act initiated market conduct regulation, and will forever be heralded as providing minimum protection standards for consumers, the net of protection is now cast as widely as possible with clear rules in most cases and strategies as to how protection should be extended to *all* consumers of financial services.

It should be mentioned that burial societies and *stokvels* remain a problem. The Policy Paper states that while the FSCA will be able to "appropriately accommodate smaller players in the financial sector (for example, burial societies and *stokvels*)", such entities should be exempted or excluded. While one can understand the need not to "cookie-cutter" all financial institutions, it is submitted that burial societies pose significant risks as these societies hold large amounts of cash and serve poor consumers who attach a particular cultural significance to funerals. Therefore, without being insensitive to the roots of burial societies and *stokvels*, although it undeniably is not ideal to subject *stokvels* and burial societies to a state-of-the-art compliance regime, some protection is urgently needed.[49] A failure to recognise this

41 *Ibid.*
42 19 of 2012.
43 Policy Paper (n 26) 25.
44 45 of 2002.
45 Policy Paper (n 26) 27.
46 24 of 1956.
47 Policy Paper (n 26) 26.
48 *Ibid* 28.
49 See the FAIS Ombud's decision in *Sipho Nchukana v African Compass Funeral Services Andcebisile Mfado* case no FSOS 00238/17-18/WC 2 (www.faisombud.co.za (accessed 1 July 2019)).

need casts serious aspersions on the COFI Bill's aspirations for formal institutions while ignoring pressing needs in informal institutions.

3.3 *Promotion of the fair treatment and protection of financial customers and supporting fair, transparent and efficient financial markets*

The drive to promote fairness is closely related to the objective to promote trust and confidence in the financial sector as fair treatment is likely to strengthen consumer confidence. It is submitted that the former objective in actual fact is a prerequisite for the latter.

The Regulator's pre-occupation with fairness is a topic that has received considerable attention over the past few years.[50] At first, the Financial Services Board (FSB) (now FSCA) published a toolkit on its website and urged financial service providers to use the kit to determine whether they complied with the principles of "Treating Customers Fairly". This toolkit was not law but was used to measure the conduct of institutions.

With the promulgation of the latest Policy Holder Protection Rules in terms of the Long-Term Insurance Act[51] and the Short-Term Insurance Act[52] respectively, these rules formally became part of the market conduct standards for insurance companies. Having said this, many of the stipulations in the General Code of Conduct in terms of the FAIS Act already aim to ensure that consumers of financial products are being treated fairly, and the FAIS Ombud has since its inception used fairness as a standard to evaluate the conduct of financial institutions.[53]

As was to be expected, the COFI Bill aims to further promote fairness outcomes by introducing *several* provisions that address fairness. For instance, clause 19 specifies that the FSCA may set licensing conditions as necessary to ensure the fair treatment of consumers.[54] The proposed Chapter 3 of the Bill on culture and governance contains a number of stipulations on fair treatment. Clause 29(1) specifically states the following:

> "The purpose of this Chapter [3] and any conduct standards prescribed under this Chapter is to set out governance requirements for financial institutions, so that financial customers can be confident that they are dealing with firms and persons where the fair treatment of customers is central to the corporate culture, and in particular, to—
>
> *(a)* improve confidence in the financial sector by promoting governance that supports the fair treatment of financial customers;
>
> *(b)* promote the supply of financial products and financial services that are appropriate for targeted financial customers; and
>
> *(c)* enhance transparency and improved market conduct in the sector."

50 See Millard "Through the looking glass: Fairness in insurance contracts – a caucus race?" 2015 *Journal of Contemporary Roman-Dutch Law* 547–566.
51 52 of 1998.
52 53 of 1998.
53 See Millard and Hattingh (n 10) 183 ff.
54 Clause 19(4)*(h)*.

The crux of the proposed stipulations on fairness is perhaps found in the proposed clause 33, "Unfair contract terms in contracts with retail financial customers", which contains the following stipulations on contracts:

> "(1) A financial institution that provides financial products or financial services to retail financial customers—
>
> *(a)* must ensure that the terms, and conditions of a contract or agreement in respect of a financial product or financial service are fair, reasonable and transparent.
>
> (2) A term or condition of a contract or agreement referred to in subsection (1), or a notice to which a term or condition is purportedly subject, is unfair or unreasonable if—
>
> *(a)* it would cause a significant and unreasonable imbalance in the parties' rights and obligations under the contract;
>
> *(b)* the terms of the contract or agreement are so adverse to the retail customer that they are inequitable;
>
> *(c)* it is not reasonably necessary to protect the legitimate interests of the financial institution, who would be advantaged by the term or condition;
>
> *(d)* it would cause undue detriment, whether financial or otherwise, to a retail financial customer if it were applied or relied on;
>
> *(e)* a retail financial customer is required, on terms that are unfair, unreasonable, or as a condition to entering into a transaction, to—
>
> (i) waive any rights;
>
> (ii) assume any obligation; or
>
> (iii) waive any obligation or liability of the financial institution who is providing a financial product or financial service; or
>
> *(f)* the transaction or agreement is subject to a term or condition, or a prescribed requirement to a retail financial customer, and—
>
> (i) the term, condition or requirement is unfair, unreasonable, unjust or unconscionable; or
>
> (ii) the fact, nature and effect of that term, condition or requirement was not appropriately disclosed to the retain financial customer in a manner that satisfied the prescribed notice requirements."[55]

The above rules elaborate on the existing provisions of the General Code of Conduct (GCC) in terms of the FAIS Act and, in fact, are not new. Furthermore, clause 33(4) stipulates that a term or condition of a contract or agreement is transparent if the term is expressed in reasonably plain language,[56] if it is legible,[57] if it is presented clearly and unambiguously,[58] and if it is readily available to any party affected by

[55] Clause 33(3) further stipulates that "[w]here a financial institution, including a sponsor of a pension fund, provides financial products or financial services to a pension fund or similar member based entity, or to another financial customer that is acting for or on behalf of other retail financial customers, all requirements in subsections (1) and (2) relating to retail financial customers apply equally in relation to the members of that pension fund or other member based entity, or in relation to those other retail customers".

[56] Clause 33(3)*(a)*.

[57] Clause 33(3)*(b)*.

[58] Clause 33(3)*(c)*.

the term.[59] These provisions, at their very core, address the disparity in parties' unequal bargaining positions and hope to level the playing field between parties. On persisting unfairness in the making of contracts, Bradfield postulates the following:

> "It is probably fair to say that unfairness in the making of a contract is generally related to the problem of inequality of bargaining power, which is a problem that has long troubled contract lawyers throughout the world because it often seems unfair to enforce a contract when it is obvious that the one party was in such a weak bargaining position that consent, even if genuine, was at best reluctant. The common law evolved a number of techniques that could be applied in circumstances that may fall within the general ground of inequality of bargaining power. Foremost among these techniques are relaxation of the *caveat subscriptor* rule, limitations on the enforcement of exemption clauses, construction of *contra proferentem*, duress, undue influence and public policy. The common law has now tackled the problem head on by relying on that ground alone, or even primarily, to investigate whether a contract should be regarded as unenforceable."[60]

Much as these remedies have been used, it is submitted that common law alone is not able to provide sufficient remedies to clients. This will perhaps explain why legislation such as the Consumer Protection Act[61] is needed to ensure effective communication, contracts that are easy to read and wording that effectively summarises the heart of the deal.[62] This also holds true for the financial services industry. Although plain language has long been a requirement in terms of the FAIS Act, it is submitted that the wording of insurance contracts (policies) and other contracts that embody complex agreements pertaining to financial products are drafted by services providers, not intermediaries and advisors, with the result that this requirement is hardly ever fulfilled. Perhaps the time is now right for a law of general application for the financial services industry that aims to provide understandable contracts in plain language. This will perhaps have a positive effect on consumers' subjective understanding of contracts for financial services, products and instruments.

3.4 *Promotion of innovation and the development of and investment in innovative technologies, processes and practices*

Advancements in the financial services industry see innovations where so-called "technology companies" compete with "traditional brick-and-mortar financial services businesses".[63] The law should sufficiently enable this new competition while adequately protecting consumers against abuse. While this no doubt is the future, any attempt at financial market conduct regulation should encourage new developments while at the same time ensuring the safety of consumers. More specifically, the Policy Paper states that the Financial Stability Board has defined

[59] Clause 33(3)*(d)*.

[60] Bradfield *Christie's Law of Contract in South Africa* (2016) 14–15.

[61] 68 of 2008.

[62] Newman "The influence of plain language and structure on the readability of contracts" 2010 *Obiter* 738 at 739.

[63] Policy Paper (n 26) 35.

fintech as "technologically-enabled financial innovation that can result in new business models, applications, processes, products, or services with an associated material effect on financial markets and institutions, and the provision of financial services". The Policy Paper further explains that fintech can be applied to a wide range of areas, including electronic payments, automated advice, delivery channels, cyber security and peer-to-peer lending.[64] These technological innovations present a challenge to regulators as they need to understand how to approach fintech. As much as it is the future, regulators should ensure that the innovations do not create unlevel playing fields, or negatively affect competition. In addition, from a market conduct perspective, regulators should consider how customers themselves understand and interact with innovative financial products and platforms, and which customer protection principles do and should apply in these circumstances.[65]

Accordingly, Chapter 2 of the Bill provides for the licensing of new entrants in such a way as to support the entry of "innovative new firms" into the financial sector.[66] The FSCA will accordingly be able to set licensing conditions, as per clause 19, that "provide scope for innovation and the development of and investment in innovative technologies, processes, and practices".[67]

According to the Policy Paper, these draft provisions enable the regulator to develop a supportive approach to fintech entrants while also ensuring that fair customer treatment remains paramount to the entire operation.[68] It goes without saying that this method of regulation is essential to any modern-day market conduct regulation.

3.5 *Promotion of sustainable competition in the provision of financial products and financial services*

One would think that competition in a free market economy is not an issue, as long as such competition is not unfair and does not amount to collusion or any other undesirable practice. The Policy Paper seems to suggest that the FSCA is able to stimulate competition by using a combination of conduct standards, interpretation rulings and guidance notices to steer sector diversification and competition.[69] To this end, the entire Chapter 7 provides for distribution, advice and discretionary investment management. According to the Policy Paper, in addition to providing adequate and clear information, financial institutions must ensure that their distribution models are appropriate "to ensure the delivery of appropriate products and services and, where applicable, provide access to suitable advice".[70] The Policy Paper further states that customers should be able to understand and compare the

[64] *Ibid* 51.
[65] *Ibid*.
[66] *Ibid* 54.
[67] *Ibid.*
[68] The Policy Paper (n 26) 54 makes mention of regulatory sandboxes. A regulatory sandbox is a mechanism for developing a regulation that keeps up with the fast pace of innovation. See www.bbva.com (accessed 5 August 2019).
[69] Policy Paper (n 26) 12.
[70] *Ibid*.

costs and contractual implications associated with sales and distribution models, and distribution models should enhance standards of professionalism, encourage fair competition, and support sustainable business models for financial institutions. The chapter also deals in detail with the provision of advice, and with discretionary investment management.[71]

To this end, clause 64 provides that services related to the distribution of and advice regarding financial products and instruments to customers must be provided in a manner that is as objective as possible,[72] must not be conflicted[73] and must support the delivery of appropriate financial products and instruments.[74] It is submitted that the profit motive in selling remains a reality, and this is the factor that drives the selling of goods and services. This clause no doubt intends to bring fairness into the financial services industry, but it is not clear exactly how the regulator intends to make this work. Perhaps clause 65(1) provides a clue. This clause stipulates that financial institutions must satisfy themselves that the methods used to distribute or provide advice on financial products and instruments are appropriate to the nature and complexity of the product and to the targeted financial customers. In addition, clause 65(4) stipulates that the information given by a financial institution that provides a financial product or financial instrument to a financial institution or any other person who is involved in distributing or providing advice on a financial product or financial instrument must be sufficient to enable it to understand and adequately place the financial product, instrument or financial service in the relevant target market; identify the target market for which the financial product, instrument or financial service is designed; and identify types of financial customers for whom the financial product or financial instrument is likely to be inappropriate. This simply amounts to "know your (potential) customer". Financial institutions already use market research and in future will be more reliant on this to ensure effective marketing.

71 *Ibid*.

72 Clause 64*(a)*.

73 Clause 64*(b)*. Conflict of interest requirements already form part of the FAIS Act. Board Notice 58 of 10 April 2010 introduced several matters into the GCC. "Conflict of interest" is described as "any situation in which a provider or a representative has an actual or potential interest that may, in rendering a financial service to a client, (a) *influence* the *objective performance* of his, her or its obligations to that client, or (b) prevent a provider or representative from rendering an *unbiased and fair* financial service to that client, or from acting in the interests of that client, including, but not limited to (i) a financial interest; (ii) an ownership interest; (iii) any relationship with a third party". Board Notice 58 of 19 April 2010 defines "financial interest" as any cash, cash equivalent, voucher, gift, service, advantage, benefit, discount, domestic or foreign travel, hospitality, accommodation, sponsorship, other incentive or valuable consideration other than an ownership interest or training. Training on products and legal matters relating to these products, general financial and industry information, and specialised technological systems of a third party necessary for the rendering of a financial service are not to be considered a financial interest. It is therefore evident that the FAIS Act did make a concerted effort to ensure that financial advice is not rendered with a view to earning a secret profit. See also Millard and Hatting (n 10) 30.

74 Clause 64*(c)*.

The Bill also stipulates that appropriate contracts should be in place between financial institutions to ensure accountability.[75] Other stipulations specify standards for advice or discretionary investment management.[76]

As the focus of the FAIS Act has for so long been on advice, one instinctively searches for provisions that regulate this activity. The important clause on advice is clause 68, which stipulates that a provider of a financial product or financial instrument must take reasonable measures to monitor the quality of advice provided to customers;[77] ensure that financial institutions providing advice in respect of the provider's financial products or financial instruments have adequate knowledge of those products or instruments, and those who are targeted;[78] and ensure that financial institutions providing advice in respect of the provider's financial products or financial instruments comply with any applicable prescribed competency requirements in relation to the provision of such advice.[79]

Overall, it is evident that the standard that was set for advice in terms of the FAIS Act stood the test of time, and it in fact is doubted whether a higher standard is needed. However, it is suggested that advice is only *one* stage in the product life cycle, and although it is just as important as it has always been, the context now is such that all the other aspects in the cycle are of equal importance and a failure to advise a client is not the only failure that may potentially lead to mis-selling and mis-buying.

3.6 *Promotion of financial inclusion and transformation of the financial sector*

The COFI Bill Policy Paper explains that research on financial exclusion goes back to the Financial Sector Charter, which came into effect in January 2004.[80]

The Financial Sector Code, which was first gazetted in 2012, envisages the development of "a transformed, vibrant and globally competitive financial sector that reflects the demographics of South Africa".[81] The revised code of December 2017 aims to introduce refined and improved transformation targets for the financial sector, and the Policy Paper suggests that the FSCA, as a regulator, has the ideal characteristics to transform the sector. As to how this is to be achieved, the Policy Paper suggests that it should be asked how the sector supports real economic activity or, more specifically, what services are provided to consumers;[82] who owns the firms that manage the assets; how sensitive these firms are to South Africa's needs; and how the assets in the system are put to use.[83] These assets include procurement, empowerment financing and socio-economic development, as well as the management control, employment equity and skills development of assets that

[75] Clause 66(1).
[76] Clause 67.
[77] Clause 68*(a)*.
[78] Clause 68*(b)*.
[79] Clause 68*(c)*.
[80] Policy Paper (n 26) 29.
[81] *Ibid.*
[82] *Ibid.*
[83] *Ibid.*

belong to financial service providers.[84] Another important transformative imperative of the COFI Bill is to ensure that the sector is regulated in such a way that conduct regulation allows the regulator to apply rules in a way that is proportionate to the size, complexity and risk posed by financial institutions.[85] This means that COFI should not impose compliance requirements that make product and service costs too high to service low-income markets, and it should not limit diversification, competition and sustainable growth in financial services by imposing unreasonable barriers to new businesses that want to enter the financial services industry.[86]

While it is accepted that the COFI Bill has commendable objectives and that transformation no doubt forms part of it, it is suggested that transformation of the industry should be balanced with ideals such as the safety of the industry as a whole and that access to new role players should never jeopardise the other objectives contained in the Bill. In addition, these imperatives increase governance costs and will lead to more expensive financial services.

3.7 *Assisting the South African Reserve Bank in maintaining financial stability*

This objective is a natural consequence of the Twin Peaks method of financial regulation. It is submitted that there are many ways in which these two authorities must work together in order to ensure that prudential standards support the objectives contained in the COFI Bill. However, it is suggested that the entire Twin Peaks model is an over-reaction to incidents such as Fidentia and Sharemax. The inability of the Regulator to provide redress[87] to the victims of these scams should not have led to the exaggerated, knee-jerk reaction that is Twin Peaks.

4 Evaluation and conclusion

It is evident from the exposition above that the COFI Bill definitely is not a "fluffed-out" version of the FAIS Act. The emphasis on advice under the FAIS Act resulted in other aspects in the value chain being under-regulated. It therefore can be said that the manner in which products are designed, distributed and sold are far from ideal and that there are aspects that can be improved, most notably those antiquated provisions in pre-constitutional financial legislation, such as the Pension Funds Act and the Medical Schemes Act, that are out of touch with the needs of South African consumers.

[84] *Ibid* 30.

[85] *Ibid* 31.

[86] See *eg* clause 38, which stipulates: "If a financial institution is subject to requirements of the Broad-Based Black Economic Empowerment Act, 2003 (Act No 53 of 2003) and the Financial Sector Code for Broad-Based Black Economic Empowerment issued in terms of section 9(1) of that Act, it must have a policy and plan in place to meet its stated commitments in terms of promoting transformation of the financial sector in line with those requirements."

[87] Chapter 11 of the COFI Bill provides for redress and specifically mentions that the Authority may undertake specified measures to remedy the effects of a contravention of this Act, including through the provision of appropriate redress to financial customers. Clause 103 (rather unnecessarily) mentions that consumers who have suffered a loss are entitled to redress.

At the same time, it is submitted that some of the objectives of the Bill are simply too ambitious. While product standards will compel service providers to compete on a more equal footing, it is ill-advised, if not perilous, to micro-manage an industry that continues to deliver a significant contribution to an otherwise flailing economy. By manipulating financial service providers in an ill-fated attempt to rejuvenate an economy that has been driven to the brink of collapse by bad governance in most other sectors may very well prove to be the final straw that breaks the camel's back.

Transformative imperatives also will not force South African consumers of financial instruments, products and services to better manage their resources. Rather, consumer confidence will far more likely be restored if there are better tax benefits available to those who participate in financial services, such as pension funds and annuities, which have the potential to empower them financially.

Overall, while the Bill is commended for not just being a "cut-and-paste" exercise and for truly considering all aspects of market conduct regulation, including better communication with consumers at an appropriate level, the Bill is overly ambitious and is likely to provide problems for solutions.

The taxation of South African expatriate employees: Dispelling the myth of financial emigration

LG TREDOUX*

1 Introduction

The emigration of skilled individuals from South Africa (SA) increased significantly in the 1990s and still continues.[1] Reliable statistics on the number of individuals that leave SA are scarce and admittedly underestimated by the South African government.[2] Statistics SA confirmed in 2019 that:

> "Detailed information on the departure of travellers is not available in the movement control system. Data on the purpose of travel and the number of days South African residents intend to spend or spent abroad are not collected by the DHA. Hence, it is not possible to categorise South African residents as tourists or non-tourists."[3]

A significant loss of skilled individual taxpayers could lead to erosion of the South African tax base. In the 2017/2018 year of assessment, personal income tax constituted the largest percentage (38%) of the total tax revenue collected.[4] The South African Revenue Service (SARS) collected an amount of R321.4 billion during this tax year from only 4.9 million individual taxpayers.[5] This relatively

* Senior Lecturer in Tax Law, University of South Africa.

1 Myburgh "Explaining emigration from South Africa" 2004 *SA Journal of Economics* 122; see also Statistics South Africa "Documented migration" Report No 03-51-03 (2003) xiv and iv which indicates that three definite peaks in self-declared emigration occurred in 1977, 1986 and 1994 and that a gradual increase is still taking place after 1994. In 2003 it was estimated that the number of recorded emigrations out of SA was 48.4% higher compared to the same period for the year 2002. Höppli "New evidence on the brain drain from South Africa" *Policy Research on International Services and Manufacturing* (2014) PRISM Working Paper Series No 1/2014 1 at 17 indicates that the brain drain still continues; also see Zamayriha "Sharp rise in number of South Africans leaving the country" (https://city-press.news24.com/News/sharp-rise-in-number-of-south-africans-leaving-the-country-20190117 (accessed 17 January 2019)) 1.

2 Bhorat, Meyer and Mlatsheni *Skilled Labour Migration from Developing Countries: Study on South and Southern Africa* International Migration Paper 52, International Migration Programme, International Labour Office, Geneva (2018) 8, where the Department of Home Affairs (DHA) conceded that they only have limited statistics available; Mattes and Mniki "Restless minds: South African students and the brain drain" (March 2007) 24(1) *Development Southern Africa* 25 at 26; Statistics South Africa 2019 "Tourism and migration" Statistical Release P0351 (May 2019) 4; Businesstech "South Africa's emigration problem – no one knows how big the brain drain really is" (https://businesstech.co.za/news/lifestyle/318736/south-africas-emigration-problem-no-one-knows-how-big-the-brain-drain-really-is/ (accessed 29 July 2019)) 1 at 2.

3 Statistics South Africa 2019 (n 2) 4.

4 National Treasury and SARS "Tax Statistics 2018" (December 2018) 19, 30–31 and 34.

5 *Ibid*.

ABLU 2019 (Part ONE).indb 147 2019/10/09 9:49 AM

small amount of individuals carry the brunt of the tax burden, in spite of 20 million individuals being registered for tax[6] and a current population count of approximately 58.78 million.[7] An estimated 900 000 individuals that were born in SA currently live and work abroad.[8] This estimate is believed to be conservative and the true number is considered to be three times as high.[9] Most of the individuals that leave SA are skilled, high-income earners that contribute to the tax base.[10]

Besides emigrating, many other South-African-born individuals work in other states, without relinquishing all ties with SA. From a tax perspective, this benefits those expatriates as they need not wind up their affairs and either qualify for unilateral relief from international double taxation[11] in terms of the South African Income Tax Act,[12] or relief in terms of a double tax agreement (DTA). This affects the South African tax base negatively as expatriates often pay very little South African income tax while working abroad, although they may return to SA from time to time and still consider SA their home country.

To address this potential tax base erosion, National Treasury announced in 2017 that the exemption from South African income tax for certain expatriate employees will be amended with effect from 1 March 2020.[13] This proposed amendment was widely reported in the press where the impression was created by certain reports that this change will affect all expatriates.[14] Many tax advisors also advocate that all expatriates should financially emigrate to avoid tax liability in SA.[15]

[6] *Ibid*.

[7] Statistics South Africa “Mid-year population estimates” Statistical Release P0302 (29 July 2019) 1 at 10.

[8] Businesstech “4 real facts about emigration in South Africa” 28 August 2018 (https://businesstech.co.za/news/wealth/267845/4-real-facts-about-emigration-in-south-africa/ (accessed 11 August 2019)) 1.

[9] Bhorat, Meyer and Mlatsheni (n 2) 13, 27; Businesstech “5 scary facts about emigration in South Africa” 3 May 2019 (https://businesstech.co.za/news/lifestyle/314656/5-scary-facts-about-emigration-in-south-africa/ (accessed 11 August 2019)) 1 at 3.

[10] Businesstech (n 9) 3.

[11] International double taxation in an economic sense refers to the same income being taxable twice, whereas in the juridical sense, double taxation refers to the same taxpayer being taxed twice in two different states; see Olivier and Honiball *International Tax: A South African Perspective* (2011) 6.

[12] 58 of 1962. Section 6*quat* of the Income Tax Act provides for a tax credit that may be subtracted from the expatriate’s South African tax liability; an exemption from South African income tax is provided for individuals in the shipping industry (s 10(1)*(o)*(i)) and certain expatriate employees (s 10(1)*(o)*(ii)); an individual may also escape tax liability or pay a reduced amount of tax based on the provisions contained in a DTA (s 1: proviso to the definition of “resident”).

[13] Section 16(1)*(g)* of the Taxation Laws Amendment Act 17 of 2017 amends s 10(1)*(o)*(ii) of the Income Tax Act with effect from 1 March 2020 for all years of assessment commencing after this date.

[14] Daniel “South Africans are emigrating abroad in record high numbers” 6 August 2019 (https://www.thesouthafrican.com/news/south-africans-are-emigrating-abroad-in-record-high-numbers/ (accessed 6 August 2018) 2; www.taxconsulting.co.za/south-africans-are-emigrating-abroad-in-record-high-numbers/ (accessed 8 May 2019) 1).

[15] Ryan “Financial emigration is the new way out. The taxman targets foreign earnings in excess of R1m – but at what cost to the country?” 11 June 2019 (https://www.moneyweb.co.za/news/south-africa/financial-emigration-is-the-new-way-out/ 1/ (accessed 8 May 2019)) 1; Smith “Financial emigration: What it involves” 7 April 2019 (https://www.fin24.com/Money/Tax/financial-

This chapter analyses the taxation of expatriates in terms of South African domestic tax law and examines the concept of "financial emigration" to determine its relevance from an income tax perspective. It further explains the tests to determine tax residency in South African domestic tax law, the current and proposed exemption from South African income tax for expatriate employees, and the interaction of the residence requirements in tax treaties with the relief from double taxation granted in domestic law. It also provides a very brief overview of the role of the commercial banks, the South African Reserve Bank (SARB) and the SARS upon "financial emigration" within the current legislative and policy framework. The contribution concludes by clarifying the law, pointing out problematic aspects and making a few final remarks concerning financial emigration and the liability of expatriates for income tax in SA.

2 The tax liability of individuals in South Africa

The starting point to determine whether an individual is liable for income tax in SA is the domestic law as contained in the Income Tax Act. Section 5 of the Income Tax Act (the charging section) levies income tax on the taxable income of any person during a year of assessment consisting of 12 months and ending on the last day of February every year.[16] The taxable income of an individual is calculated by determining the gross income of a taxpayer and subtracting certain exempt amounts from it.[17] The result of this computation is defined as income.[18] From this income, allowable deductions and allowances,[19] as well as assessed losses,[20] are subtracted, any capital gain or loss is added or subtracted,[21] and the result is classified as taxable income.[22] The starting point of the income tax calculation for individuals is thus "gross income".

emigration-what-it-involves-20190322 (accessed 8 May 2019)) 1; Businesstech "Massive jump in the number of South Africans applying to financially emigrate" 30 July 2019 (https://businesstech.co.za/news/finance/321847/massive-jump-in-the-number-of-south-africans-applying-for-financial-emigration/ (accessed 8 May 2019)) 1; Leon "Opinion: Tax residency and financial emigration: Not everyone is an expert!" *IOL Personal Finance* 30 July 2019 (https://www.iol.co.za/personal-finance/tax/opinion-tax-residency-and-financial-emigration-not-everyone-is-an-expert-19698072 1/3 (accessed 8 May 2019)) 1 at 1–3.

16 Section 5(1)*(c)* of the Income Tax Act.

17 Section 1(1) of the Income Tax Act: definition of "income"; Stiglingh et al *Silke South African Income Tax 2019* (2019) 22–23.

18 Section 1(1) of the Income Tax Act: definition of "income"; Stiglingh (n 17) 22–23.

19 These deductions are contained in s 11–19 and s 21–21P of the Income Tax Act; Stiglingh (n 17) 22–23.

20 Section 20–20B of the Income Tax Act; Stiglingh (n 17) 22–23.

21 Section 26A of the Income Tax Act; De Koker and Williams *Silke on South African Tax* (2019) par 1.3 and 2.0.

22 Section 1(1) of the Income Tax Act: definition of "taxable income"; Stiglingh (n 17) 22–23.

Section 1(1) of the Income Tax Act defines "gross income" as:

> "'gross income' in relation to any year or period of assessment, means—
>
> (i) in the case of any resident, the total amount, in cash or otherwise, received by or accrued to or in favour of such resident; or
>
> (ii) in the case of any person other than a resident, the total amount, in cash or otherwise, received by or accrued to or in favour of such person from a source within the Republic,
>
> during such year or period of assessment, excluding receipts or accruals of a capital nature …"

This definition distinguishes between residents who are taxed on all amounts received from all worldwide sources and non-residents who are taxed in SA if the source of their income is located in the Republic.[23] Many of the terms in the definition of "gross income" are not defined further, and reliance is placed on case law to determine the meaning of such concepts.[24] As the taxation of expatriates centres on the question of residence, this concept requires further analysis.

3 The tax residence of an expatriate

An expatriate is generally described as an individual who is a citizen or resident of one country but is living in another country.[25] The term "expatriate" is not used in South African tax legislation. When someone moves across borders it can be challenging to ascertain in which state that person is liable for income tax at a specific moment in time. A *nexus* is required with the state that levies income tax on this individual.[26] As many states levy income tax on a residence or source basis, the individual working abroad has to consider the domestic law of the state from where he or she originates, the state where he or she is resident, and the state or states where he or she is working.[27] If a double tax agreement (DTA) was concluded between the two states concerned, the expatriate must also consider the provisions of that treaty.[28] A DTA, however, cannot found tax liability and the domestic legislation concerned must contain a charging provision which imposes tax liability on the person working or living in that particular state.[29]

Although an expatriate might work outside SA, he or she could be liable for income tax in SA if he or she is "resident" in the Republic as defined in the Income Tax Act.[30] The meaning of this term is unique to the application of tax legislation.

23 Oguttu *International Tax Law: Offshore Tax Avoidance in South Africa* (2015) 71.
24 Stiglingh (n 17) 29.
25 *Oxford Online Dictionary Lexico* "expatriate" (https:/www.lexico.com/en/definition/expatriate (accessed 21 August 2019)); *Cambridge Online Dictionary* "expatriate" (https:/www.dictionary.cambridge.org>expatriate (accessed 21 August 2019)).
26 Oguttu (n 23) 67–68.
27 Olivier and Honiball (n 11) 9–10.
28 *Ibid*.
29 Olivier and Honiball (n 11) 30; in South Africa this is found in s 5 of the Income Tax Act as explained in par 2 above.
30 Section 1 of Income Tax Act: definition of "resident".

It is not the same as domicile or citizenship and differs from the meaning of resident used in the exchange control regulations.[31] Two tests determine whether an individual is tax resident in SA, namely the "ordinarily resident" and "physical presence" tests.[32] In addition, the definition of "resident" determines when residency ceases and ensures that an individual who is resident in another state in terms of a DTA is not classified as a resident of SA.[33]

3.1 *The meaning of "ordinarily resident"*

An individual is resident in SA if he or she is "ordinarily resident" within the Republic.[34] The phrase "ordinarily resident" is not defined in the Income Tax Act and the guidelines found in case law is considered to establish its meaning in each specific situation.[35] In 1946, the Appellate Division, as it then was, confirmed in *Cohen v CIR*[36] that a person is resident in the state to which he

> "as a matter of fact returns from his wanderings, as contrasted with other lands it might be called his usual or principal residence and would be described more aptly than other countries as his real home."[37]

The South African courts have confirmed the *Cohen* decision and in addition found that a person may be ordinarily resident in SA, despite being temporarily absent during a specific tax year.[38] Similarly, the House of Lords in the English case of *Shah v Barnet London Borough Council*[39] referred to habitual and normal residence as a test "apart from temporary or occasional absences of long or short duration".[40]

Olivier and Honiball are of the view that an individual can be ordinarily resident in SA in spite of being physically absent during a specific year of assessment or even several years.[41] If an individual has the intention to return to SA as his or

[31] Olivier and Honiball (n 11) 19.

[32] Section 1 of Income Tax Act: definition of "resident"; Olivier and Honiball (n 11) 23.

[33] Section 1 of the Income Tax Act: proviso to the definition of "resident".

[34] Section 1 of the Income Tax Act: definition of 'resident' par *(a)*; Olivier and Honiball (n 11) 20. The term "Republic" is further defined in s 1 of the Income Tax Act as "the Republic of South Africa, and when used in geographical sense, includes the territorial sea thereof as well as any area outside the territorial sea which has been or may be designated, under international law and the laws of South Africa, as areas within which South Africa may exercise sovereign rights or jurisdiction with regard to the exploration or exploitation of natural resources".

[35] Olivier and Honiball (n 11) 20; Stiglingh (n 17) 29.

[36] 1946 AD 174.

[37] *Cohen v CIR* (n 36) 185; this decision was confirmed in *CIR v Kuttel* 1992 (3) SA 242 (A), 54 SATC 298 at 306.

[38] *CIR v Kuttel* (n 37) 306.

[39] [1983] 1 All ER 226 (HL).

[40] *Shah v Barnet London Borough Council* (n 39) 234b–c. The leading Canadian case, discussed by Olivier and Honiball (n 11) 20, namely *Thompson v Minister of National Revenue* 2 DTC 812 (SCC), described the ordinary residence of an individual as a place "where in the settled routine of his life he regularly normally and customarily lives", put differently, the place "at which he in mind and in fact settles into or maintains or centralizes his ordinary mode of living with its accessories in social relations, interest and conveniences".

[41] Olivier and Honiball (n 11) 22–23.

her true home, such an individual is still regarded as ordinarily resident in SA.[42] This interpretation of the phrase "ordinary residence" is in line with the courts' interpretation and can cause an expatriate to be liable for tax in SA, despite not being physically present in this tax jurisdiction.

SARS is of the view that two requirements must met for an individual to be ordinarily resident in SA: first, the intention to be ordinarily resident and, secondly, the taking of steps that indicate the practical implementation of this intention.[43] In determining this intention, the actions of the person are examined based on several factors and the circumstances as whole.[44] These non-exhaustive factors include the location of a fixed residence, habitual abode, business or personal interests, personal belongings, nationality, family and social ties, political cultural involvement, a person's status in the country and the question whether such a person applied for permanent residence elsewhere.[45] These factors indicate that SARS considers more than the mere place of employment or location of work activities when determining the place of ordinary residence of an expatriate. SARS further opines that it is not possible to add a specific time limit to ascertain ordinary residence.[46] If a person was physically absent, the intention, purpose, nature and duration of such absence should be investigated.[47]

In the modern economy an individual could have many multinational economic opportunities causing a person to be constantly on the move and without a permanent home.[48] The expatriate bears the onus of proof to show that he or she is a non-resident of SA, or that an amount, transaction, event or item is not taxable or is exempt.[49]

3.2 *The physical presence test*

The physical presence test is based on an objective approach, which considers the time an individual spends in SA and disregards the intention of the taxpayer. This test imposes tax resident status on individuals who are not ordinarily resident in terms of the common law, but are sufficiently physically present in SA.[50] If a person is ordinarily resident in SA, this expatriate is tax resident, in spite of the person possibly not complying with the amount of days required to be resident based on physical presence. A person who is physically present in SA for more than 91 days in aggregate in the year of assessment and each of the preceding five years, as well as an aggregate of more than 915 days in total during the five years preceding

42 Olivier and Honiball (n 11) 23.
43 SARS Income Tax Interpretation Note 3 dated 4 February 2002 (http://www.sars.gov.za/home.asp?pid=54958 (accessed 20 July 2019)) 4.
44 *Ibid*.
45 *Ibid*.
46 *Ibid*.
47 *Ibid*.
48 Olivier and Honiball (n 11) 23; SARS Interpretation Note 3 (n 43) 4.
49 Section 102 of the Tax Administration Act 28 of 2011.
50 Section 1 of the Income Tax Act: definition of "resident" par (*a*)(ii).

the year of assessment, is resident in SA.[51] The days need not be continuous.[52] Practically, a person can only become a resident of SA for tax purposes based on this test in the sixth year of assessment that the person was sufficiently present.[53] This test was introduced to encourage foreign skilled workers to work temporarily in SA to address the skill shortage.[54] From an expatriate's perspective, an individual who is or was resident based on this test should carefully monitor the amount of days spent in SA to determine whether he or she is classified as resident in SA.

3.3 *The exclusion of individuals that are exclusively resident outside SA in terms of a DTA*

Individuals who are resident in another country might be excluded from the South African definition of resident in domestic law. The proviso to the definition of resident in the Income Tax Act excludes an individual from South African tax residence if the person is exclusively resident in another state due to the application of the provisions of a DTA concluded between SA and such other state.[55] This treaty override applies to the definition of a resident irrespective of whether the individual was a South African tax resident based on ordinary residence or physical presence. It is a useful clarification for expatriates who find themselves in jurisdictions that have treaties with SA. It was necessary to include this specific treaty override in legislation, as normally the provisions of a DTA have the same effect and status as national legislation.[56] This equal status was confirmed in 1975 by the Supreme Court of Appeal in *Secretary for Inland Revenue v Downing*,[57] and by the Western Cape High Court in *Commissioner South African Revenue Service v Van Kets*[58] in 2012. The latter case further confirmed that the correct application of section 231 of the Constitution requires that any conflict between a provision in a DTA and the South African domestic law must be solved by allowing the provisions in the treaty to prevail.[59] This approach was also applied by the Supreme Court of Appeal in *CSARS v Tradehold*.[60]

51 Section 1 of the Income Tax Act: definition of "resident" par *(a)*(ii)*(aa)* and *(bb)*.

52 Section 1 of the Income Tax Act: definition of "resident" par *(a)*(ii)*(aa)* and *(bb)*; Olivier and Honiball (n 11) 23 state that part of a day will be treated as a whole day.

53 Olivier and Honiball (n 11) 23.

54 *Ibid*.

55 Section 1 of the Income Tax Act: definition of "resident" par *(b)*.

56 Section 108(2) of the Income Tax Act; *Secretary for Inland Revenue v Downing* 1975 (4) SA 518 (A) 523A; s 231(4) of the Constitution of the Republic of South Africa, 1996; *Glenister v President of the Republic of South Africa and Others* 2011 (3) SA 347 (CC); Arnold and McIntyre *International Tax Primer* (2002) 104; Rothagi *Basic International Taxation Volume One: Principles of International Taxation* (2005) 17–19; Olivier and Honiball (n 11) 305. See further Du Plessis "The incorporation of double taxation agreements into South African domestic law" 2015 *PELJ* 1188–1204. Du Plessis "Some thoughts on the interpretation of tax treaties in South Africa" 2012 *SA Merc LJ* 31 describes the diverse views on this question and the interpretation thereof by of the courts.

57 1975 (4) SA 518 (A) 523A.

58 2012 (3) SA 399 (WCC).

59 *Commissioner South African Revenue Service v Van Kets* (n 58) par 25.

60 (2012) 74 SATC 263 at 265.

For treaty purposes, the residence of an individual is in the state where a person is liable to tax based on a nexus of either citizenship, domicile, residence or a similar criterion, but excluding the source criterion.[61]

In terms of the South African domestic law, the question whether a person can be resident in more than one location at the same time (dual residence) was not expressly decided by the court in *Cohen*, but one could argue that a person cannot have more than one "ordinary" residence. It is, however, possible for an expatriate to be a dual resident and be connected to two different states. For example, the person could be physically present and working in another state and considered tax resident there but could also have a home and family in SA, resulting in classification as ordinarily resident in SA, based on the common law. This dual residency is caused by differing resident tests and definitions in the domestic laws of both states.[62] This dual residency of expatriates must be addressed and/or eliminated to allocate taxing rights to the appropriate state. The proviso to the definition of resident in South African domestic law also requires exclusive residence in another state for an individual not to be resident in SA.

3.4 *Dual residence and the tie-breaker provisions in the OECD Model Tax Convention*

If an individual is a dual resident, the application of tie-breaker rules may classify this person as exclusively resident in one of these specific states. The DTAs concluded between SA and other states all make use of the Organisation for Economic Co-operation and Development (OECD) Model Tax Convention's tie-breaker rules to determine the state in which an individual is exclusively resident for tax purposes.[63]

Article 4(2) of the OECD Model Tax Convention contains a four-step procedure to determine the residency of individuals. First, an individual is deemed resident in the state where that person has a permanent home.[64] Secondly, if a permanent home is found in both states, or the individual has no permanent home, the state where the individual's personal and economic interests are closer – his or her "centre of vital interests" – is considered his or her country of residence.[65] Thirdly, if no dominant state is identified based on the vital interests of the individual, the state where his or her "habitual abode" is situated, will be his or her ordinary place of residence.[66] Fourthly, if the "habitual abode" cannot be determined or is in both states, the state of which the person is a national is his or her state of ordinary residence.[67] If the matter cannot be resolved by applying the four steps in this article, a person who is a national of both contracting states, or of neither of the two states, may contact the authorities of both the contracting states to settle the issue via a mutual agreement

[61] Article 4(1) of the Organisation for Economic Co-operation and Development (OECD) Model Tax Convention (2017); Olivier and Honiball (n 11) 29; Oguttu (n 23) 67–68.

[62] Olivier and Honiball (n 11) 32.

[63] Olivier and Honiball (n 11) 36.

[64] Article 4(2) of the OECD Model Tax Convention on Income and Capital (2017).

[65] *Ibid*.

[66] *Ibid*.

[67] *Ibid*.

procedure.[68] From the perspective of the expatriate, it is clear that more than the mere location of his or her work activities is considered when tax residence is allocated to a specific state.

For an expatriate to be tax resident in another state, the factual breaking of ties with SA is required for both the ordinary residence test and the application of the tax treaty override in the definition of resident. The English court in *Sheperd v Revenue and Customs Commissioners*[69] found that evidence of a "distinct break" with the previous jurisdiction is required for a person to be treated as non-resident, even if that person's intention was to reduce his or her tax liability.[70]

When an individual works in many different jurisdictions, or is possibly tax resident in three different states, bilateral tax treaties cannot offer a solution.[71] The OECD developed a Multilateral Instrument to serve as a guideline in cases where more than two jurisdictions are involved.[72] Although SA is a signatory of the OECD Multilateral Instrument,[73] SA has not entered into any multilateral tax treaties yet. The treaty override in the proviso to the definition of resident also does not contain a reference to the OECD Multilateral Instrument.[74] In this scenario, I submit that the normal ordinary residence test will apply and one would have to determine residency based on the provisions of the DTAs that are applicable (if any) through a process of elimination. The dominant factual ties and/or citizenship should be decisive, and if not, unilateral relief in SA tax legislation or the domestic law of the other state could reduce double taxation. From a revenue collector's perspective, a risk of double (or triple) non-taxation remains prevalent in this scenario.

4 Unilateral tax relief granted to expatriate employees

The South African Income Tax Act grants relief from international double taxation to specific categories of individuals that work outside the Republic. Expatriates who have not ended their SA tax residency (either expressly or via the operation of a DTA) remain liable for income tax, but may subtract the amount of foreign tax paid from their South African income tax liability as a credit,[75] or may claim a specific exemption[76] unilaterally regulated in the Income Tax Act. The same relief is available to expatriates who have not broken all ties with SA but are unsure of their intention to return to SA. Contrary to the impression created in the media,[77]

68 *Ibid*.
69 2006 STC 1821 and on appeal in 2006 EWHC 1512 (Ch) this decision was confirmed.
70 *Sheperd v Revenue and Customs Commissioners* (n 69).
71 Olivier and Honiball (n 11) 32.
72 Musviba "South Africa signs the multilateral BEPS convention" 4 September 2019 (https://www.sataxguide.co.za/south-africa-signs-the-multilateral-beps-convention/ (accessed 16 September 2019)).
73 *Ibid*.
74 Section 1 of the Income Tax Act: proviso to the definition of "resident".
75 Section 6*quat* of the Income Tax Act.
76 Stiglingh (n 17) 79; De Koker and Williams (n 21) par 9.1. The exemptions are found in s 10 of the Income Tax Act.
77 See par 1 above.

section 10(1)*(o)* of the Income Tax Act does not apply to all expatriates.[78] This section only exempts certain expatriate employees and individuals involved in the shipping industry under specific circumstances.[79] In addition, an exemption is also granted to certain government employees.[80]

4.1 *Members of crew ships, mining operations at sea, international shipping and fishing*

The remuneration of an officer or crew member of a ship which transports passengers or goods for reward, performs certain mining operations on the sea bed, or is involved in international shipping or fishing outside SA is exempt from income tax, if the person performing such activities was outside SA for an aggregate of more than 183 full days during a year of assessment.[81] This exemption only applies to remuneration as defined in paragraph 1 of the Fourth Schedule to the Income Tax Act, which regulates provisional tax that is paid in advance by employers on behalf of employees, also known as "pay as you earn" (PAYE).[82] This remuneration includes salary, wages, leave pay, overtime pay, commission, fees, emolument, gratuity, bonus, superannuation, pension and allowances received by any person, whether it is received for services rendered or not.[83] It further includes certain identified amounts that are classified as specific inclusions in gross income.[84] The term "remuneration", as defined in the Fourth Schedule to the Income Tax Act, also includes allowances in respect of meals, accommodation, business expenses, other specific travel allowances, gains and benefits that accrue to employees that participate in specific share incentive schemes, as well as certain dividends.[85]

Save for the specific requirements contained in the detailed sections that regulate the nature of remuneration received or accrued, the nature of the employment relationship is not decisive in determining whether individuals qualify for an exemption in terms of section 10(1)*(o)*(i) of the Income Tax Act. The emphasis is on the nature of the activities that an individual is involved in. This exemption applies to "any person" that was outside SA for the requisite number of days, namely 183 full days.[86]

78 Section 10(1)*(o)*(i) and (ii) of the Income Tax Act.
79 Section 10(1)*(o)* of the Income Tax Act.
80 Section 10(1)*(p)* of the Income Tax Act.
81 Section 10(1)*(o)*(i) of the Income Tax Act.
82 Section 10(1)*(o)*(i) of the Income Tax Act.
83 Par 1 of the Fourth Schedule to the Income Tax Act: definition of "remuneration".
84 Par 1 of the Fourth Schedule to the Income Tax Act: definition of "remuneration" par *(a)* includes amounts that are included in gross income in terms of par *(a)*, *(c)*, *(c*A*)*, *(c*B*)*, *(d)*, *(e)*, *(e*A*)*, and *(f)* of the definition of gross income. These include regular amounts that are part of the general definition of gross income, amounts received or accrued for services rendered in respect of the holding of office or employment (par *(c)*), labour brokers, personal service providers, personal service companies or trusts (par (*c*A)), restraint of trade payments (par *(c*B*)*), voluntary awards for the termination of service and certain policies of insurance (par *(d)*), retirement fund lump sum benefits and withdrawals (par *(e)*), payments by pension funds (par *(e*A*)*), and amounts paid in terms of a contract of employment or service (par *(f)*).
85 Par 1 of the Fourth Schedule to the Income Tax Act: definition of "remuneration".
86 Section 10(1)*(o)*(i) of the Income Tax Act.

4.2 *Services rendered outside SA by expatriate employees*

The remuneration of an employee for services rendered outside SA for or on behalf of any employer is exempt if that employee was outside SA for more than 183 days in aggregate during the year of assessment and for a period exceeding 60 consecutive full days during the 12 months of that specific tax year.[87] The type of remuneration is listed in the section itself and includes salary, wages, leave pay, overtime, bonus, commission, fees, emolument or allowance, amounts in terms of paragraph (i) of the definition of gross income,[88] or amounts referred to in sections 8, 8B and 8C of the Income Tax Act.[89] Section 8 includes a wide variety of expenses (fringe benefits) that an employer pays on behalf of its employee, section 8B taxes certain amounts that employees receive in terms of broad-based employee share schemes, and section 8C regulates the taxation of the vesting of equity instruments that employees might receive by virtue of their employment.

To qualify for this exemption the services must be rendered outside SA by the employee for or on behalf of the employer.[90] When determining whether a person was outside SA, employees who are in transit between two states via a South African airport and who do not formally enter SA through a port of entry, are not considered to enter SA.[91] If the services of the employee were rendered over two different years of assessment, the remuneration is deemed to accrue evenly/equally over the period of the service.[92]

This exemption does not apply to holders of a public office,[93] or to independent services, work or labour.[94] A person who is self-employed is not entitled to the exemption. Income for professional services rendered is taxed in accordance with the normal residence rules,[95] and included in gross income in the worldwide income of the expatriate if he or she is resident.

The term "employee" is not defined in section 1 of the Income Tax Act for purposes of the general application of this Act, but is defined for purposes of the application of the PAYE rules in the Fourth Schedule to the Income Tax Act.[96] Here, an employee is defined (for purposes of the application of this schedule only) as a person that is not a company to whom remuneration is paid or accrues, certain labour brokers, persons declared to be employees by the Minister of Finance and any director of a private company.[97] In section 10(1)*(o)*(ii), however, the type of remuneration is not defined with reference to the Fourth Schedule. Olivier and Honiball opine that the normal meaning of employment would require a form of

[87] Section 10(1)*(o)*(ii)*(aa)* and *(bb)* of the Income Tax Act.
[88] See par 2 above where this part of the definition of gross income is quoted.
[89] Section 10(1)*(o)*(ii)*(aa)* and *(bb)* of the Income Tax Act.
[90] Section 10(1)*(o)*(ii)*(aa)* and *(bb)* of the Income Tax Act.
[91] Section 10(1)*(o)*(ii)(A) of the Income Tax Act.
[92] Section 10(1)*(o)*(ii)(C) of the Income Tax Act.
[93] Section 10(1)*(o)*(ii)(B)(AA) of the Income Tax Act excludes income derived due to the holding of a public office in terms of s 9(2)*(g)*.
[94] Section 10(1)*(o)*(ii)(B)(BB) of the Income Tax Act excludes amounts that accrue in terms of s 9(2)*(h)*.
[95] Olivier and Honiball (n 11) 413.
[96] Olivier and Honiball (n 11) 415.
[97] Par 1 of the Fourth Schedule to the Income Tax Act: definition of "employee".

control by the employer over the employee.[98] This accords with *ITC 1174*[99] where the Tax Court found that independent contractors are not included in the meaning of the term employee.[100]

In *BMW South Africa (Pty) Ltd v The Commissioner for the South African Revenue Service*[101] the Supreme Court of Appeal found that amounts paid by an employer to tax advisors who advised their expatriate employee constituted a taxable benefit which should be included in the gross income of the expatriate employee.[102] This decision indicates that all possible benefits will be taxable in terms of South African domestic law. In spite of this, and unlike the wide exemption in respect of remuneration in the shipping industry, the scope of the exemption in respect of employees is still more limited, as, for example, dividends, certain pension payments, and payments received on the termination of services are not exempt.[103]

4.3 *The suggested amendment of the taxation of expatriate employees (from March 2020)*

On 22 February 2017, National Treasury announced that it would repeal the exemption from income tax in section 10(1)*(o)*(ii) of the Income Tax Act which currently applies to expatriate employees.[104] SARS stated that the reason for the repeal is that the current exemption creates opportunities for abuse, as expatriates could qualify for double non-taxation where foreign host countries do not levy income tax on employment income or taxes it at a reduced rate.[105] It is also possible to avoid tax through the manipulation of the days of absence of both jurisdictions. Initially, in 2017, it was proposed that the exemption in section 10(1)*(o)*(ii) of the Income Tax Act be repealed and that expatriate employees claim a tax credit for foreign taxes paid under section 6*quat* of the Income Tax Act.[106] This amendment, at the time, would have applied with effect from 1 March 2019 for years of assessment that commenced after this date.[107]

The final amendment kept this exemption in the Income Tax Act, yet changed it to only exempt the first R1 million earned outside SA from income tax for all years of assessment commencing from 1 March 2020.[108] All resident expatriate individuals must report and pay South African income tax on an amount earned

98 Olivier and Honiball (n 11) 415.
99 (1972) 34 SATC 135.
100 See n 99 above.
101 [2019] ZASCA 107 (6 September 2019).
102 *BMW South Africa (Pty) Ltd v The Commissioner for the South African Revenue Service* (n 101) par 26.
103 Section 10(1)*(o)*(ii)*(aa)* and *(bb)* of the Income Tax Act.
104 National Treasury Media Statement "Publication of the Draft Taxation Laws Amendment Bill and the 2017 Draft Tax Administration Laws Amendment Bill for public comment" (19 July 2017) 1.
105 National Treasury and SARS "Explanatory Memorandum on the Taxation Laws Amendment Bill 2017" (19 July 2017) 6.
106 National Treasury and SARS (n 105) 6–7.
107 *Ibid*.
108 Section 6(1)*(g)* of the Taxation Laws Amendment Act 17 of 2017 amends s 10(1)*(o)*(ii) of the Income Tax Act, with effect from 1 March 2020.

outside SA that exceeds R1million, if they were absent from the Republic for the requisite amount of days. This amendment only affects expatriate employees as the suggestion amends only section 10(1)*(o)*(ii) and not section 10(1)*(o)*(i), which grants an exemption to persons involved in the shipping industry.

4.4 *The unilateral exemption of government officials*

Under certain circumstances, non-resident employees who render services outside SA to an employer in the national or provincial sphere of government or a municipality, or any national or provincial public entity, are exempt from income tax.[109] This applies if "not less than 80% of the expenditure of such entity is defrayed directly or indirectly from funds voted by Parliament".[110] The income is only exempt in SA if it is chargeable to tax in the jurisdiction where the employee is ordinarily resident (not SA) and the employee personally paid tax there and this tax was not paid by the employer on behalf of the employee.[111]

5 The unilateral tax credit for foreign taxes paid

All other South African resident expatriate individuals, who are not employees or members of the shipping industry or government, may claim a tax credit for foreign taxes paid in another state when calculating their South African tax liability.[112] This rebate is available to residents who earn any type of income from any source outside SA,[113] a proportional amount attributed to an individual in terms of the controlled foreign company (CFC) rules,[114] a taxable capital gain from a source outside SA[115] or certain deemed accruals in terms of section 7,[116] certain capital gains from sources outside SA that are attributed to a resident,[117] and certain capital gains of trusts that are attributable to a resident.[118] Typically this could apply to an employed person who spends the required amount of time outside SA to qualify for section 10(1)*(o)* relief but spends the remainder of the tax year in SA or is ordinarily resident in SA.

[109] Section 10(1)*(p)* of the Income Tax Act.
[110] *Ibid*.
[111] *Ibid*.
[112] Section 6*quat* of the Income Tax Act.
[113] Section 6 *quat*(1)*(a)* of the Income Tax Act.
[114] Section 6*quat*(1)*(b)* of the Income Tax Act; see also s 9D and the rules that attribute income to SA residents in specific circumstances.
[115] Section 6*quat*(1)*(e)* of the Income Tax Act.
[116] Section 6*quat*(1)*(f)*(i) of the Income Tax Act.
[117] Section 6*quat*(1)*(f)*(ii) states that these amounts must be attributable in terms of par 68, 69, 70, 71, 72 or 80 of the Eighth Schedule to the Income Tax Act.
[118] Section 6*quat*(1)*(f)*(iii) of the Income Tax Act.

6 DTA and source-state relief for expatriate employees

Article 15 of the OECD Model Tax Convention, upon which many South Africa's DTAs are based, provides relief from international double taxation of employment income that is not earned by directors, sportspersons, artists and government officials. Both the UN and OECD Model Tax Conventions have other articles that contain specific rules for income earned by the latter categories of taxpayers.[119] In short, salaries, wages and remuneration are taxable in the state where the employment is exercised,[120] yet, this source state will not always have an unlimited right to levy income tax on expatriates. Taxation at source does not apply if the employee is present in the state where the employment is exercised for less than 183 days in a twelve-month period that commences and ends in accordance with the relevant tax year, or the remuneration was paid by an employer that is not resident in that state or not paid by a permanent establishment[121] of the employer in the state where the employment is exercised.[122] The physical presence of the employee is decisive when determining the number of days for purposes of the 183-day exception, and not the purpose of the presence.[123] Part days are included in the computation of the 183-day period[124] (unlike the SA domestic law, which requires 183 full days[125] for the section 10(1)*(o)*(ii) exemption to apply).[126] If the individual exercises employment on an aircraft or a ship, the remuneration is taxed in the state where the enterprise as employer has its place of effective management.[127]

Employment in terms of the OECD Model Tax Convention requires a contractual relationship as well as supervision and control by the employer.[128] Unlike the SA domestic law definition of employee, which includes directors of private companies, the OECD Model Tax Convention regulates the taxation of directors' fees separately in article 16, which allocates the taxing rights to the source state. Similarly, sportspersons, artists and certain of their support staff are taxable in the source state,[129] while pensions and annuities are generally taxable in the state of residence.[130] Specific, detailed rules apply to government employees, students and members of diplomatic services and consular services and the receipt of dividends, royalties or interest, which in most instances allocate taxing rights to the source state.[131] In many of these cases, domestic law imposes a withholding tax in the source state where the income is paid.

[119] Olivier and Honiball (n 11) 414–415.
[120] Article 15(1) of the OECD Model Tax Convention (2017).
[121] Olivier and Honiball (n 11) 419.
[122] Article 15(2) of the OECD Model Tax Convention (2017).
[123] Par 5 on article 15 of the OECD Model Tax Convention Commentaries (IBFD 2014).
[124] Par 5 on article 15 of the OECD Model Tax Convention Commentaries (IBFD 2014); days on which the employee is in transit or intended to leave and fell ill are excluded from this computation.
[125] Olivier and Honiball (n 11) 417.
[126] *Ibid*.
[127] Article 15(3) of the OECD Model Tax Convention (2017).
[128] Olivier and Honiball (n 11) 415.
[129] Article 17 of the OECD Model Tax Convention (2017).
[130] Article 18 of the OECD Model Tax Convention (2017).
[131] See articles 10, 11, 12, 19 and 20 of the OECD Model Tax Convention (2017).

ABLU 2019 (Part ONE).indb 160 2019/10/09 9:49 AM

Each specific treaty between SA and the state in which an expatriate exercises employment, as well as the domestic tax law of both (or more) states, should be scrutinised to determine the specific tax consequences for each individual taxpayer. It is clear that the nature of the position of the employee, the industry, the type of "income" received and time period spent abroad should be considered, prior to determining which specific article of a treaty (of which the OECD Model Tax Convention is an example) applies. This should be done after the country entitled to levy tax (based on either residence or domicile or another nexus besides source) on the individual was determined, as the application of the specific rules depends on the basic nexus which founds tax jurisdiction in each of the states involved.

7 The cessation of tax residence in South Africa

7.1 *Cessation of residence*

Although the definition of the term resident excludes individuals that are resident in another state in terms of a DTA, the timing of the exact moment or date when ordinary residency ends is not clearly formulated in legislation.[132] The proviso to the definition of the term "resident" in section 1 of the Income Tax Act states that

> "where any person that is a resident ceases to be a resident during a year of assessment, that person must be regarded as not being a resident from the day on which that person ceases to be resident …"

If an expatriate was resident based on the ordinary resident test, the taxpayer must prove his or her subjective intention not to return, or show that he or she is exclusively tax resident in another state.[133] SARS states that a person will cease to be a resident from the date that he or she emigrates.[134] This view, however, is not reflected in the Income Tax Act, which makes no mention of emigration as a determining factor for the cessation of tax residence. I am of the view, however, that, on a balance of probabilities, proof of formal emigration by a taxpayer will discharge the expatriate taxpayer of this evidentiary burden.

If an individual was classified as resident based on the physical presence test, the timing of the cessation of residence is easier to determine as the Income Tax Act clearly states that he or she is deemed not to be resident from the first day of physical absence.[135] If an individual is physically outside SA for a continuous time of more than 330 days directly after the day when he or she first ceases their physical presence in SA, that person is not classified as resident based on the physical presence test.[136] This non-resident status applies retrospectively from the first day that he or she stops being physically present in SA.[137] The counting of the days must be applied over a period of two years due to the fact that the person is

[132] Olivier and Honiball (n 11) 44.
[133] Section 102 of the Tax Administration Act 28 of 2011.
[134] SARS Interpretation Note 3 (n 43) 6.
[135] Section 1 of the Income Tax Act: definition of "resident" par *(a)*(ii)*(bb)*(B).
[136] See s 1 of the Income Tax Act: definition of "resident" par *(a)*(ii)*(bb)*(B).
[137] *Ibid*.

already resident based on physical presence of 91 days in the year that he or she first leaves SA.[138] There are not enough days to satisfy the requirements unless the next tax year is included in the calculation of the absent days.[139] This cessation of residency only applies if this person is not ordinarily resident based on the common law, and stops being classified as resident based on physical presence.[140]

For purposes of the application of a DTA, one could argue that the moment an individual becomes exclusively tax resident in another country would be the moment upon which his or her SA tax residency ceases. This moment would differ depending on the specific DTA's provisions, the tax legislation of the state where that expatriate is employed and the practices that are applied by the revenue authority in that jurisdiction. In my view, a certificate of tax residence of another country constitutes sufficient proof to end tax residence in SA, despite the fact that the expatriate may have an intention to return to SA. In this case, the onus of proof is discharged by the taxpayer and the exclusion applies, irrespective of the subjective intention of the taxpayer.

7.2 *Final tax liabilities upon cessation of residency*

Prior to his or her final exit out of SA, an expatriate is required to pay exit tax on all his or her capital assets.[141] The cessation of tax residence is a deemed disposal for purposes of capital gains tax.[142] It is deemed that an individual disposed of all his or her assets on the day before he or she ceases to be resident, at the market value of those assets as it is on the day before cessation of residence.[143] The departing person is also deemed to reacquire the assets at their market value on the day that residency ceases.[144] This fiction does not apply to immovable property situated in SA, assets that are attributable to a permanent establishment after the person ceases residency, qualifying equity shares in terms of a broad-based employee share scheme, equity instruments that vested in an employee, or a right to obtain a marketable security.[145] Although the timing of this "exit tax" liability is clearly indicated as the day before residence ceases, all the residency rules still need to be applied to determine this specific date. As the definition of resident excludes persons exclusively resident in another state in terms of a DTA, this creation of residency in the other state will trigger exit tax liability.

[138] Stiglingh (n 17) 31.
[139] *Ibid*.
[140] *Ibid*.
[141] Section 9H(2) of the Income Tax Act.
[142] *Ibid*.
[143] *Ibid*.
[144] *Ibid*.
[145] Section 9H(4)*(a)–(f)* of the Income Tax Act.

8 The formal emigration of individuals from South Africa

8.1 *The regulation of formal emigration from South Africa*

The right to leave SA and the right to freedom of movement is enshrined in the Constitution.[146] There is currently no separate legislation in SA that regulates the formal requirements for the emigration of individuals (save for certain exchange control regulations which regulate the export of cash, capital and assets from SA). Section 9 of the Immigration Act provides that no one shall depart from the Republic other than at a port of entry,[147] but contains no further requirements for such departure. The Immigration Act also states that the borders of SA must be monitored for illegal entry and exit of persons.[148] As is evident from the quote by Statistics SA in paragraph 1 of this chapter, this monitoring is limited and does not record the purpose of departure when individuals leave SA. Emigration remains a question of fact based on the specific circumstances of each individual. Myburgh states that emigration is deemed permanent when a person leaves SA with an intention to remain abroad for two years or longer.[149] Although the Income Tax Act contains no reference to emigration as a manner in which to end tax residency, the ending of residence based on the physical presence test is also determined over a two-year time period. If a person was, however, resident based on the principle of ordinary residence, the time period of their physical absence is not decisive.

In an attempt to address this lack of regulation, the White Paper on International Migration for South Africa of 2017 addressed certain policy gaps that exist in the management of emigration and stated that policy interventions were needed in this area to align the management of emigration with the achievement of SA's national development goals.[150] An analysis of this White Paper, however, reveals that South Africa lags behind many other African states in the amendment of policies to accommodate regional instruments signed in the SADC region.[151] This White Paper specifically states that it is the policy goal of government to manage their relations with individuals that have migrated so that their skills, capital and connections can be used to further the developmental priorities of SA.[152] The envisioned outcomes include an improved management of admission and departures from SA, the management of residency, and the management of ties with expatriates to attract them to invest in and return to SA.[153] Targeted persons include those who show an intention to emigrate, individuals that have emigrated and wish to invest in SA,

146 Section 21(1)–(2).

147 Act 13 of 2002; a port of entry is defined in s 1 of the Immigration Act as "a place designated by the Minister where all persons have to report before they may enter, sojourn or remain within or depart from the Republic".

148 Section 9 of the Immigration Act.

149 Myburgh (n 1) 122.

150 Department of Home Affairs "White Paper on International Migration for South Africa" (2017) iii and vi.

151 Department of Home Affairs (n 150) 18.

152 Department of Home Affairs (n 150) 33.

153 Department of Home Affairs (n 150) 34.

and persons who do live abroad and intend to return to SA.[154] It is proposed that the South African Passports and Travel Documents Act be amended to incorporate compulsory requirements for all expatriates to register on the Registration of South Africa (ROSA) system and that it become mandatory for all persons wishing to emigrate for more than three months to register on ROSA.[155] This register should be monitored by the Department of International Relations and Co-operation (DIRCO).[156] However, at the time of writing this chapter, no such amendment had been promulgated. It is stated that the skills and experience that expatriates gained abroad could improve the economy if they invest it in SA and that "South African citizens will benefit from the increase in economic growth and development resulting from effective emigration management".[157]

Besides mandatory registration and the deploying of additional staff to missions abroad to strengthen dysphoria centres,[158] there are no other practical suggestions to regulate emigration or to regulate the timing of formal emigration. On the contrary, the White Paper admits that the Department of Home Affairs has severe capacity constraints that make it impossible to fulfil its current mandate.[159] This remains an obvious impediment to the implementation of any new suggestions.

Emigration is, however, defined in this White Paper as "the act of departing or exiting from one's country (country of origin or of habitual residence) with a view to settling in another (host country)".[160]

8.2 *Financial emigration*

The term "financial emigration" is not used in legislation. Individuals who wish to formalise their status and emigrate from SA need to comply with the exchange control regulations and certain financial requirements in the regulations of the SARB.[161] The process involves several steps, the first of which is an application for foreign capital allowance through the completion of the MP336(b) form of the SARB, at a commercial bank of the emigrant's choice.[162] The emigrant is required to open a capital account with an authorised dealer (commercial bank) to which certain exchange control regulations apply.[163] The purpose of this account is to monitor all capital transfers abroad once a person's emigration status is recorded.[164] Once this account is opened the emigrant should apply for a Tax Clearance

154 Department of Home Affairs (n 150) 7 and 50.
155 Act 4 of 1994; Department of Home Affairs (n 150) 50.
156 Department of Home Affairs (n 150) 50.
157 Department of Home Affairs (n 150) 51.
158 Department of Home Affairs (n 150) 70.
159 The mandate of the DHA is described as "a critical enabler of economic development, security, service delivery and access to rights"; see Department of Home Affairs (n 150) 71.
160 Department of Home Affairs (n 150) 73.
161 South African Reserve Bank (SARB) "Emigrants" (https://www.resbank.co.za/RegulationAndSupervision/FinancialSurveillanceAndExchangeControl/FAQs/Pages/Emigrants.aspx (accessed 16 September 2019)) 1.
162 *Ibid*.
163 *Ibid*.
164 *Ibid*.

Certificate – Emigration from SARS, based on the MP336(b) form, a certified copy of which must be submitted to SARS.[165] Once both the Tax Clearance Certificate and the MP336(b) forms are received by the individual, it must be submitted to the commercial bank as authorised dealer.[166] After completion of the necessary formalities, the practical moving of assets must be done in accordance with several detailed rules, as prescribed in the exchange control regulations.

8.3 *The exchange control regulations*

If an expatriate decides to emigrate, he or she must comply with certain exchange control regulations issued by the Financial Surveillance Department of the SARB in terms of the Currency and Exchanges Act.[167] This short Act regulates issues related to banking, currency and exchanges, but does not contain detailed rules. The exchange control regulations are more detailed and, amongst other functions, play an important role to protect the tax base[168] and prevent capital flight from SA. There is no definition of the term "resident" in either the exchange control regulations or the Currency and Exchanges Act. There are, however, four categories of residents for purposes of the application of these exchange control regulations, each with its own restrictions that apply when a variety of assets and cash is moved out of SA.

9 Analysis

9.1 *Ordinary resident*

The *Cohen* decision, which is often quoted as authority for the statement that a real home is the guideline for tax residency, was delivered in 1946 at a time when international travel was not done with the ease that it is today. It was also not a trend to go abroad indefinitely for gap years, for working holidays, with no real future plan to return or to stay, or to study abroad and remain there indefinitely. It was thus easier to link a location to the notion of a real home. In modern societies, many professionals do not have real homes of the type envisaged in the famous quote of the *Cohen* case. It is possible to travel abroad frequently, to have homes in several states, to obtain dual citizenship, dual passports, property, as well as personal and financial links to several states. Modern technology, the global workplace and the increased trend towards migration, coupled with a highly effective internet communication system, have made anything possible. The traditional test for residency cannot be applied effectively in these circumstances. What was considered "ordinary" (to return to a "real home") in 1946 cannot be said to be "ordinary" in 2019, when the "normal" or "ordinary" manner of conducting oneself has changed dramatically. In 2019 it can be ordinary to live in many different places during one tax year, or even several tax years, without having one dominant or even two dominant locations to which a person is bound by intention

165 *Ibid*.
166 *Ibid*.
167 9 of 1933 s 9.
168 Olivier and Honiball (n 11) 71.

or fact. It does, however, remain a question of fact and the taxpayer bears the onus of proving where he or she is tax resident. The solution does not necessarily lie in creating more restrictive provisions in income tax legislation, as this might infringe on the constitutional right of persons to leave SA.

9.2 *Physical presence test*

The physical presence test is not without challenges as it is very difficult for SARS to trace a person's comings and goings to and from the Republic within the six-year period. A person with two passports could enter the country using one passport and leave using the other passport, thus leaving SARS in a difficult position to determine actual presence in South Africa.[169] As mentioned in paragraph 1 of this chapter, the purpose of an individual's travel abroad is not recorded by the Department of Home Affairs (DHA) at the ports of entry, resulting in a clear lack of information and statistics. This places SARS in a precarious position as it has to rely on the integrity of taxpayers and their voluntary compliance. Considering the instability at SARS recently, rumours of corruption and reports in the media that taxpayers are on the verge of a tax revolt,[170] the incentive for taxpayers to comply is very low and the potential for tax avoidance and/or evasion remains present. This situation is certainly not ideal, as SA is in desperate need of tax revenue to fund the activities of the state. A lack of accountable spending by government, high levels of corruption and a continued trend of increasing tax types and rates tend to cause reluctance on the side of taxpayers to comply voluntarily.

9.3 *The exclusion of expatriates resident in terms of the provisions of a DTA and the tie-breaker rules*

While the proviso to the resident definition does provide clarity in determining the residence of an individual working in two states, this treaty override afforded to a specific provision of a DTA still does not solve the situation where a taxpayer has no dominant location to which he or she can be tied. Similarly, the tie-breaker rules do not solve the multinational presence of individuals that work in more than two states.

9.4 *The exemption from income tax for expatriate employees in section 10(1)(*o*)(ii)*

The current exemption in section 10(1)*(o)*(ii) of the Income Tax Act as it applies to employees is quite simple as it contains a total exemption from income tax. After 1 March 2020, however, certain practical difficulties arise due to the fact that expatriate employees should now calculate the R1 million exempt amount and a

[169] See also Olivier and Honiball (n 11) 24 where the problem associated with dual passports of an individual is explained.

[170] Watson "The tax revolt is well under way, with expats already cutting their SARS ties" 9 September 2019 (https://citizen.co.za/business/2176615/the-tax-revolt-is-well-under-way-with-expats-already-cutting-their-sars-ties/(accessed 16 September 2019)).

tax credit that is applied to the amount above R1 million. For example, if a person is employed as a business analyst and earned the equivalent of R5 million abroad in a specific tax year, yet was absent from the Republic for more than 183 days and 60 continuous days, and a portion of R1 million constituted allowances such as accommodation, this amount is not necessarily taxable in the other jurisdiction where he or she lives. The domestic law of the other jurisdiction might differ from that of SA. Yet, the amount that is used to determine the exemption includes certain allowances in terms of section 8, one of which is accommodation. The net result for the taxpayer is that this amount is included for tax purposes in SA, but no section 6*quat* tax credit can be claimed to reduce the tax liability in SA as no foreign tax was paid.

Furthermore, it is not clear how the tax credit would be calculated if this person paid tax on the entire R5 million and wanted to claim the full amount of tax paid abroad. Logic dictates that a proportional credit paid in relation to the amount in excess of R1 million should be applied for in SA (in other words, only the foreign tax paid on the R4 million should be claimed) but there is currently no legislative authority for such an assumption. Practically, it could also be difficult for a taxpayer to obtain proof that foreign tax was paid in this proportional manner.

It is also not clear how the PAYE will be determined in advance by employers who agree to pay tax on behalf of their employees. PAYE requires an estimation of the tax liability in advance, whereas the claiming of the section 6*quat* tax credit is usually available ex post facto after proof of payment of foreign tax is provided to SARS. A further practical challenge lies in differing tax years or periods of assessment for different tax jurisdictions, causing the timing of tax liability not to correspond. It remains to be seen how SARS and taxpayers will deal with these obstacles. The decision in *BMW South Africa (Pty) Ltd v The Commissioner for the South African Revenue Service*[171] indicates that SARS will take all steps necessary to enforce the tax provisions that apply to expatriates.

In spite of these challenges, expatriate employees are still better off than self-employed expatriates, who will not qualify for the proposed R1 million exemption but can only claim the section 6*quat* credit. Expatriate employees who are still SA tax resident qualify for both the exemption of R1 million and the section 6*quat* credit on the amount above R1 million, leaving a smaller percentage of their earnings subject to SA income tax. Expatriates who find this unacceptable may end their tax residency to ensure that the South African Income Tax Act does not apply to them. This, as is illustrated above, requires more than merely working outside SA but entails the breaking of all ties with SA.

9.5 *Cessation of residence for tax purposes*

Three situations may arise to indicate an end in tax residency: first, when a person who was ordinarily resident in SA has the intention not to return to SA on a permanent basis (timing unclear); secondly, when an individual obtained

[171] See n 101 above.

tax residency by means of the physical presence test and such a person remains outside SA for 330 full days (timing clear); and thirdly, when a person is treated as exclusively resident in another contracting state (besides SA) in terms of a DTA. Financial emigration does not equal the cessation of tax residence. Formal emigration may, however, carry evidential weight, as the actions taken by the taxpayer are judged to determine if a distinct break was made with the previous tax jurisdiction. Expatriates should take note that physically leaving SA and moving assets (financial emigration) does not automatically absolve that individual from tax liability as the tie-breaker tests require the breaking of personal ties as well. Expatriates should be aware that if they do not sever all ties with SA formally (by, for example, closing bank accounts, selling primary residences, giving up club memberships and moving their families) they could still be regarded as tax resident in SA by SARS, despite working elsewhere permanently.

In my view, the definition of "resident" in section 1 of the Income Tax Act could, however, be amended to create a more specific moment in time when residency is regarded to cease. The current definition is somewhat circular as it refers to residence ceasing when it ceases. This calls for a factual analysis which forces expatriates to resort to expert advice for assistance to discharge their onus of proof, which can significantly increase their cost of compliance. One way of solving this is to define the cessation of residence with reference to emigration, and further define emigration in a manner which aligns it with the suggested definition in the White Paper on International Migration for South Africa of 2017. It remains to be seen whether the suggestions in the White Paper to regulate expatriates and enforce the registration of potential emigrants will be promulgated. This White Paper is certainly not without shortcomings. In my view, legislators should beware of infringing on the constitutional right of citizens to leave the Republic. It also seems very unrealistic for the government to expect expatriates or emigrants to contribute to the developmental goals of SA after they have left, if one considers the current socio-economic and political climate in SA.

9.6 *The exit tax liability upon cessation of residence*

The levying of exit tax is rather burdensome for the taxpayer as it could impair his or her ability to pay the tax in circumstances where the assets that are taxable based on this fiction are not realised. It is also quite harsh that a taxpayer who obtains exclusive residence in another state based on the proviso to the resident definition be subject to exit tax. Expatriates who work abroad for short periods of time without formally emigrating and giving up their South African homes or assets, but who are classified as non-resident based on the provisions of a tax treaty, are severely prejudiced by this tax liability.

It would be more appropriate to allow deferral of this tax liability in cases where the expatriate has not formally emigrated or is only abroad for a short period of time. The proviso to the resident definition and the provisions of section 9H of the Income Tax Act are not currently harmonised to regulate this issue in an equitable manner.

9.7 *The myth of "financial emigration"*

The assumption that "financial emigration", as was widely reported in the press and suggested to expatriates by certain tax and financial advisors, provides a solution to end South African tax residency or avoid tax liability in terms of the proposed section 10(1)*(o)*(ii) that will apply from 1 March 2020, is shockingly incorrect. This phrase is not used in tax legislation or the exchange control regulations of the South African Reserve Bank (SARB). This term does not exist in legislation. Yet, it was so widely used by legal and tax advisors that it almost became a household term used by concerned individuals who work abroad and considered their options when analysing their tax position. The term is nothing more than a marketing tool to entice taxpayers to obtain "expert" advice from persons who are inevitably driven by their own ulterior motives. It is a myth which has no application in tax law. The public outcry and frenzy in the press and the generalisation of the amendment as if it applies to all individuals working abroad is an unwarranted oversimplification. It is clear that only expatriate employees who remain tax resident and earn in excess of R1 million are affected by the suggested amendment of section 10(1)*(o)*(ii) of the Income Tax Act.

It remains the choice of the expatriate whether or not he or she would like to move assets and cash abroad. In cases where the expatriate is not resident due to the application of the provisions of a DTA, tax liability is determined in accordance with the normal resident rules and interaction of various sections in domestic law as well as articles in tax treaties. Financial emigration is not a legal requirement, nor does it affect the determination of the person's tax residency. Once an individual does formally emigrate, compliance with exchange control regulations, through selected commercial banks as authorised dealers, is required.

10 Conclusion

The loss of income tax contributions from individuals who leave SA remains concerning if one considers the fact that 900 000 people have left (conservatively speaking) and 4.9 million individuals contributed towards tax in 2017/2018, amounting to an approximate loss of one fifth of the individual tax base in SA. In my view, the amendment of section 10(1)*(o)*(ii) is justified as it provides slight protection against base erosion in respect of expatriates who split their time between SA and another jurisdiction, without breaking ties with SA. The practical difficulties of its application with regard to tax credits, fringe benefits, jurisdictions that SA does not have treaties with, and taxes that are not covered in the DTAs in respect of which expatriates wish to claim a tax credit remains to be seen. These aspects have created uncertainty and either require further legislative clarification or the publication of guidelines by National Treasury and SARS.

A balance between the right of the fiscus and those of the taxpayer remains present in the SA tax system as expatriates are still entitled to arrange their tax affairs to the best of their advantage so that they pay the least possible amount of tax. This can be achieved by simply applying the definition of the term resident correctly, which is the founding principle upon which the nexus to levy tax in SA is based.

Expatriate South Africans could, through effective tax planning, eliminate their South African tax liability entirely, within the parameters of the law. This requires a decision, evidenced by the correct documents, which shows that the expatriate broke all financial, residential and personal ties with South Africa, so that it is clear that SA is no longer the place to which he (or she) "as a matter of course returns from his wanderings".[172]

[172] *Cohen v CIR* (n 36) 185.

Table of Cases

Australia

National Australia Bank Ltd v Hokit Pty Ltd [1996] 39 NSWLR 377 [Booysen]

Varker v Commercial Banking Co of Sydney Ltd [1972] 2 NSWLR 967 [Booysen]

Canada

Arrow Transfer Co Ltd v Royal Bank of Canada [1972] 27 DLR (3d) 81 [Booysen]

Canadian Pacific Hotels Ltd v Bank of Montreal [1988] 40 DLR (4th) 385 [Booysen]

Canadian Pacific Hotels Ltd v Bank of Montreal [1987] 1 SCR 711 [Booysen]

Canadian Pacific Hotels v Bank of Montreal [1987] 40 DLR (4th) 385 [Booysen]

Columbia Gramophone Co v Union Bank of Canada [1916] 34 DLR 743 [Booysen]

Thompson v Minister of National Revenue 2 DTC 812 [Tredoux]

Hong Kong

Dex Asia Ltd v DBS Bank (Hong Kong) Ltd [2007] HKCA 516 [Booysen]

Tai Hing Cotton Mill Ltd v Liu Chong Hing Bank [1986] 1 AC 80 [Booysen]

Malaysia

Malaysia Plastics Sdn Bhd v United Overseas Bank (M) Bhd [2012] 9 MLJ 336 (High Court, Kuala Lumpur) [Booysen]

United Asian Bank Bhd v Tai Soon Heng Construction Sdn Bhd (United Asian) [1993] 1 MLJ 182 [Booysen]

Namibia

Bank of Namibia v Small & Medium Enterprises Bank Ltd (2) 2018 (1) NR 183 [Zongwe]

Bank of Namibia v Small & Medium Enterprises Bank Ltd (3) 2018 1 NR 193 (HC) [Zongwe]

New Zealand

National Bank of New Zealand Ltd v Walpole and Patterson Ltd [1975] 2 NZLR 7 [Booysen]

New Zealand v Auckland Information Bureau (Inc) [1996] 1 NZLR [Booysen]

ABLU 2019 (Part ONE).indb 171 2019/10/09 9:49 AM

Singapore

Hsu Ann Mei Amy v OCBC [2011] 2 SLR 178 [Booysen]
Jiang Ou v EFG Bank AG [2011] 4 SLR 246 [Booysen]
Khoo Tian Hock v Oversea-Chinese Banking Corporation Ltd [2000] 4 SLR 673 [Booysen]
Pertamina Energy Trading Limited v Credit Suisse [2006] 4 SLR(R) 273 [Booysen]
Tjoa Elis v United Overseas Bank [2003] 1 SLR(R) 747 [Booysen]

South Africa

BMW Financial Services (SA) (Pty) Ltd v National Credit Regulator; In Re BMW Financial Services v National Credit Regulator; In Re Volkswagen Financial Services (SA) (Pty) Ltd v National Credit Regulator [2018] ZANCT 49 (5 June 2018) [Van Heerden and Renke]
BMW South Africa (Pty) Ltd v The Commissioner for the South African Revenue Service [2019] ZASCA 107 (6 September 2019). [Tredoux]
CIR v Kuttel 1992 (3) SA 242 (A) [Tredoux]
Cohen v CIR 1946 AD 174 [Tredoux]
Commissioner South African Revenue Service v Van Kets 2012 (3) SA 399 (WCC). [Tredoux]
CSARS v Tradehold [2012] 74 SATC 263 [Tredoux]
Edcon Holdings Ltd v The National Consumer Tribunal unreported, GNP case no A237/2017 (23 May 2018) [Van Heerden and Renke]
Eskom Holdings v Halstead-Cleak 2017 (1) SA 333 (SCA) [Vessio]
Glenister v President of the Republic of South Africa and Others 2011 (3) SA 347 (CC) [Tredoux]
Miya v Miway NCT/43934/2016/75(1)(b) [Vessio]
National Credit Regulator v Edcon Holdings Ltd [2017] ZANCT 58 (24 April 2017) [Van Heerden and Renke]
National Credit Regulator v Lewis Stores (Pty) Ltd [2017] ZANCT 78 (25 May 2017) [Van Heerden and Renke]
Rossouw and Another v FirstRand Bank 2010 (6) SA 439 (SCA) [Van Heerden and Renke]
Secretary for Inland Revenue v Downing 1975 (4) SA 518 (A) [Tredoux]
Sipho Nchukana v African Compass Funeral Services Andcebisile Mfado case no FSOS 00238/17-18/WC 2 [Millard]
Volkswagen Financial Services v National Consumer Commission NCT/94937/2017/56(1) [Van Heerden and Renke]

United Kingdom

Barclays Bank Plc v Quince Care Limited [1992] 4 All ER 363 [Booysen]
Goldman v Cox (1924) 40 TLR 744 [Booysen]
Greenwood v Martins Bank [1933] AC 51 [Booysen]
Hall v Fuller (1826) 5 B & C 750; 108 ER 279 [Booysen]

Karak Rubber Co Ltd v Burden (No2) [1972] 1 WLR 602 [Booysen]
Kelly v Solari (1841) 9 M&W 54, 152 ER 24 [Booysen]
Kleinwort Benson Limited v Lincoln City Council [1999] 2 AC 349 [Booysen]
Lipkin Gorman v Karpnale & Co [1989] 1 WLR 1340 [Booysen]
London Joint Stock Bank Ltd v Macmillan & Arthur [1918] AC 777 [Booysen]
Marfani & Co Ltd v Midland Bank Ltd [1968] 1 WLR 956 [Booysen]
National Westminster Bank Ltd v Barclays Bank International Ltd [1975] 1 QB 654 [Booysen]
Selangor United Rubber Estates Ltd v Cradock (No3) [1968] 1 WLR 1555 [Booysen]
Shah v Barnet London Borough Council [1983] 1 All ER 226 (HL) [Tredoux]
Sheperd v Revenue and Customs Commissioners 2006 STC 1821 [Tredoux]
Smith v Lloyd's TSB Group plc [2001] QB 541 [Booysen]
Smith v Mercer (1815) 6 Taunt 76; 128 Eng Rep 961 [Booysen]
Tidal Energy Ltd v Bank of Scotland [2014] 2 Lloyd's Rep 549 [Booysen]
W J Simms Son & Cook (Southern) Limited [1980] 1 QB 677 [Booysen]

United States

Associated Home & RV Sales, Inc v Bank of Belen 2013-NMCA-018 [Booysen]
Critten v Chemical Nat'l Bank [1902] 171 NY 219 [Booysen]
Frank v Chemical National Bank of New York [1881] 84 NY 209 [Booysen]
Leather Manufacturers' Bank v Morgan [1886] 117 US 96 [Booysen]
Maryland Casualty Co v Central Trust Co 297 NY 294 [Booysen]
Potts & Co v Lafayette National Bank 269 NY 181 [Booysen]
Thomson v New York Trust Co 293 NY 58 [Booysen]

Legislation

Europe

European Union. Commission Delegated Regulation (EU) 2018/389 of 27 November 2017 supplementing Directive (EU) 2015/2366 of the European Parliament and of the Council [Booysen]

Hong Kong

Control of Exemption Clauses Ordinance [Booysen]
Ch 71 [Booysen]

Unfair Contract Terms Ordinance [Booysen]

Malaysia

Bills of Exchange Act 1949 [Booysen]
s 73A [Booysen]

Namibia

Bank of Namibia Act 15 of 1997 [Zongwe]
s 32 [Zongwe]

Singapore

Bills of Exchange Act (Chapter 23, 2004 Revised Edition) [Booysen]
s 86 [Booysen]

Unfair Contract Terms Act [Booysen]
Ch 396, 1994 Revised Edition [Booysen]

South Africa

Banks Act 94 of 1990 [Van Lill], [Vessio]
s 1: definition of "deposit" [Vessio], [Millard]

Broad-Based Black Economic Empowerment Act 53 of 2003 [Millard]
s 9(1) [Millard]

Collective Investment Schemes Control Act 45 of 2002 [Vessio], [Millard]

Competition Act 89 of 1998 [Van Heerden and Renke], [Vessio]
s 8(1)*(a)* [Van Heerden and Renke]

Constitution of the Republic of South Africa, 1996 [Vessio]
s 21(1)–(2) [Tredoux]
s 23 [Vessio]
s 231 [Tredoux]

ABLU 2019 (Part ONE).indb 175 2019/10/09 9:49 AM

Consumer Protection Act 68 of 2008 [Vessio]
- s 1: definition of "transaction" [Vessio]
- s 5(1)*(a)* [Vessio]
- s 5(1)*(b)* [Vessio]
- s 5(1)*(c)–(d)* [Vessio]
- s 5(2)*(b)* [Vessio]
- s 5(7) [Vessio]
- s 5(8) [Vessio]
- s 6(1) [Vessio]
- s 60 [Vessio]
- s 61 [Vessio]

Cooperative Banks Act 40 of 2007 [Van Lill], [Vessio]

Credit Agreements Act 75 of 1980 [Van Heerden and Renke]

Credit Rating Services Act 24 of 2012 [Vessio]

Currency and Exchanges Act 9 of 1933 [Tredoux]
- s 9 [Tredoux]

Financial Advisory and Intermediary Services Act 37 of 2002 [Vessio], [Millard]
- s 1(1): definition of "advice" [Millard]
- s 1(1): definition of "financial product" [Millard]
- s 1(1): definition of "intermediary services" [Millard]
- s 1(1) [Millard]

Financial Institutions (Protection of Funds) Act 28 of 2001 [Vessio]

Financial Intelligence Centre Act 38 of 2001 [Vessio]

Financial Markets Act 19 of 2012 [Vessio], [Millard]
- s 1(1): definition of "market infrastructure" [Vessio]
- s 1(1): definition of "securities" [Millard]

Financial Sector Regulation Act 9 of 2017 [Van Lill], [Van Heerden and Renke], [Vessio]
- Ch 1 [Vessio]
- Ch 2 [Vessio]
- Ch 3 [Vessio]
- Ch 4 [Vessio]
- Ch 5 [Vessio]
- Ch 8 [Vessio]
- Ch 14 [Vessio]
- Ch 15 [Vessio]
- Ch 17 [Vessio]
- s 1: definition of "financial customer" [Vessio]
- s 1: definition of "financial product" [Vessio], [Millard]
- s 1: definition of "financial sector law" [Vessio]
- s 1: definition of "financial system" [Vessio]
- s 1: definition of "regulatory instrument" [Vessio], [Millard]
- s 1: definition of "systemic event" [Van Lill]
- s 2: definition of "financial products" [Vessio]

s 3: definition of "financial services" [Vessio], [Millard]
s 4: definition of "financial stability" [Vessio]
s 5(2) [Vessio]
s 7 [Vessio]
s 9(1) [Vessio]
s 9(2) [Vessio]
s 10(1) [Vessio]
s 10(2) [Vessio]
s 13 [Vessio]
s 57 [Millard]
s 97–104 [Vessio]
s 175 [Vessio]
s 290 [Vessio]
Sch 1 [Vessio]
Sch 2 [Vessio]
Sch 4 [Vessio]

Financial Sector Regulations, 2018 (GN R405 in *GG* 41550 of 29 March 2018) [Vessio]

Financial Services Board Act 97 of 1990 [Vessio]

Financial Services Laws General Amendment Act 45 of 2013 [Vessio]

Financial Services Ombud Schemes Act 37 of 2004 [Vessio]

Financial Supervision of the Road Accident Fund Act 8 of 1993 [Vessio]

Friendly Societies Act 25 of 1956 [Vessio]

Immigration Act 13 of 2002 [Tredoux]
s 1: definition of "port of entry"
s 9 [Tredoux]

Income Tax Act 58 of 1962 [Tredoux]
s 1(1): definition of "gross income" [Tredoux]
s 1(1): definition of "income" [Tredoux]
s 1(1): definition of "Republic" [Tredoux]
s 1(1): definition of "resident" [Tredoux]
s 1(1): definition of "taxable income" [Tredoux]
s 5(1)*(c)* [Tredoux]
s 6*quat* [Tredoux]
s 8 [Tredoux]
s 9H(2) [Tredoux]
s 9H(4)*(a)–(f)* [Tredoux]
s 10 [Tredoux]
s 10(1)*(o)*(i) [Tredoux]
s 10(1)*(o)*(ii) [Tredoux]
s 11–19 [Tredoux]
s 20–20B [Tredoux]
s 21–21P [Tredoux]
s 26A [Tredoux]
s 108(2) [Tredoux]

Fourth Schedule
par 1: definition of “employee” [Tredoux]
par 1: definition of “remuneration” [Tredoux]
Eighth Schedule [Tredoux]
par 68 [Tredoux]
par 69 [Tredoux]
par 70 [Tredoux]
par 71 [Tredoux]
par 72 [Tredoux]
par 80 [Tredoux]

Inspection of Financial Institutions Act 80 of 1998 [Vessio]

Insurance Act 18 of 2017 [VESSIO]
s 1: definition of “non-life insurance policy” [Vessio]

Labour Relations Act 66 of 1995 [Vessio]
s 213 [Vessio]

Long-term Insurance Act 52 of 1998 [Vessio], [Millard]
s 1: definition of “long-term policy” [Vessio]
s 1A [Vessio]

Medical Schemes Act 131 of 1998 [Millard]

Mutual Banks Act 124 of 1993 [Van Lill], [Vessio]

National Credit Act 34 of 2005 [Van Heerden and Renke], [Vessio], [Millard]
Ch 5, part C [Van Heerden and Renke]
s 1: definition of “collection costs” [Van Heerden and Renke]
s 1: definition of “cost of credit insurance” [Van Heerden and Renke]
s 1: definition of “default administration charge” [Van Heerden and Renke]
s 1: definition of “initiation fee” [Van Heerden and Renke]
s 1: definition of “interest” [Van Heerden and Renke]
s 1: definition of “mortgage agreement” [Van Heerden and Renke]
s 1: definition of “service fee” [Van Heerden and Renke]
s 3*(e)* [Van Heerden and Renke]
s 8(4) [Van Heerden and Renke]
s 9(3)*(b)* [Van Heerden and Renke]
s 9(4)*(b)* [Van Heerden and Renke]
s 27*(a)*(ii) [Van Heerden and Renke]
s 56(2) [Van Heerden and Renke]
s 89(2)*(c)* [Van Heerden and Renke]
s 90 [Van Heerden and Renke]
s 90(2)2 [Van Heerden and Renke]
s 91(2)*(a)–(c)* [Van Heerden and Renke]
s 92 [Van Heerden and Renke]
s 92(2)*(a)* [Van Heerden and Renke]
s 92(2)*(b)* [Van Heerden and Renke]
s 93 [Van Heerden and Renke]
s 100(1)*(a)–(d)* [Van Heerden and Renke]
s 100 [Van Heerden and Renke]

s 101 [Van Heerden and Renke]
s 101(1)*(b)–(g)* [Van Heerden and Renke]
s 102 [Van Heerden and Renke]
s 102(2)*(a)–(c)* [Van Heerden and Renke]
s 103 [Van Heerden and Renke]
s 104 [Van Heerden and Renke]
s 105 [Van Heerden and Renke]
s 106 [Van Heerden and Renke]
s 121(3) [Van Heerden and Renke]
s 140 [Van Heerden and Renke]
s 150 [Van Heerden and Renke]
s 151 [Van Heerden and Renke]
s 160(1) [Van Heerden and Renke]
s 161 [Van Heerden and Renke]

National Credit Amendment Act 9 of 2014 [Van Heerden and Renke]
s 29 [Van Heerden and Renke]

National Credit Regulations, 2006 (GN R489 in *GG* 28864 of 31 May 2006) [Van Heerden and Renke]
Regulation 29 [Van Heerden and Renke]
Regulation 31 [Van Heerden and Renke]
Regulations 39–48 [Van Heerden and Renke]
Regulation 39(1): definition of "deferred amount" [Van Heerden and Renke]
Regulation 40 [Van Heerden and Renke]
Regulation 46 [Van Heerden and Renke]

National Payment System Act 78 of 1998 [Vessio]
s 6A [Vessio]

Pension Funds Act 24 of 1956 [Vessio], [Millard]
s 1: definition of "benefit" [Vessio]

Policy Board for Financial Services and Regulation Act 141 of 1993 [Vessio]

Regulations relating to the Levies on Financial Institutions (GN 384 in *GG* 42579 of 12 July 2019) [Vessio]

Share Blocks Control Act 59 of 1980 [Millard]
s 1: definition of "share block company" [Millard]

Short-term Insurance Act the 53 of 1998 [Vessio], [Millard]
s 1: definition of "short-term policy" [Vessio]

South African Passports and Travel Documents Act 4 of 1994 [Tredoux]

South African Reserve Bank Act 90 of 1989 [Zongwe], [Vessio]
s 10(1)*(f)* [Zongwe]
s 11 [Zongwe]
s 21 [Zongwe]
s 22 [Zongwe]
s 23 [Zongwe]
s 24 [Zongwe]

Tax Administration Act 28 of 2011 [Tredoux]
s 102 [Tredoux]
Taxation Laws Amendment Act 17 of 2017 [Tredoux]
s 6(1)*(g)* [Tredoux]
s 16(1)*(g)* [Tredoux]
Usury Act 68 of 1973 [Van Heerden and Renke]

South West Africa

Administration of Justice Proclamation 21 of 1919 [Zongwe]
s 1(1) [Zongwe]
Mandate for South West Africa, 1919 [Zongwe]
art 2 [Zongwe]
Proclamation 20 of 1919 [Zongwe]

United Kingdom

Bills of Exchange Act 1882 (BEA) [Booysen]
s 2 [Booysen]
s 8 [Booysen]
s 24 [Booysen]
s 64(1) [Booysen]
Cheques Act 1957 [Booysen]
s 4 [Booysen]
Consumer Credit Act 1974 [Booysen]
s 75 [Booysen]
Payment Services Regulations 2009 [Booysen]
Payment Services Regulations 2017 [Booysen]
reg 74(1) [Booysen]
Unfair Contract Terms Act 1977 [Booysen]

United States

Banking Act of 1933 [Zongwe]
ch 48 [Zongwe]
ch 89 [Zongwe]
Electronic Funds Transfer Act of 1978 [Booysen]
Truth in Lending Act of 1968 [Booysen]
Uniform Commercial Code (UCC) [Booysen]
Article 3 [Booysen]
§ 4-406(f) [Booysen]
Uniform Negotiable Instruments Act 1896 [Booysen]

Conventions and Treaties

Organisation for Economic Co-operation and Development (OECD) Model Tax Convention (2017) [Tredoux]
a 4(1) [Tredoux]
a 4(2) [Tredoux]
a 10 [Tredoux]
a 11 [Tredoux]
a 12 [Tredoux]
a 15(3) [Tredoux]
a 17 [Tredoux]
a 18 [Tredoux]
a 19 [Tredoux]
a 20 [Tredoux]

ABLU 2019 (Part ONE).indb 181 2019/10/09 9:49 AM